FOREST GENOMICS AND BIOTECHNOLOGY

EDITED BY: Isabel Allona, Matias Kirst, Wout Boerjan, Steven Strauss and Ronald Sederoff
PUBLISHED IN: Frontiers in Plant Science

frontiers Research Topics

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

FOREST GENOMICS AND BIOTECHNOLOGY

Topic Editors:

Isabel Allona, Universidad Politécnica de Madrid, Spain
Matias Kirst, University of Florida, United States
Wout Boerjan, Ghent University, Belgium
Steven Strauss, Oregon State University, United States
Ronald Sederoff, North Carolina State University, United States
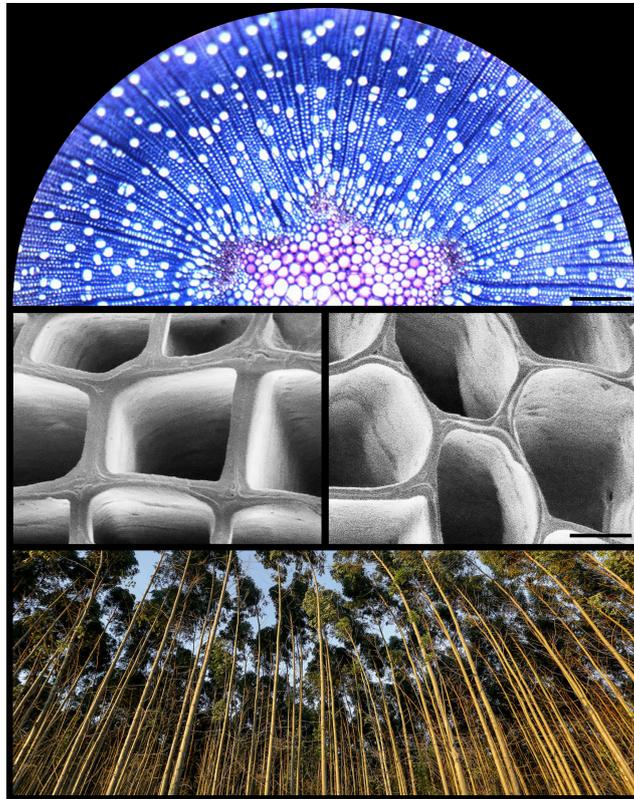
Top: Stem cross-section of three-month old *P. tremula x alba* genotype INRA 717-1B4 plantlets. Cell walls are stained with toluene blue and vessels are white. Black bar represents 1 mm. Picture taken by Cintia Leite Ribeiro from Matias Kirst's lab.

Middle: The SEM images are stem cross-sections of six-month-old *P. trichocarpa* solubilized by Caldicellusiruptor bescii for wildtype (left) and transgenics (right) downregulated in the expression of PtrCAD1 and PtrCAD2 (Straub et al., 2019; Nat. Commun. 10:2548). Black bar represents 5 μm. The SEM pictures were taken by Dr. Ilona Peszlen (NCSU)

Bottom: *E. grandis x urophylla* hybrid plantation, five-years old trees, VERACEL CELULOSE S.A. Municipality of Eunapolis, Southern Bahia, Brazil. Picture taken by Steve Strauss.

Image Design: Jack Wang and Ronald Sederoff

This Research Topic addresses research in genomics and biotechnology to improve the growth and quality of forest trees for wood, pulp, biorefineries and carbon capture.

Forests are the world's greatest repository of terrestrial biomass and biodiversity. Forests serve critical ecological services, supporting the preservation of fauna and flora, and water resources. Planted forests also offer a renewable source of timber, for pulp and paper production, and the biorefinery. Despite their fundamental role for society, thousands of hectares of forests are lost annually due to deforestation, pests, pathogens and urban development. As a consequence, there is an increasing need to develop trees that are more productive under lower inputs, while understanding how they adapt to the environment and respond to biotic and abiotic stress.

Forest genomics and biotechnology, disciplines that study the genetic composition of trees and the methods required to modify them, began over a quarter of a century ago with the development of the first genetic maps and establishment of early methods of genetic transformation. Since then, genomics and biotechnology have impacted all research areas of forestry. Genome analyses of tree populations have uncovered genes involved in adaptation and response to biotic and abiotic stress. Genes that regulate growth and development have been identified, and in many cases their mechanisms of action have been described. Genetic transformation is now widely used to understand the roles of genes and to develop germplasm that is more suitable for commercial tree plantations. However, in contrast to many annual crops that have benefited from centuries of domestication and extensive genomic and biotechnology research, in forestry the field is still in its infancy. Thus, tremendous opportunities remain unexplored.

This Research Topic aims to briefly summarize recent findings, to discuss long-term goals and to think ahead about future developments and how this can be applied to improve growth and quality of forest trees.

Table of Contents

5 BIOTIC AND ABIOTIC STRESS

6 CYBERINFRASTRUCTURE

frontiers
in Plant Science

Editorial: Forest Genomics and Biotechnology

Isabel Allona[1,2]*[†], Matias Kirst[3], Wout Boerjan[4,5], Steven Strauss[6]
and Ronald Sederoff[7]*[†]

[1] Centro de Biotecnología y Genómica de Plantas (CBGP, UPM-INIA), Universidad Politécnica de Madrid (UPM) - Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Campus de Montegancedo, Pozuelo de Alarcón, Madrid, Spain, [2] Departamento de Biotecnología-Biología Vegetal, Escuela Técnica Superior de Ingeniería Agronómica, Alimentaria y de Biosistemas, Universidad Politécnica de Madrid (UPM), Madrid, Spain, [3] School of Forest Resources and Conservation, University of Florida, Gainesville, FL, United States, [4] Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent, Belgium, [5] Center for Plant Systems Biology, VIB, Ghent, Belgium, [6] Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR, United States, [7] Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC, United States

Editorial on the Research Topic

Forest Genomics and Biotechnology

Forest biotechnology can be said to have begun with the construction of the first transgenic tree. A bacterial gene imparting glyphosate resistance (EPSP synthase) was introduced into a hybrid poplar (Fillatti et al., 1987). This achievement required the development of three technologies: gene discovery, gene transfer, and *in vitro* plant regeneration. These powerful tools advanced the investigation of the unique biology of forest trees, including their unusual reproductive features, woody and perennial growth habit, and their mechanisms of adaptation to abiotic and biotic stress. In addition, biotechnology has advanced the practical application of molecular genetics for tree breeding, by expanding options for the production of pulp, wood, and energy products (Huang et al., 1993; Baucher et al., 2010; Allwright and Taylor, 2016).

While biotechnology established the tools necessary to modify genes, genomics provided a new platform for the high-throughput genetic analysis of forest trees. Genomics was founded on two technologies: genetic mapping and DNA sequencing. Genetic mapping provided the location of genes, allowing the association of their position to function. Previously, genetic maps were made using isozyme loci as markers (Conkle, 1980) but the number of loci that could be sampled was limited. Restriction fragment length polymorphisms (RFLP), PCR amplification and high throughput DNA sequencing led to an extraordinary expansion in the regions of the genome that could be surveyed (Botstein et al., 1980; Williams et al., 1990; Vos et al., 1995; Wang et al., 1998). Consequently, genetic maps can now be readily constructed for forest trees where genetic analysis of quantitative traits had been previously impossible (Kirst et al., 2004). Through genetic mapping and segregation analysis, a trait could be readily shown to be monogenic, oligogenic, or polygenic in genetic architecture.

Large scale genome mapping of forest trees was first carried out using haploid genetic analysis using conifer megagametophytes and PCR-based anonymous markers (Carlson et al., 1991; Grattapaglia et al., 1992; O'Malley et al., 1996). The concept of mapping with anonymous dominant or codominant markers was extended to diploid crosses using "pseudotestcross" strategies (Grattapaglia and Sederoff, 1994). As sequencing technology advanced, anonymous markers were replaced by sequence-based markers associated with genes, single nucleotide polymorphisms (SNPs) (Eckert et al., 2009) or variation in repeated sequences (Echt et al., 2011).

Forest biotechnology has been strongly influenced by the human genome project (HGP) (Venter et al., 2001). The new technology of the HGP advanced the studies of many species and was important for species such as forest trees, which had been previously difficult due to large sizes and long generation times. Genome sequencing of forest trees (Tuskan et al., 2006) has brought about a host of new technologies ("omics") where gene expression and function could be studied for single genes or for large gene families and even for all of the expressed genes in a specific tissue or cell type. Omics have been developed in tree species for populations of RNA molecules (transcriptomics), proteins (proteomics), and metabolites (metabolomics). All of these methods have been or are being applied to forest tree species (Wagner et al., 2012; Wang et al., 2014; Wang et al., 2018).

While the new "omics" technologies characterize the activity of biological systems, the understanding about the relationship among them is complex and require mathematical modeling to provide predictive power (Wang et al., 2014). Systems and synthetic biology of forest trees reflect the growing collaboration of engineering and molecular biology. Systems biology integrates levels of information and technology deriving from genomics, such as genome sequencing, transcriptomics, epigenomics, proteomics, metabolomics, and imaging. While systems biology uses information from existing species, synthetic biology goes beyond what already exists in nature by reassembling or inventing novel gene sequences and functions. The goals of systems and synthetic biology are to improve the efficiency of metabolic flux, to redesign pathways, or to create novel ones.

A thorough understanding of metabolic and developmental pathways may not only provide innovations in systems design, but could ultimately prove critical for survival of natural and planted forests. Forests around the world continue to be threatened by an increase in the introduction of nonnative pests and pathogens due to world trade and travel. Epidemics of native species of pests and pathogens have also increased due to the destabilizing effects of climate change, which imposes increased abiotic stress on tree populations. Many species of forest trees may soon be lost, affecting entire ecosystems and leading to loss in ecosystem services and biodiversity (NASEM, 2019). Biotechnology could increase our understanding of host pathogen interactions and could also aid in the development of new genotypes able to resist new biotic and abiotic stresses. This will require major new allocation of resources to studies of forest tree biology and the genetic modification of forest tree populations require major changes in highly restrictive regulations and preclusion from markets (Strauss et al., 2015)

This volume is organized in six sections. (1) Gene Discovery, (2) Gene Transfer and Genetic Engineering, (3) Genetic Mapping and Quantitative Trait Analysis, (4) Growth and Development, (5) Biotic and Abiotic Stress, and (6) Cyberinfrastructure.

(1) Gene Discovery. Great attention has been paid to discovery and manipulation of genes involved in wood formation. Motivated by industrial activity in pulp and paper processing, and by the use of wood for biofuels and solid wood products, it has become critical to identify and characterize genes involved in the chemical and physical properties of wood. Wang et al. (in this volume) have reviewed the status of protein–protein interactions in lignin precursor biosynthesis, and Chanoca et al. (in this volume) have reviewed the efforts made to reduce biomass recalcitrance by engineering lignin quantity and composition. Engineering of noncellulosic polysaccharides in wood is discussed by Donev et al. (in this volume), with the goal of modifying wood for chemical content (particularly sugars) and physical properties, which affect the interactions of hemicelluloses with lignin and cellulose. Many forest trees accumulate high levels of secondary metabolites for defense against pests and pathogens. Conifers and other trees such as eucalypts accumulate terpenes in wood that can be extracted and utilized as a renewable chemical feedstock. Peter (in this volume) has reviewed the status of genetic engineering and breeding approaches to increase the abundance of terpenes and thereby increase the value of plantation forest trees. Myburg et al. (in this volume) integrated systems biology, systems genetics, and synthetic biology to propose a new paradigm for the production of chemical feedstocks from woody biomass and for a multitude of other wood products.

(2) Gene Transfer and Genetic Engineering. The slow and inefficient transfer, incorporation, and expression of specific genes into forest trees continues to be a major barrier to progress. Even more demanding is the subsequent requirement for the transformed cell to dedifferentiate, divide, and regenerate organs or embryos expressing the inserted gene. Only a small number of tree genotypes and species have *in vitro* regeneration systems able to support the stress of DNA insertion and express the plasticity needed for embryogenesis or organogenesis. Nagle et al. (in this volume) have reviewed the challenges and opportunities existing for DNA transformation in forest trees. Use of development-stimulating genes such as WUSCHEL appear highly promising, as do *in vivo* approaches. Bewg et al. (in this volume) summarized the status of gene editing in trees using CRISPR-cas9 technology. CRISPR technology is highly efficient for making knock-outs and other alterations, with few off-target effects in trees. Sterility has been the strategy to mitigate gene flow from future plantations that contain genetically modified trees. The minireview by Fritsche et al. (in this volume) outline approaches to the containment of modified genes, and of exotic and invasive forest tree species, by sterility.

(3) Genetic Mapping and Quantitative Trait Analysis. One major benefit of genomics has been the resulting integration of quantitative and molecular genetics. Now that genomic sequencing can, in principle, identify all the genes, the genetic basis of complex quantitative traits can be potentially identified and characterized more directly. Wood properties are under moderate to strong genetic control (Porth et al., 2012) but are also influenced by environmental factors such as season, rainfall, and variation of the gravity vector (Plomion et al., 2001). Du et al. (in this volume) use wood formation as a model system for investigation of the genetic architecture and regulatory mechanisms of quantitative traits

in forest trees. They reviewed recent progress in genome-wide association studies (GWAS) of wood properties as a tool for functional genomics and the potential for molecular breeding. Grattapaglia et al. (in this volume) reviewed the application of genomics to tree breeding and describe how genomic information may be used to improve selection. Genomic selection (GS) may allow predictive markers to accelerate the selection of elite genotypes and discovery of the genetic factors contributing to quantitative traits.

(4) Growth and Development. The growth habit and architecture of forest trees has a major effect on the ecological roles and commercial value of forest trees (Zobel and Jett, 1995; Holliday et al., 2017). Busov (in this volume) has reviewed the regulation of crown architecture, secondary growth, wood formation, and adventitious rooting, all complex traits, based on a number of molecular mechanisms.

Chromatin modification is thought to be the basis for a wide spectrum of changes in gene expression in response to developmental or environmental signaling. Only recently have the extent and roles of chromatin remodeling in forest trees been explored. In this volume, Conde et al. have reviewed this progress focusing on the signatures of chromatin regulation during active growth and seasonal dormancy.

Molecular switches, sensitive to day length and temperature, drive poplar phenology. Short days and low temperatures trigger the sequential induction of ethylene and abscisic acid signaling pathways, bud maturation and the establishment of dormancy. Transcriptional profiling and genetic association studies in poplar by Maurya et al. (in this volume) describe a platform for studying how the environment affects the molecular switches that drive phenology.

A major distinction in forest trees is based on the independent origin of gymnosperms (softwoods) and woody angiosperms (hardwoods). Tuskan et al. (in this volume) have described how, as genome sequence information increases and gene function is better understood, much will be learned about the evolution of these very different lineages with respect to their adaptation to variable environments.

(5) Biotic and Abiotic Stress. The world's forest tree species are under threat from the globalization of pests and pathogens. At the same time, tree species are increasingly susceptible to these threats due to the increased stress from global climate change. Increased efforts are being made to find or create genetic resistance, either through breeding or genetic engineering. Naidoo et al. (in this volume) have proposed a strategy for the investigation of defense mechanisms that could lead to the development of superior genotypes with enhanced resistance to biotic stress, integrating quantitative and qualitative resistance with the additional contributions of microbial endophytes and the root-associated microbiomes.

Cold hardiness can affect the natural range of forest trees or the establishment of exotic species in new environments (Hinchee et al., 2011). Wisniewski et al. (in this volume) have reviewed the current knowledge of cold hardiness, and the efforts to improve it through transgenic approaches. Cold hardiness is a complex trait involving avoidance, tolerance, seasonal stages, and dormancy.

Nitrogen availability, as ammonia or nitrate, and nitrogen use efficiency are limiting factors for growth and development of forest trees. Nitrogen reserves also affect dormancy and nitrogen cycling in their ecosystem. Nitrogen affects both primary and secondary metabolism. Canovas et al. (in this volume) summarize advances in forest trees for the functional characterization of genes affecting molecular regulation of acquisition, assimilation, and internal recycling of nitrogen.

One of the major effects of climate change is manifested in the water cycles where droughts are expected to be more frequent and more extreme. Polle et al. (in this volume) review approaches to identify genes that could modify drought tolerance through knowledge of the molecular physiology of the responses to drought stress.

(6) Cyberinfrastructure. Databases have become essential to forest biotechnology, as genomic analysis, transcriptomics, metabolomics, and image analysis become accessible tools for genetic engineering and systems biology of forest trees. Wegrzyn et al. (in this volume) describe the existing individual databases, each focus of interest and how they interact to provide synergistic cyberinfrastructure for the forest tree biotechnology community.

Forest genomics and biotechnology is a highly diverse international endeavor, which is advancing rapidly as new technology enables novel approaches and insights. It is an exciting time, and the reviews in this volume provide an excellent update about where things are, and where they are likely to go. Enjoy reading this special issue on Forest Genomics and Biotechnology!

AUTHOR CONTRIBUTIONS

All authors participated either in the writing or editing of the editorial.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Allwright, M. R., and Taylor, G. (2016). Molecular breeding for improved second generation bioenergy crops. *Trends Plant Sci.* 21 (1), 43–54. doi: 10.1016/j.tplants.2015.10.002

Baucher, M., Monties, B., Van Montagu, M., and Boerjan, W. (2010). Biosynthesis and genetic engineering of lignin. *Crit. Rev. Plant Sci.* 17, 125–197. doi: 10.1080/07352689891304203

Botstein, D., White, R. L., Skolnik, M. H., and Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Amer. J. Hum. Genet.* 32, 314–331.

Carlson, J. E., Tulsieram, L. K., Glaubitz, J. C., Luc, V. W. K., Kauffeldt, C., and Rutledge, R. (1991). Segregation of random amplified DNA markers in F1 progeny of conifers. *Theoret. Appl. Genet.* 83, 194. doi: 10.1007/BF00226251

Conkle, M. T. (1980). "Amount and distribution of isozyme variation in various conifer species," in *Proceedings 17th Meeting of the Canadian Tree Improvement Association*, ed. M.A.K. Khalil (Canada: Canadian Forest Service, Environment) 109–117. https://www.fs.fed.us/psw/publications/conkle/psw_1980_conkle001.pdf https://www.fs.usda.gov/treesearch/pubs/32855

Echt, C. S., Saha, S., Krutovsky, K. V., Wimalanathan, K., Erpelding, J. E., Liang, C., et al. (2011). An annotated genetic map of loblolly pine based on microsatellite and cDNA markers. *BMC Genet.* 12, 17. doi: 10.1186/1471-2156-12-17

Eckert, A. J., Pande, B., Ersoz, E. S., Wright, M. H., Rashbrook, V. K., Nicolet, C. M., et al. (2009). High-throughput genotyping and mapping of single nucleotide polymorphisms in loblolly pine (*Pinus taeda* L.). *Tree Genet. Genomes* 5, 225–234. doi: 10.1007/s11295-008-0183-8

Fillatti, J. J., Selmer, J., McCown, B., Hassig, B., and Comai, L. (1987). Agrobacterium mediated transformation and regeneration of Populus. *Mol. Gen. Genet.* 206, 192–199. doi: 10.1007/BF00333574

Grattapaglia, D., Chaparro, J. X., Wilcox, P., McCord, S., Werner, D., Amerson, H., et al. (1992). "Mapping in woody plants with RAPD markers: applications to breeding in forestry and horticulture," in *Proceedings of the Symposium on Applications of RAPD Technology to Plant Breeding* (CSSA/ASHS/AGA), 37–40.

Grattapaglia, D., and Sederoff, R. (1994). Genetic linkage maps of Eucalyptus grandis and Eucalyptus urophylla using a pseudo-testcross: mapping strategy and RAPD markers. *Genetics* 137 (4), 1121–1137.

Hinchee, M., Zhang, C., Chang, S., Cunningham, M., Hammond, W., and Nehra, N. (2011). Biotech Eucalyptus can sustainably address society's need for wood: the example of freeze tolerant Eucalyptus in the southeastern U.S. IUFRO Tree Biotechnology Conference 2011: From Genomes to Integration and Delivery. *BMC Proceedings* vol.5 (Suppl 7):A1. doi: 10.1186/1753-6561-5-S7-I24

Holliday, J. A., Aitken, S. N., Cooke, J. E. K., Fady, B., González-Martínez, S. C., Heuertz, M., et al. (2017). Advances in ecological genomics in forest trees and applications to genetic resources conservation and breeding. *Mol. Ecol.* 26 (3) 706–717 doi: 10.1111/mec.13963

Huang, H., Karnofski, D. F., and Tauer, C. G. (1993). Applications of biotechnology and molecular genetics to tree improvement. *J. Arboriculture* 19 (2), 84–98.

Kirst, M., Myburg, A., and Sederoff, R. (2004). "Genetic mapping in forest trees: markers, linkage analysis and genomics," in *Genetic Engineering: Principles and Methods.*, vol. 26 . Ed. J. K. Setlow (Boston, MA: Springer). doi: 10.1007/978-0-306-48573-2_7

NASEM. The National Academies of Science, Engineering and Medicine (2019). "Forest health and biotechnology: possibilities and considerations," in *Forest Health and Biotechnology: possibilities and considerations* (Washington, DC: The National Academies Press). doi: 10.17226/25221

O'Malley, D. M., Grattapaglia, D., Chaparro, J., Wilcox, P., Amerson, H. V., Liu, B.-H., et al. (1996). "Molecular markers, forest genetics and tree breeding, 21st *Stadler Genetics Symposium*: University of Missouri–Columbia)," in *Genomes of plants and animals*. Eds. J.P. Gustafson and R.B. Flavell. (NY: Plenum Press) doi: 10.1007/978-1-4899-0280-1_7

Plomion, C., Leprovost, G., and Stokes, A. (2001). Wood formation in trees. *Plant Physiol.* 127, 1513–1523. doi: 10.1104/pp.010816

Porth, I., Klapste, J., Skyba, O., Lai, B. S. K., Geraldes, A., Muchero, W., et al. (2012). *Populus trichocarpa* cell wall chemistry, and ultrastructure trait variation, genetic control and genetic correlations. *New Phytol.* 197, 777–790. doi: 10.1111/nph.12014

Strauss, S., Costanza, A. and Séguin, A. (2015). Genetically engineered trees: paralysis from good intentions. Science 349 (6250), 794–795. doi: 10.1126/science.aab0493

Tuskan, G. A., DiFazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313 (5793), 1596–1604. doi: 10.1126/science.1128691

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., et al. (2001). The sequence of the human genome. *Science* 291 (5507), 1304–1351. doi: 10.1126/science.1058040

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., and Hornes, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* 23 (21), 4407–4414. doi: 10.1093/nar/23.21.4407

Wagner, A., Donaldson, L., and Ralph, J. (2012). "Lignification and lignin manipulations in conifers," in *Advances in Botanical Research*, vol. 61 . Eds. L. Jouanin and C. Lapierre (Burlington: Academic Press), 37–76. ISBN:978-0-12-416023-1. doi: 10.1016/B978-0-12-416023-1.00002-1

Wang, D. G., Fan, J. B., Siao, C. J., Berno, A., Young, P., Sapolsky, R., et al. (1998). Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science* 280 (5366), 1077–1082. doi: 10.1126/science.280.5366.1077

Wang, J. P., Naik, P. P., Chen, H.-C., Shi, R., Lin, C.-Y., and Liu, J. (2014). Complete proteomic-based enzyme reaction and inhibition kinetics reveal how monolignol biosynthetic enzyme families affect metabolic flux and lignin in *Populus trichocarpa*. *The Plant Cell* 26, 894–914. doi: 10.1105/tpc.113.120881

Wang, J. P., Matthews, M. L., Williams, C. M., Shi, R., Yang, C., Tunlaya-Anukit, S., et al. (2018). Improving wood properties for wood utilization through multi-omics integration in lignin biosynthesis. *Nat. Commun.* 9, 1579. doi: 10.1038/s41467-018-03863-z

Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A., and Tingey, S. V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acid Res.* 18, 6531–6535. doi: 10.1093/nar/18.22.6531

Zobel, B. J., and Jett, J. B. (1995). "Genetics of wood production," in *Springer Series in Wood Science*. Ed. T. E. Timmel (Berlin, Heidelberg: Springer-Verlag) 337. doi: 10.1007/978-3-642-79514-5

Enzyme-Enzyme Interactions in Monolignol Biosynthesis

Jack P. Wang[1,2], Baoguang Liu[2,3], Yi Sun[2], Vincent L. Chiang[1,2] and Ronald R. Sederoff[1]*

[1]*Forest Biotechnology Group, Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC, United States,* [2]*State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University, Harbin, China,* [3]*Department of Forestry, Beihua University, Jilin, China*

The enzymes that comprise the monolignol biosynthetic pathway have been studied intensively for more than half a century. A major interest has been the role of pathway in the biosynthesis of lignin and the role of lignin in the formation of wood. The pathway has been typically conceived as linear steps that convert phenylalanine into three major monolignols or as a network of enzymes in a metabolic grid. Potential interactions of enzymes have been investigated to test models of metabolic channeling or for higher order interactions. Evidence for enzymatic or physical interactions has been fragmentary and limited to a few enzymes studied in different species. Only recently the entire pathway has been studied comprehensively in any single plant species. Support for interactions comes from new studies of enzyme activity, co-immunoprecipitation, chemical crosslinking, bimolecular fluorescence complementation, yeast 2-hybrid functional screening, and cell type–specific gene expression based on light amplification by stimulated emission of radiation capture microdissection. The most extensive experiments have been done on differentiating xylem of *Populus trichocarpa*, where genomic, biochemical, chemical, and cellular experiments have been carried out. Interactions affect the rate, direction, and specificity of both 3 and 4-hydroxylation in the monolignol biosynthetic pathway. Three monolignol P450 monooxygenases form heterodimeric and heterotetrameric protein complexes that activate specific hydroxylation of cinnamic acid derivatives. Other interactions include regulatory kinetic control of 4-coumarate CoA ligases through subunit specificity and interactions between a cinnamyl alcohol dehydrogenase and a cinnamoyl-CoA reductase. Monolignol enzyme interactions with other pathway proteins have been associated with biotic and abiotic stress response. Evidence challenging or supporting metabolic channeling in this pathway will be discussed.

Keywords: monolignol biosynthesis, protein-protein interaction, lignin, enzyme kinetics, BiFC, co-immunoprecipitation

INTRODUCTION

Metabolic pathways are typically conceived as sequences of enzymatic events that are linear, branched, or sometimes circular, that is, how they are described in charts and textbooks. It is a standard procedure in biochemistry to study enzyme reactions in highly dilute aqueous solutions, that is, what is needed to isolate and characterize specific enzymes. As a result, the

roles of structure and compartmentation may not be adequately considered. We often assume that each enzyme acts independently, in part because we must purify each enzyme to learn its essential properties, unconfounded by activity of other enzymes. In biochemical reality, enzymes and pathways occur in three-dimensional space and are likely to be associated with other proteins, polymers, membranes, and intracellular structures.

Much of our genetic theory of selection is based on concepts of genes acting independently, at least in part because it provides a computational simplification, although interactions of genes and proteins abound. Genomics provides a new platform for the investigation of molecular interactions and structure because all members of gene and protein families can now be identified and characterized, so that assays for

molecular interactions can now be investigated in a far more comprehensive way. We can investigate the role of structure for pathways, beginning from the investigation of protein-protein interactions using new tools to detect genome wide enzymatic and physical interactions.

The lignin biosynthetic pathway in vascular plants (**Figure 1**) has been intensively studied for the greater part of a century (Payen, 1839) because of its fundamental role in the formation of the plant secondary cell wall (Sarkanen and Ludwig, 1971; Harada and Côté, 1985; Boerjan et al., 2003; Weng and Chapple, 2010) and for its practical implications in wood formation (Sarkanen and Ludwig, 1971; Hu et al., 1999; Chiang, 2002; Rastogi and Dwivedi, 2006; Chen and Dixon, 2007; Vanholme et al., 2008; Novaes et al., 2010; Wang et al., 2014; Wang

FIGURE 1 | Interacting enzyme in the monolignol biosynthetic pathway. Colored boxes represent the interacting enzyme families. Red lines represent interactions indicated in the literature that involve proteins outside of the core monolignol pathway. Abbreviations: phenylalanine ammonia-lyase (PAL); cinnamate 4-hydroxylase (C4H); coumarate 3-hydroxylase (C3H); 4-coumarate: coenzyme A ligase (4CL); hydroxycinnamoyl-CoA shikimate hydroxycinnamoyl transferase (HCT); caffeoyl shikimate esterase (CSE); caffeoyl coenzyme A O-methyltransferase (CCoAOMT); cinnamoyl-CoA reductase (CCR); coniferaldehyde 5-hydroxylase (CAld5H); 5-hydroxyconiferaldehyde O-methyltransferase (COMT); cinnamyl alcohol dehydrogenase (CAD); membrane steroid-binding protein (MSBP); cytochrome P450 reductase (CPR); isoflavone synthase (CYP93C1); plant disease resistance gene 1 (Rp1); Rac family small GTPase 1 (Rac1).

et al., 2018). The lignin biosynthetic pathway begins in the cytoplasm by deaminating phenylalanine and proceeds stepwise through the modification of the phenyl ring and the reduction of the propanoid side chain to produce the monolignol precursors, which are translocated to the apoplast to form the lignin polymer (**Figure 1**) (Higuchi, 1985; Boerjan et al., 2003; Wang et al., 2014). While each step of the pathway has reversible reactions, which can be described by Michaelis-Menten kinetics (Wang et al., 2014), the formation of the lignin polymer by oxidative polymerization is not reversible.

The lignin pathway has been variously described as linear, branched, or metabolic grid. The pathway is different in different taxa, related to the distinction of hardwoods (angiosperms) that make a mixed polymer from guaiacyl and syringyl monolignols, compared to softwoods (gymnosperms) that make a purely guaiacyl lignin (Mottiar et al., 2016). Different cell types have different lignin compositions (Fergus and Goring, 1970a,b; Zhou et al., 2011; Shi et al., 2017). Vessel cells in poplar xylem have a guaiacyl rich lignin, while adjacent fiber cells have a syringyl rich lignin content. Studies typically have not been comprehensive in which all the enzymes in the pathway have not been studied in the same cell types, and comparisons across species have been both informative and confusing. The full extent of variation in lignin at the cell, tissue, and species level has not yet been determined.

Our previous studies have shown that unexpected and dramatic effects on the metabolic flux through the monolignol pathway have been the result of protein-protein interactions (Chen et al., 2011, 2014; Naik et al., 2018; Yan et al., 2018). These interactions can affect the extent and direction of flux. The purpose of this short review is to describe the interactions that have been found to date, to guide the discovery of new interactions.

Early Concepts and Evidence for Interactions

The aqueous environment within cells and the concentration of proteins in typical biochemical reactions are quite different. Protein concentrations in cells are very high, and the cellular environment in the *in vitro* reactions is dilute by comparison (Mendes et al., 1995). In theory, compartmentation would increase the concentration of metabolites and proteins leading to higher efficiencies (Ralston and Yu, 2006). In addition to compartmentation, molecular organization of the enzymes could lead to greater efficiencies by direct physical interaction of enzymes or by organizing structures bringing metabolites into greater proximity. "Supramolecular complexes of sequential metabolic enzymes and cellular structural elements" called metabolons have been proposed (Kuzin, 1970; Srere, 1985, 1987).

In 1974, Stafford presumed the existence of multienzyme complexes for phenylpropanoid metabolism because of the diversity of secondary products in the same cells and the need for a mechanism that regulated the complex series of biosynthetic pathways (Stafford, 1974). Support for multienzyme complexes was observed as high molecular weight aggregates, differences in utilization of endogenous versus exogenous origin of metabolites, and existence of multiple isoforms of enzymes.

None of this, however, was unequivocal direct evidence for multienzyme complexes (Stafford, 1974). Early evidence for interactions as membrane bound enzyme complexes in monolignol biosynthesis came from studies of the formation of *p*-coumaric acid from phenylalanine in potato slices (Czichi and Kindl, 1975). Cinnamic acid added to a soluble reaction mixture was a less effective substrate for cinnamate 4-hydroxylase (C4H) than cinnamic acid formed by the phenylalanine ammonia-lyase (PAL) reaction (Czichi and Kindl, 1975, 1977). This observation indicated that endogenously formed cinnamic acid may only be partially equilibrated with exogenous cinnamic acid and that the interaction is dependent on microsomal integrity. PAL and C4H were subsequently proposed to interact and channel substrates for increased catalytic efficiency and regulation of flux (Rasmussen and Dixon, 1999; Achnine et al., 2004; Ralston and Yu, 2006). The concept of "metabolic channeling" refers to the transfer of a metabolite as a product of one enzyme to its interacting enzyme as substrate, without entering a metabolic pool (Winkel, 2004; Jorgensen et al., 2005). However, reconstituted PAL and C4H in yeast revealed that metabolic channeling is not required for the conversion of phenylalanine to *p*-coumaric acid (Ro and Douglas, 2004). Measurements of relative abundance of the 24 monolignol precursors also did not support the hypothesis of metabolic channels in this pathway (Chen et al., 2011). The intriguing hypothesis of channeling in monolignol biosynthesis remains to be substantiated.

Proteins in the endoplasmic reticulum (ER) are thought to act as anchors for multi-protein complexes (Ralston and Yu, 2006). Candidates for protein anchors in the phenylpropanoid pathway are the cytochrome P450s (Durst and Benveniste, 1993). P450s are the largest family of enzymes in plants. Enzymes catalyzing three of the oxidative steps of the monolignol biosynthetic pathway are cytochrome P450s. They are C4H, coumarate 3-hydroxylase (C3H), and ferulate 5-hydroxylase (F5H) also known as coniferaldehyde 5-hydroxylase (CAld5H) (**Figure 1**). C4H and C3H are early steps in the monolignol pathway, common to much of phenylpropanoid metabolism (Vogt, 2010), while CAld5H acts late at lignin-specific steps (Osakabe et al., 1999). These P450s insert an oxygen into hydrophobic regions of proteins and as a result increase reactivity and solubility. All cytochrome P450s are mono-oxygenases that split molecular oxygen through NADPH-cytochrome P450 reductases (CPRs), which need to be tightly coupled to the P450 (Omura, 2010). A hydrophobic membrane-spanning domain anchors the CPR to the ER. Evidence for a direct interaction of CPRs with P450s has been provided using atomic force microscopy and reconstituted phospholipid bilayer disks (Bayburt et al., 1998; Bayburt and Sligar, 2002). Surface residues of both the CPR and the P450 are thought to contribute to the CPR-P450 interaction. The number of P450s in plants far exceeds the number of CPRs therefore that any given CPR might couple with many different P450s, involved with other pathways (Ralston and Yu, 2006). Such interactions are suggested to be dynamic and transient because of the disparity in abundance.

Direct Physical Interaction of Ptr4CL3 and Ptr4CL5 in *Populus trichocarpa*

Genome sequence analysis identified 17 putative genes in *P. trichocarpa* with similarity to known 4CL encoding genes (Shi et al., 2010). Of the 17, only two (*PtrPtr4CL3* and *PtrPtr4CL5*) encode functional enzymes for CoA ligation in stem differentiating xylem (Chen et al., 2013).

A simple experiment indicated that the activity of a mixture of the Ptr4CL3 and Ptr4CL5 isoforms was not additive (Chen et al., 2014). Further experiments demonstrated that the isoforms interact to form a tetramer, *in vitro* and *in vivo*, which appears to have a regulatory role (Chen et al., 2014). Laser microdissection was used to collect different cell types to determine that both isoforms of 4CL are present in the same cells. Most of the transcripts in differentiating xylem are in fiber cells (Chen et al., 2014; Shi et al., 2017; Wang et al., 2018). Isolated fiber cells obtained by microdissection showed that transcripts from both 4CL genes are co-expressed and abundant. Bimolecular fluorescence complementation (BiFC) was carried out using complementing fragments of yellow fluorescence protein (YFP) co-transfected into *P. trichocarpa* protoplasts from differentiating xylem (Chen et al., 2014; Lin et al., 2014). Strong complementation was observed, indicating close spatial proximity (6–10 nm, Fan et al., 2008) between Ptr4CL3 and Ptr4CL5 in the protoplasts.

Further evidence of an interaction between Ptr4CL3 and Ptr4CL5 has been obtained from chemical crosslinking using dithiobis (succinimidyl propionate) (DSP), which makes crosslinks equivalent to 8 carbon linkages, about 12 angstroms (Lomant and Fairbanks, 1976). A crosslinked mixture of Ptr4CL3 and Ptr4CL5 produced a band detected on SDS-PAGE greater than 200 kDa, consistent with a heterotetramer (Chen et al., 2014). Co-immunoprecipitation (Co-IP) also supported the existence of a protein complex involving Ptr4CL3 and Ptr4CL5 (Chen et al., 2014). Antibody prepared against either Ptr4CL3 or Ptr4CL5 was able to co-precipitate both Ptr4CL3 and Ptr4CL5 from extracts of differentiating xylem. Therefore, the complex could be formed *in vitro* and *in vivo*. The stoichiometry of the complex was calculated to be a Ptr4CL3: Ptr4CL5 ratio of 3:1 based on a size estimate of ~240 kDa by SDS-PAGE (Chen et al., 2014). The size of the monomer is ~60 kDa. A ratio of molecular mass of the complex from crosslinking, purification, and protein cleavage-isotope dilution mass spectrometry was 2.69 to 1 of Ptr4CL3 to Ptr4CL5, supporting the conclusion of a hetero-tetramer with three subunits of Ptr4CL3 and one of Ptr4CL5. A mathematical model was developed that predicts the metabolic flux for mixtures of the two enzymes, in the presence of multiple substrates, incorporating the activity of the monomers and tetramers including competitive inhibition, uncompetitive inhibition, and self-inhibition (Chen et al., 2013, 2014). The model suggests that Ptr4CL3-Ptr4CL5 complex improves the homeostatic properties of the pathway, increasing the stability of the pathway by 22% (Naik et al., 2018).

Membrane Protein Complexes Catalyze 4- and 3-Hydroxylation of Cinnamic Acids in *P. trichocarpa*

As discussed earlier, the entry steps to the pathway for monolignol biosynthesis were hypothesized to involve metabolic channeling (Rasmussen and Dixon, 1999; Achnine et al., 2004; Ralston and Yu, 2006). Genomic sequence and transcriptome analysis indicate the involvement of multiple P450 proteins in these steps in *P. trichocarpa* (Shi et al., 2010). Two paralogs of C4H, designated PtrC4H1 and PtrC4H2, are abundantly and specifically expressed in fiber and vessel cells of stem differentiating xylem (Chen et al., 2011; Shi et al., 2017; Wang et al., 2018). One gene encodes the activity of C3H (PtrC3H3). All three are resident ER proteins (Chen et al., 2011).

Co-expression of some combinations of the three hydroxylases had increased activities (Chen et al., 2011) compared to individual enzymes in a yeast system (Urban et al., 1994). Co-expression of PtrC4H1 and PtrC3H3 showed up to 40-fold increases in V_{max} of 4-hydroxylase activity compared to the individual proteins, while PtrC4H2 and PtrC3H3 showed nearly a 100-fold increase in 4-hydroxylase activity. The combination of the three enzymes had a lower K_m with cinnamic acid and more than a 100-fold increase in catalytic efficiency (V_{max}/K_m) compared to the two C4H isoforms (Chen et al., 2011). Co-expression of PtrC4H1 and PtrC3H3 showed nearly a 500-fold increase in catalytic efficiency for 3-hydroxylation of *p*-coumaroyl shikimic into caffeoyl shikimic acid. Co-expression of all three hydroxylases, the increase was over 6,500-fold. Taken together with the activities in extracts of stem differentiating xylem, the results support two hydroxylation pathways, one for the conversion of *p*-coumaric to caffeic acid and the other for the conversion of *p*-coumaroyl shikimic acid to caffeoyl shikimic acid. These results suggest that when co-expressed in the same membrane system, the hydroxylases interact through protein-protein interactions to modulate enzyme activity and metabolic flux (Chen et al., 2011). This inference was further supported by reciprocal co-immunoprecipitation of protein complexes in yeast microsomes.

When full length His-tagged PtrC4H1 and untagged PtrC3H3 were co-expressed in yeast and lysates incubated with anti-His antibody, precipitation of PtrC4H1 was highly enriched in PtrC3H3 (Chen et al., 2011). The results were supported by reciprocal tagging followed by affinity purification and quantitative MS (spectral counting). Further experiments using BiFC, chemical crosslinking, MS, and reciprocal immunoprecipitation in stem differentiating xylem extracts provide a strong evidence for a multi-protein complex of PtrC4H1, PtrC4H2, and PtrC3H3 in yeast microsomes, differentiating xylem protoplasts, and in the differentiating xylem of the growing plant (Chen et al., 2011). The evidence for a trimeric protein complex suggests a structure for metabolic channeling; however, the complex was able to convert exogenous *p*-coumarate into caffeic acid, which argues against channeling.

Interaction of Monolignol Enzymes With Other Pathway Proteins

Protein interactions involving monolignol biosynthetic enzymes have been associated with plant defense signaling. Rice *Oryza sativa*

CCR1 interacts with a Rac family small GTPase (Rac1) in yeast and *in vitro* in a GTP-dependent manner (Kawasaki et al., 2006). Rac1 is a signaling protein that regulates the production of reactive oxygen species mediated by NADPH oxidase and has an important role in defense response. The interaction of Rac1 with CCR1 (**Figure 2**) leads to the enzymatic activation of CCR1 *in vitro* and in rice suspension cell cultures, which results in a higher accumulation of lignin (Kawasaki et al., 2006). Maize CCoAOMT and HCT interact with a plant disease resistance (R) protein, Rp1, which is a nucleotide binding Leu-rich-repeat (NLR) protein that confers pathogen resistance (Wang and Balint-Kurti, 2016). Physical interaction among CCoAOMT, HCT, and Rp1 in a multi-protein complex (**Figure 2**) suppresses the hypersensitive response to *Agrobacterium tumefaciens* infection conferred by Rp1. Downregulation of CCoAOMT or HCT in tobacco *N. benthamiana* disrupts the protein complex and re-activates Rp1, leading to a severe hypersensitive response to the infection (Wang and Balint-Kurti, 2016).

Recent work in Arabidopsis (Gou et al., 2018) has implicated two membrane steroid binding proteins (MSBP1 and MSBP2) in the structural organization of the three lignin P450 hydroxylases. These MSBPs were shown to reside on the ER membrane and to interact with C4H, C3H, and F5H forming MSBP-P450 complexes (**Figure 2**). The MSBPs are proposed to be essential for the stability and activity of the P450s and necessary for channeling of metabolic flux through monolignol biosynthesis (Gou et al., 2018). In Arabidopsis, two cytochrome

P450 reductases (ATR1 and ATR2) were shown to interact with the P450 enzymes (Sundin et al., 2014). One of these, ATR2, is associated with lignin biosynthesis and other phenylpropanoid biosynthetic genes. Plants with loss of function mutations in ATR2 were slightly reduced (6%) in total lignin, were enriched in *p*-coumaric acid derivatives, and were reduced in coniferyl alcohol derivatives. The results were attributed to reduced AtC3H and AtF5H activities (Sundin et al., 2014). Using a yeast-2 hybrid (Y2H) screen, strong interactions with ATR2 were found for AtC4H and AtC3H. Y2H evidence was also obtained indicating direct interactions of AtC4H and AtC3H with At4CL-1 as well as AtC4H with AtCCR1 (Gou et al., 2018). Affinity chromatography and co-purification of the three cytochrome P450 enzymes supported an *in vivo* association of these enzymes. All three P450s were found to interact with the MSBPs presumably on the ER membrane. BiFC experiments show an association of the P450s with MSBPs on the ER. Simultaneous downregulation of both MSBP genes reduced lignin biosynthesis (Gou et al., 2018).

Outlook

A fundamental challenge of biochemistry is the reconstruction of an *in vitro* system, which 1) reproduces *in vivo* function, 2) is consistent with predictions of mathematical models, and 3) accounts for variation in place and time, as well as concentration of enzymes and metabolites (Mendes et al., 1995; Faraji et al., 2018). We are still far from being able to recreate

FIGURE 2 | Enzyme-enzyme interaction network for the monolignol biosynthetic pathway. Blue lines represent protein-protein interactions indicated in the literature that involve monolignol biosynthetic enzymes. Colored circles represent the interacting enzyme families (refer to **Figure 1**).
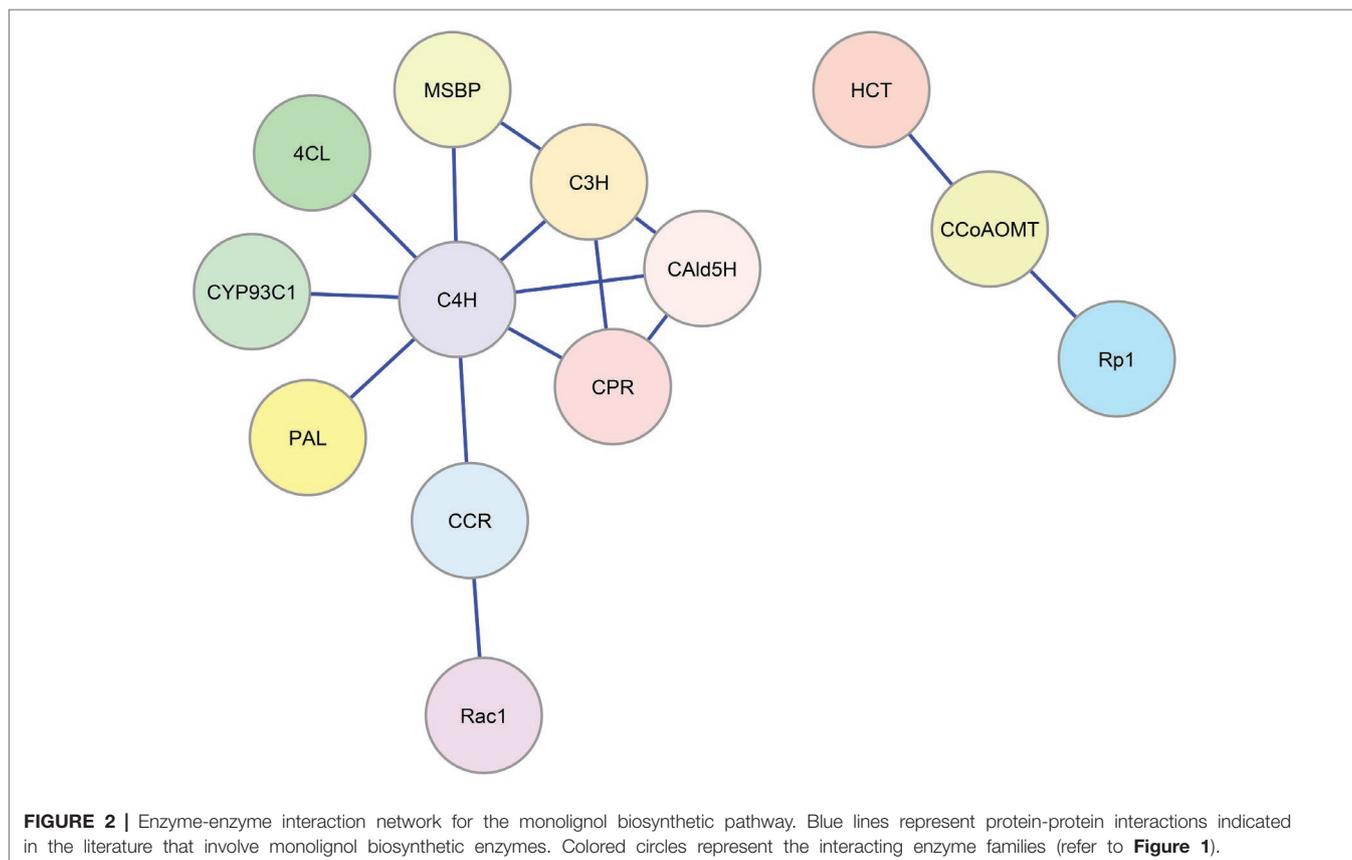
such *in vivo* systems. One long-term strategy would be to build models based on *in vitro* evidence and to make predictions that can be verified in transgenic plants. Reconstruction of the pathway *in vitro* could also provide the necessary evidence. At this time, models of lignin biosynthesis have defined most, but most likely not all, of the components of the system for biosynthesis of monolignols. The best models have incorporated the absolute concentrations of the enzymes and metabolites, including their roles as substrates and inhibitors (Wang Jack et al., 2016; Wang et al., 2018). The variation in cell types and the relative abundance of transcripts of all known monolignol biosynthetic genes have been described (Shi et al., 2017). However, the current models are obtained from data where all enzymes are derived from a single wood-forming tissue (containing fiber, vessel, and ray cells) at the same place and stage of development. The models assumed enzyme functions to be largely independent (Wang et al., 2014; Wang et al., 2018). Cell type–specific data have not yet been incorporated, except to show that proteins that interact *in vitro* are co-expressed in the same cell types (Shi et al., 2017; Wang et al., 2018). Discovery of new regulatory components, such as the recently identified heterodimeric interaction between CAD and CCR that activates the sequential reduction of phenolic CoA esters to aldehydes and monolignols

(**Figure 1**) (Yan et al., 2018), will continue to improve our understanding of the pathway. Incorporating the interactions of enzymes (**Figure 2**) and their metabolites requires a higher level of mechanistic insight based on dynamic interactomic models. The experimental basis for such models is available using inducible promoters that allow induction and repression of gene-specific expression (Borghi, 2010).

AUTHOR CONTRIBUTIONS

FUNDING

REFERENCES

Achnine, L., Blancaflor, E. B., Rasmussen, S., and Dixon, R. A. (2004). Colocalization of L-phenylalanine ammonia-lyase and cinnamic acid 4-hydroxylase for metabolic channeling in phenylpropanoid biosynthesis. *Plant Cell* 16, 3098–3109. doi: 10.1105/tpc.104.024406

Bayburt, T. H., Carlson, J. W., and Sligar, S. G. (1998). Reconstitution and imaging of a membrane protein in a nanometer-size phospholipid bilayer. *J. Struct. Biol.* 123, 37–44. doi: 10.1006/jsbi.1998.4007

Bayburt, T. H., and Sligar, S. G. (2002). Single-molecule height measurements on microsomal cytochrome P450 in nanometer-scale phospholipid bilayer disks. *Proc. Natl. Acad. Sci. USA* 99, 6725–6730. doi: 10.1073/pnas.062565599

Boerjan, W., Ralph, J., and Baucher, M. (2003). Lignin biosynthesis. *Annu. Rev. Plant Biol.* 54, 519–546. doi: 10.1146/annurev.arplant.54.031902.134938

Borghi, L. (2010). Inducible gene expression systems for plants. *Methods Mol. Biol.* 655, 65–75. doi: 10.1007/978-1-60761-765-5_5

Chen, F., and Dixon, R. A. (2007). Lignin modification improves fermentable sugar yields for biofuel production. *Nat. Biotechnol.* 25, 759–761. doi: 10.1038/nbt1316

Chen, H. C., Li, Q., Shuford, C. M., Liu, J., Muddiman, D. C., Sederoff, R. R., et al. (2011). Membrane protein complexes catalyze both 4- and 3-hydroxylation of cinnamic acid derivatives in monolignol biosynthesis. *Proc. Natl. Acad. Sci. USA* 108, 21253–21258. doi: 10.1073/pnas.1116416109

Chen, H. C., Song, J., Wang, J. P., Lin, Y. C., Ducoste, J., Shuford, C. M., et al. (2014). Systems biology of lignin biosynthesis in *Populus trichocarpa*: heteromeric 4-coumaric acid: coenzyme A ligase protein complex formation, regulation, and numerical modeling. *Plant Cell* 26, 876–893. doi: 10.1105/tpc.113.119685

Chen, H. C., Song, J., Williams, C. M., Shuford, C. M., Liu, J., Wang, J. P., et al. (2013). Monolignol pathway 4-coumaric acid: coenzyme A ligases in *Populus trichocarpa*: novel specificity, metabolic regulation, and simulation of coenzyme A ligation fluxes. *Plant Physiol.* 161, 1501–1516. doi: 10.1104/pp.112.210971

Chiang, V. L. (2002). From rags to riches. *Nat. Biotechnol.* 20, 557–558. doi: 10.1038/nbt0602-557

Czichi, U., and Kindl, H. (1975). Formation of *p*-coumaric acid and *o*-coumaric acid from L-phenylalanine by microsomal membrane fractions from potato: Evidence of membrane-bound enzyme complexes. *Planta* 125, 115–125. doi: 10.1007/BF00388698

Czichi, U., and Kindl, H. (1977). Phenylalanine ammonia lyase and cinnamic acid hydroxylases as assembled consecutive enzymes on microsomal membranes of cucumber cotyledons: Cooperation and subcellular distribution. *Planta* 134, 133–143. doi: 10.1007/BF00384962

Durst, F., and Benveniste, I. (1993). "Cytochrome P450 in plants" in *Cytochrome P450*. eds. J. B. Schenkman and H. Greim (Berlin, Heidelberg: Springer Berlin Heidelberg), 293–310.

Fan, J. Y., Cui, Z. Q., Wei, H. P., Zhang, Z. P., Zhou, Y. F., Wang, Y. P., et al. (2008). Split mCherry as a new red bimolecular fluorescence complementation system for visualizing protein-protein interactions in living cells. *Biochem. Biophys. Res. Commun.* 367, 47–53. doi: 10.1016/j.bbrc.2007.12.101

Faraji, M., Fonseca, L. L., Escamilla-Trevino, L., Barros-Rios, J., Engle, N., Yang, Z. K., et al. (2018). Mathematical models of lignin biosynthesis. *Biotechnol. Biofuels* 11:34. doi: 10.1186/s13068-018-1028-9

Fergus, B. J., and Goring, D. A. I. (1970a). The distribution of lignin in birch wood as determined by ultraviolet microscopy. *Holzforschung* 24, 118–124. doi: 10.1515/hfsg.1970.24.4.118

Fergus, B. J., and Goring, D. A. I. (1970b). The location of guaiacyl and syringyl lignins in birch xylem tissue. *Holzforschung* 24, 113–117. doi: 10.1515/hfsg.1970.24.4.113

Gou, M., Ran, X., Martin, D. W., and Liu, C. J. (2018). The scaffold proteins of lignin biosynthetic cytochrome P450 enzymes. *Nat. Plants* 4, 299–310. doi: 10.1038/s41477-018-0142-9

Harada, H., and Côté, W. A. (1985). "Structure of wood" in *Biosynthesis and biodegradation of wood components*. ed. T. Higuchi (Orlando, FL: Academic Press), 1–42. doi: 10.1007/BF01964833

Higuchi, T. (1985). "Biosynthesis of lignin" in *Biosynthesis and biodegradation of wood components*. ed. T. Higuchi (New York: Academic Press), 141–160.

Hu, W. J., Harding, S. A., Lung, J., Popko, J. L., Ralph, J., Stokke, D. D., et al. (1999). Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees. *Nat. Biotechnol.* 17, 808–812. doi: 10.1038/11758

Jorgensen, K., Rasmussen, A. V., Morant, M., Nielsen, A. H., Bjarnholt, N., Zagrobelny, M., et al. (2005). Metabolon formation and metabolic channeling in the biosynthesis of plant natural products. *Curr. Opin. Plant. Biol.* 8, 280–291. doi: 10.1016/j.pbi.2005.03.014

Kawasaki, T., Koita, H., Nakatsubo, T., Hasegawa, K., Wakabayashi, K., Takahashi, H., et al. (2006). Cinnamoyl-CoA reductase, a key enzyme in lignin biosynthesis, is an effector of small GTPase Rac in defense signaling in rice. *Proc. Natl. Acad. Sci. USA* 103, 230–235. doi: 10.1073/pnas.0509875103

Kuzin, A. M. (1970). *Strukturno-Metabolicheskaya Gipoteza v Radiobiologii.* Moscow: Nauka.

Lin, Y. C., Li, W., Chen, H., Li, Q., Sun, Y. H., Shi, R., et al. (2014). A simple improved-throughput xylem protoplast system for studying wood formation. *Nat. Protoc.* 9, 2194–2205. doi: 10.1038/nprot.2014.147

Lomant, A. J., and Fairbanks, G. (1976). Chemical probes of extended biological structures: synthesis and properties of the cleavable protein cross-linking reagent [35S]dithiobis(succinimidyl propionate). *J. Mol. Biol.* 104, 243–261. doi: 10.1016/0022-2836(76)90011-5

Mendes, P., Kell, D. B., and Welch, G. R. (1995). Metabolic channeling in organized enzyme systems: Experiments and models. *Adv. Mol. Cell. Biol.* 11, 1–19. doi: 10.1016/S1569-2558(08)60246-5

Mottiar, Y., Vanholme, R., Boerjan, W., Ralph, J., and Mansfield, S. D. (2016). Designer lignins: harnessing the plasticity of lignification. *Curr. Opin. Biotechnol.* 37, 190–200. doi: 10.1016/j.copbio.2015.10.009

Naik, P., Wang, J. P., Sederoff, R., Chiang, V., Williams, C., and Ducoste, J. J. (2018). Assessing the impact of the 4CL enzyme complex on the robustness of monolignol biosynthesis using metabolic pathway analysis. *PLoS One* 13:e0193896. doi: 10.1371/journal.pone.0193896

Novaes, E., Kirst, M., Chiang, V., Winter-Sederoff, H., and Sederoff, R. (2010). Lignin and biomass: a negative correlation for wood formation and lignin content in trees. *Plant Physiol.* 154, 555–561. doi: 10.1104/pp.110.161281

Omura, T. (2010). Structural diversity of cytochrome P450 enzyme system. *J. Biochem.* 147, 297–306. doi: 10.1093/jb/mvq001

Osakabe, K., Tsao, C. C., Li, L., Popko, J. L., Umezawa, T., Carraway, D. T., et al. (1999). Coniferyl aldehyde 5-hydroxylation and methylation direct syringyl lignin biosynthesis in angiosperms. *Proc. Natl. Acad. Sci. USA* 96, 8955–8960. doi: 10.1073/pnas.96.16.8955

Payen, A. (1839). Recherches sur la matière incrustante des bois. *Acad. Sci. Paris Compt. Rend.* 8, 169–170.

Ralston, L., and Yu, O. (2006). Metabolons involving plant cytochrome P450s. *Phytochem. Rev.* 5, 459–472. doi: 10.1007/s11101-006-9014-4

Rasmussen, S., and Dixon, R. A. (1999). Transgene-mediated and elicitor-induced perturbation of metabolic channeling at the entry point into the phenylpropanoid pathway. *Plant Cell* 11, 1537–1551. doi: 10.1105/tpc.11.8.1537

Rastogi, S., and Dwivedi, U. N. (2006). Down-regulation of lignin biosynthesis in transgenic *Leucaena leucocephala* harboring O-methyltransferase gene. *Biotechnol. Prog.* 22, 609–616. doi: 10.1021/bp050206+

Ro, D. K., and Douglas, C. J. (2004). Reconstitution of the entry point of plant phenylpropanoid metabolism in yeast (*Saccharomyces cerevisiae*): implications for control of metabolic flux into the phenylpropanoid pathway. *J. Biol. Chem.* 279, 2600–2607. doi: 10.1074/jbc.M309951200

Sarkanen, K. V., and Ludwig, C. H. (1971). *Lignins: occurrence, formation, structure and reactions.* ed. H. Erdtman (New York: Wiley-Interscience). doi: 10.1002/pol.1972.110100315

Shi, R., Sun, Y. H., Li, Q., Heber, S., Sederoff, R., and Chiang, V. L. (2010). Towards a systems approach for lignin biosynthesis in *Populus trichocarpa*: transcript abundance and specificity of the monolignol biosynthetic genes. *Plant Cell Physiol.* 51, 144–163. doi: 10.1093/pcp/pcp175

Shi, R., Wang, J. P., Lin, Y. C., Li, Q., Sun, Y. H., Chen, H., et al. (2017). Tissue and cell-type co-expression networks of transcription factors and wood component genes in *Populus trichocarpa*. *Planta* 245, 927–938. doi: 10.1007/s00425-016-2640-1

Srere, P. A. (1985). The metabolon. *Trends Biochem. Sci.* 10, 109–110. doi: 10.1016/0968-0004(85)90266-X

Srere, P. A. (1987). Complexes of sequential metabolic enzymes. *Annu. Rev. Biochem.* 56, 89–124. doi: 10.1146/annurev.bi.56.070187.000513

Stafford, H. A. (1974). Possible multienzyme complexes regulating the formation of C6-C3 phenolic compounds and lignins in higher plants. *Recent Adv. Phytochem.* 8, 53–79. doi: 10.1016/B978-0-12-612408-8.50009-3

Sundin, L., Vanholme, R., Geerinck, J., Goeminne, G., Hofer, R., Kim, H., et al. (2014). Mutation of the inducible arabidopsis thaliana cytochrome P450 REDUCTASE2 alters lignin composition and improves saccharification. *Plant Physiol.* 166, 1956–1971. doi: 10.1104/pp.114.245548

Urban, P., Werck-Reichhart, D., Teutsch, H. G., Durst, F., Regnier, S., Kazmaier, M., et al. (1994). Characterization of recombinant plant cinnamic acid 4-hydroxylase produced in yeast. Kinetic and spectral properties of the major plant P450 of the phenylpropanoid pathway. *Eur. J. Biochem.* 222, 843–850. doi: 10.1111/j.1432-1033.1994.tb18931.x

Vanholme, R., Morreel, K., Ralph, J., and Boerjan, W. (2008). Lignin engineering. *Curr. Opin. Biotechnol.* 11, 278–285. doi: 10.1016/j.pbi.2008.03.005

Vogt, T. (2010). Phenylpropanoid biosynthesis. *Mol. Plant* 3, 2–20. doi: 10.1093/mp/ssp106

Wang, G. F., and Balint-Kurti, P. J. (2016). Maize homologs of CCoAOMT and HCT, two key enzymes in lignin biosynthesis, form complexes with the NLR Rp1 protein to modulate the defense response. *Plant Physiol.* 171, 2166–2177. doi: 10.1104/pp.16.00224

Wang Jack, P., Tunlaya-Anukit, S., Shi, R., Yeh, T. F., Chuang, L., Isik, F., et al. (2016). "A proteomic-based quantitative analysis of the relationship between monolignol biosynthetic protein abundance and lignin content using transgenic *Populus trichocarpa*" in *Recent advances in polyphenol research*. eds. K. Yoshida, V. Cheynier, and S. Quideau (Elsevier), 1–19. doi: 10.1002/9781118883303.ch4

Wang, J. P., Matthews, M. L., Williams, C. M., Shi, R., Yang, C., Tunlaya-Anukit, S., et al. (2018). Improving wood properties for wood utilization through multi-omics integration in lignin biosynthesis. *Nat. Commun.* 9:1579. doi: 10.1038/s41467-018-03863-z

Wang, J. P., Naik, P. P., Chen, H. C., Shi, R., Lin, C. Y., Liu, J., et al. (2014). Complete proteomic-based enzyme reaction and inhibition kinetics reveal how monolignol biosynthetic enzyme families affect metabolic flux and lignin in *Populus trichocarpa*. *Plant Cell* 26, 894–914. doi: 10.1105/tpc.113.120881

Weng, J. K., and Chapple, C. (2010). The origin and evolution of lignin biosynthesis. *New Phytol.* 187, 273–285. doi: 10.1111/j.1469-8137.2010.03327.x

Winkel, B. S. (2004). Metabolic channeling in plants. *Annu. Rev. Plant Biol.* 55, 85–107. doi: 10.1146/annurev.arplant.55.031903.141714

Yan, X., Liu, J., Kim, H., Liu, B., Huang, X., Yang, Z., et al. (2018). CAD1 and CCR2 protein complex formation in monolignol biosynthesis in *Populus trichocarpa*. *New Phytol.* doi: 10.1111/nph.15505

Zhou, C., Li, Q., Chiang, V. L., Lucia, L. A., and Griffis, D. P. (2011). Chemical and spatial differentiation of syringyl and guaiacyl lignins in poplar wood via time-of-flight secondary ion mass spectrometry. *Anal. Chem.* 83, 7020–7026. doi: 10.1021/ac200903y

Lignin Engineering in Forest Trees

Alexandra Chanoca[1,2], Lisanne de Vries[1,2] and Wout Boerjan[1,2]*

[1] Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent, Belgium, [2] VIB Center for Plant Systems Biology, Ghent, Belgium

Wood is a renewable resource that is mainly composed of lignin and cell wall polysaccharides. The polysaccharide fraction is valuable as it can be converted into pulp and paper, or into fermentable sugars. On the other hand, the lignin fraction is increasingly being considered a valuable source of aromatic building blocks for the chemical industry. The presence of lignin in wood is one of the major recalcitrance factors in woody biomass processing, necessitating the need for harsh chemical treatments to degrade and extract it prior to the valorization of the cell wall polysaccharides, cellulose and hemicellulose. Over the past years, large research efforts have been devoted to engineering lignin amount and composition to reduce biomass recalcitrance toward chemical processing. We review the efforts made in forest trees, and compare results from greenhouse and field trials. Furthermore, we address the value and potential of CRISPR-based gene editing in lignin engineering and its integration in tree breeding programs.

Keywords: lignin, forest trees, genetic engineering, CRISPR, field trial

INTRODUCTION

Fossil resources are the main feedstock for energy and organic compounds, and their use results in the emission of greenhouse gases associated with climate change. The coming climate crash calls for an urgent transition from a fossil-based to a bio-based economy in which lignocellulosic biomass rather than oil is used for the production of fuels, chemicals and materials. Wood is an important source of lignocellulosic biomass; it is mainly composed of secondary-thickened cell walls rich in cellulose, hemicelluloses, and lignin. All three polymers can be valorized in the bio-based economy. Cellulose is a source for the pulp and paper industry, and both cellulose and hemicelluloses can be depolymerized to their monosaccharides for fermentation into, e.g., bio-ethanol, lactic acid and detergents (Vanholme et al., 2013b). As lignin negatively affects the efficiency of wood processing toward these applications, trees can be engineered to accumulate less lignin, to become more amenable for the production of paper and fermentable sugars. On the other hand, lignin is increasingly being considered a valuable component in the bio-based economy. Indeed, given that lignin is the largest renewable aromatic source on Earth, the economic viability of a bio-refinery can be significantly increased if lignin is also valorized, and used as a resource for the production of chemicals (Holladay et al., 2007; Tuck et al., 2012; Davis et al., 2013; Ragauskas et al., 2014; Li C. et al., 2015; Van den Bosch et al., 2015; Rinaldi et al., 2016; Upton and Kasko, 2016; Schutyser et al., 2018).

The lignin polymer is composed of monolignols that are produced by the phenylpropanoid and monolignol biosynthetic pathways, by a series of enzymatic reactions starting with the deamination of phenylalanine (**Figure 1**). The monolignols are synthesized in the cytoplasm and translocated to the apoplast, where they are dehydrogenated to monolignol radicals by the action

of laccases and peroxidases (Berthet et al., 2011; Zhao et al., 2013). These monolignol radicals then couple with each other in a combinatorial fashion, generating a range of chemical bonds such as the aryl-ether bond (β-O-4), resinol bond (β-β), and phenylcoumaran bond (β-5) (Boerjan et al., 2003; Ralph et al., 2004; Vanholme et al., 2010). The most common monolignols are the hydroxycinnamyl alcohols p-coumaryl, coniferyl, and sinapyl alcohols, which generate the H, G, and S units upon their incorporation into the lignin polymer, respectively (Bonawitz and Chapple, 2010; Ralph et al., 2019; Vanholme et al., 2019). The relative contribution of the lignin building blocks varies among taxa, developmental stage, tissue and cell type, and even cell wall layer; lignin from softwoods (gymnosperms) is comprised almost entirely of G units with a minor fraction of H units, while lignin from hardwoods (angiosperms) has S units in addition to G units and traces of H units (Boerjan et al., 2003; Vanholme et al., 2010, 2019). Besides these traditional monolignols, a variety of other p-hydroxylated aromatic molecules can be incorporated in the lignin polymer to various levels (Vanholme et al., 2019).

Given that lignin is a major recalcitrance factor in wood delignification processes, large research efforts have been devoted to unravel the lignin biosynthetic pathway, and to study the effects of perturbations of the lignin biosynthesis genes on lignin amount and composition, and on wood processing efficiency. While modifications in genes ranging from those encoding transcription factors up to those encoding oxidative enzymes have resulted in altered lignin content, composition or deposition (Eriksson et al., 2000; Li Y.H. et al., 2003; Liang et al., 2008; Lu et al., 2013; Lin et al., 2016; Xu et al., 2017; Yang et al., 2017; Obudulu et al., 2018), this review will focus on the results obtained by engineering the lignin biosynthetic genes.

ENGINEERING THE LIGNIN PATHWAY

Table 1 provides an overview of the different studies on downregulated or mutated lignin biosynthetic genes in poplar, pine, eucalyptus and birch, with the resulting effects on wood processing efficiencies, when determined. Reducing the activity of any step of the lignin biosynthetic pathway, starting from PAL up to CAD may result in a reduction in lignin content (**Table 1**). Several parameters influence the degree of lignin reduction, such as the target gene and the degree of downregulation of the enzyme activity, which in turn depends on the efficiency of the silencing construct used, the size of the gene family, and redundancy within the gene family. Generally, the downregulation of the steps from C4H up to CCR results in a more dramatic reduction in lignin amount (Hu et al., 1999; Meyermans et al., 2000; Zhong et al., 2000; Li L. et al., 2003; Jia et al., 2004; Lu et al., 2004; Leplé et al., 2007; Coleman et al., 2008a,b; Bjurhager et al., 2010; Mansfield et al., 2012; Ralph et al., 2012; Min et al., 2014; Van Acker et al., 2014; Zhou et al., 2015, 2018; Saleme et al., 2017; Xiang et al., 2017) than downregulation of F5H, COMT and CAD (Van Doorsselaere et al., 1995; Baucher et al., 1996; Lapierre et al., 1999; Jouanin et al., 2000; Van Acker et al., 2014; Wang et al., 2018). Lignin reduction can be associated with an increase in S/G, such as in C3'H- (Coleman et al., 2008a; Ralph et al., 2012)

and CCoAOMT-downregulated trees (Meyermans et al., 2000), or a decrease in S/G ratio such as in CSE- (Saleme et al., 2017), and COMT-downregulated trees (Van Doorsselaere et al., 1995; Lapierre et al., 1999; Jouanin et al., 2000). Interestingly, low-lignin 4CL-downregulated poplars were found to have an increase in S/G (Min et al., 2014; Xiang et al., 2017), a decrease in S/G (Voelker et al., 2010; Zhou et al., 2015), or ratios comparable to wild type (Hu et al., 1999; Li L. et al., 2003). This variance cannot be associated with the promoter or the method used for downregulation, suggesting that differences in the degree of silencing, growth conditions or developmental state influence this trait. On the other hand, the strongest effects on H/G/S lignin composition have been observed for trees downregulated in C3'H and HCT, which deposit lignin with large increases in H unit content (Coleman et al., 2008a; Ralph et al., 2012; Vanholme et al., 2013a), whereas trees that overexpress F5H produce lignin strongly enriched in S units (Franke et al., 2000; Li L. et al., 2003; Stewart et al., 2009), and trees that are downregulated in COMT have dramatically reduced S unit biosynthesis (Van Doorsselaere et al., 1995; Lapierre et al., 1999).

Both the reduced lignin content and variation in the H/G/S ratios can affect the biomass processing efficiency. Consistent with the established role of lignin in determining biomass recalcitrance (Zeng et al., 2014; McCann and Carpita, 2015; Li et al., 2016; Wang et al., 2018), plants with reduced levels of lignin show increased chemical pulping and saccharification efficiency (Hu et al., 1999; Jouanin et al., 2000; Rastogi and Dwivedi, 2006; Wadenbäck et al., 2008; Wang et al., 2012; Sykes et al., 2015; Cai et al., 2016; Edmunds et al., 2017; Saleme et al., 2017; Xiang et al., 2017; Van Acker et al., 2017; Wang et al., 2018). An increased level of H units reduces lignin polymer length and, hence, increases the removal of lignin from the biomass (Mansfield et al., 2012; Sykes et al., 2015). Increased S/G results in lignin more easily cleaved and extracted in alkaline conditions, supposedly due to the lower degree of polymerization (Huntley et al., 2003; Stewart et al., 2009; Mansfield et al., 2012; Yoo et al., 2018).

The processing efficiency of the biomass can also be modified by the increased incorporation of molecules that generally represent minor components in the lignin of wild-type plants. The incorporation of ferulic acid in CCR-deficient trees results in the formation of acetal bonds in the lignin polymer, which are easily cleaved in acidic biomass pretreatments (Leplé et al., 2007; Ralph et al., 2008; Van Acker et al., 2014). Indeed, the levels of ferulic acid in lignin positively correlated with a higher saccharification efficiency (Van Acker et al., 2014). The incorporation of 5-hydroxyconiferyl alcohol and 5-hydroxyconiferaldehyde in the lignin of COMT-deficient poplars (Van Doorsselaere et al., 1995; Lapierre et al., 1999; Jouanin et al., 2000; Morreel et al., 2004; Lu et al., 2010) gives rise to benzodioxane bonds, potentially preventing covalent linkages between lignin and the polysaccharide hydroxyl groups (Weng et al., 2010; Vanholme et al., 2012; Nishimura et al., 2018). On the other hand, COMT deficiency also results in a more condensed lignin due to the relatively higher levels of the condensed β-β and β-5 bonds, and the lower levels of β-O-4 bonds, when the S unit frequency drops. Chemical pulping of wood derived from poplars strongly downregulated for COMT resulted in a higher

FIGURE 1 | Lignin biosynthetic pathway. Alternative monomers and heterologously expressed enzymes are shown in bold. PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate:CoA ligase; HCT, *p*-Hydroxycinnamoyl-CoA:quinate/shikimate-*p*-hydroxycinnamoyltransferase; C3'H, *p*-coumaroyl quinate/shikimate 3'-hydroxylase; CSE, caffeoyl shikimate esterase; CCoAOMT, caffeoyl-CoA *O*-methyltransferase; CCR, cinnamoyl-CoA reductase; F5H, ferulate 5-hydroxylase/CAld5H, coniferaldehyde 5-hydroxylase; COMT, caffeic acid *O*-methyltransferase; CAD, cinnamyl alcohol dehydrogenase; FMT, feruloyl-CoA monolignol transferase; PMT, *p*-coumaroyl-CoA monolignol transferase.

pulp yield, counterbalanced by the residual lignin content in the pulp. These trees had a lower lignin and a higher cellulose content (Jouanin et al., 2000). On the other hand, poplars that were modestly downregulated for COMT had a large decrease in pulp yield, presumably because lignin content had remained normal

while the lignin had a higher frequency of condensed bonds that negatively affected the lignin extraction (Lapierre et al., 1999; Pilate et al., 2002). The incorporation of cinnamaldehydes in the lignin polymer in CAD-deficient trees results in shorter lignin polymer chains, hence a higher proportion of free phenolic

end groups that increase the solubility of the polymer in alkali. The incorporation of cinnamaldehydes in the lignin polymer presumably also reduces the covalent interaction of the aliphatic chain with hemicellulose, again rendering the lignin more soluble. In addition, due to the extended conjugated system that is generated when a cinnamaldehyde β-O-4 couples with another monomer, the aromatic ether bond of the incorporated cinnamaldehyde becomes more susceptible to alkaline cleavage (Lapierre et al., 1989; Van Acker et al., 2017).

Lignin polymerization is a combinatorial radical coupling process, allowing a wide range of phenolic compounds to be naturally incorporated into the lignin polymer (Boerjan et al., 2003; Vanholme et al., 2019). Researchers have attempted to tailor the lignin amount and composition to improve biomass processing by expression of heterologous genes, aiming at the biosynthesis and incorporation of various compatible phenolic compounds as alternative monolignols into the lignin polymer (Ralph, 2006; Vanholme et al., 2012; Mottiar et al., 2016; Mahon and Mansfield, 2018). One example is the introduction of genes encoding enzymes that are needed for S unit biosynthesis in pine; the simultaneous expression of *F5H*, *COMT* and *CAD* successfully introduced S units in *Pinus radiata* (Wagner et al., 2015; Edmunds et al., 2017). The introduction of the gene encoding a monolignol 4-O-methyltransferase (*MOMT4*) into poplar leads to the formation of 4-O-methylated coniferyl and sinapyl alcohols, which cannot be incorporated into the growing lignin polymer because they lack the aromatic hydroxyl group. This leads to a halt in lignin polymerization and results in trees with lower lignin content and higher saccharification efficiency (Bhuiya and Liu, 2010; Cai et al., 2016). Poplars have also been engineered to contain ester linkages in the lignin polymer backbone. Coniferyl ferulate esters were introduced into the polymer via expression of a *FERULOYL-CoA:MONOLIGNOL TRANSFERASE* (*FMT*) gene derived from *Angelica sinensis* (Wilkerson et al., 2014), leading to an improved saccharification efficiency under various pretreatment conditions (Wilkerson et al., 2014; Kim et al., 2017; Bhalla et al., 2018), and an improved kraft pulping efficiency as compared to wild type (Zhou et al., 2017). Monolignol *p*-coumarate esters have also been engineered in poplar, via expression of a rice *p-COUMAROYL-CoA:MONOLIGNOL TRANSFERASE* (*PMT*) gene, resulting in a higher frequency of resistant interunit bonds and a higher frequency of G and S terminal units with free phenolic groups (Smith et al., 2015; Sibout et al., 2016). While in Arabidopsis the heterologous expression of *PMT* resulted in a reduced lignin amount accompanied by an increased saccharification efficiency (Sibout et al., 2016), there was no decrease in lignin amount in poplar and the saccharification efficiency was not determined (Smith et al., 2015).

While several modifications of the lignin amount and composition were shown to provide improvements in biomass processing, these modifications were often accompanied by a biomass yield penalty (Leplé et al., 2007; Wadenbäck et al., 2008; Wagner et al., 2009; Voelker et al., 2010; Stout et al., 2014; Van Acker et al., 2014; Sykes et al., 2015; Zhou et al., 2018). A recent metastudy perturbed 21 lignin biosynthesis genes in *P. trichocarpa*, and comprehensively integrated the results of transcriptomic, proteomic, fluxomic, and phenomic data of 221 lines. The authors concluded that tree growth is not associated with lignin amount, subunit composition or specific linkages (Wang et al., 2018), but rather correlated with the presence of collapsed xylem vessels (Coleman et al., 2008a,b; Wagner et al., 2009; Voelker et al., 2010; Vargas et al., 2016; De Meester et al., 2018), the activation of a cell wall integrity pathway (Bonawitz et al., 2014) and/or the accumulation of chemical inhibitors (Gallego-Giraldo et al., 2011; Muro-Villanueva et al., 2019).

Whereas substantial efforts have been made to decrease lignin content by downregulation of lignin biosynthetic genes, studies on the upregulation of the lignin pathway and the overproduction of lignin have been scarce. Indeed, reports on the overexpression of *F5H* show an unchanged or even a decrease in lignin content (Huntley et al., 2003; Li L. et al., 2003; Stewart et al., 2009; Mansfield et al., 2012; Edmunds et al., 2017). The overexpression of *CAD* and *COMT* has resulted in gene-silencing rather than upregulation, or no effect on expression levels was detected (Baucher et al., 1996; Lapierre et al., 1999; Jouanin et al., 2000; Leplé et al., 2007; Van Acker et al., 2014). The overexpression of the R2R3-MYB transcription factors *PtoMYB92*, *PtoMYB216*, and *PtoMYB74* all resulted in additional xylem layers, thicker xylem cell walls as well as ectopic lignin deposition, and the plants accumulated 13–50% more lignin (Tian Q. et al., 2013; Li C.F. et al., 2015; Li et al., 2018). The *MYB* overexpression lines constitutively upregulated the lignin biosynthesis pathway genes, and while plants overexpressing *MYB92* and *MYB74* had a biomass penalty, the overexpression of *MYB216* resulted in plants with up to 50% more lignin and no developmental phenotype. As lignin is increasingly being considered an important resource for the sustainable production of chemicals (Cao et al., 2018) the engineering of plants overproducing lignin should be further explored.

FIELD TRIALS

The examples discussed above clearly show that lignin engineering via down- or upregulation of phenylpropanoid pathway genes – or expression of heterologous genes – has the potential to increase the processing efficiency of lignocellulosic biomass. Due to practical and regulatory reasons, most studies report on data obtained from the analysis of trees grown in a greenhouse. However, experiments with trees grown in a greenhouse typically do not take into account developmental processes such as growth cessation and dormancy. In addition, greenhouse experiments do not provide sufficient insight into the interaction of the engineered plant with environmental factors such as soil type, wind, and pathogens. Understanding these interactions is an important step in the translation of research results toward commercial applications. Indeed, the body of work produced by studies for which permission to establish field trials was granted, highlights important differences in phenotype between greenhouse- and field grown trees. **Table 1** summarizes the reports on field trials performed with *4CL, CCoAOMT, CCR, COMT,* and *CAD* downregulated trees.

TABLE 1 | Overview of forest trees with modified expression of lignin biosynthesis genes.

Species	Gene	Method	Growth conditions	Lignin amount	Lignin composition	Saccharification efficiency	Pulping efficiency	Biomass yield	References
Pinus radiata	HCT	RNAi	TE cultures	↓42%	↑ H/G	n.d.	n.d.	n/a	Wagner et al., 2007
Pinus radiata	CCoAOMT	RNAi	TE cultures	↓20%	↑ H/G	n.d.	n.d.	n/a	Wagner et al., 2011
Pinus radiata	CCR	RNAi	TE cultures	↓46%	Trace amount of ferulic acid	n.d.	n.d.	n/a	Wagner et al., 2013
Pinus radiata	F5H + COMT	Heterologous (over) expression	TE cultures	n.d.	Incorporation of S units	n.d.	n.d.	n/a	Wagner et al., 2015
Betula pendula	COMT	Co-suppression	Greenhouse	WT	↓ S/G	n.d.	n.d.	WT/↓	Tiimonen et al., 2005
Leucaena leucocephala	COMT	Antisense	Greenhouse	↓30%	↓ S units (histochemical)	n.d.	↑	WT	Rastogi and Dwivedi, 2006
Pinus radiata	4CL	RNAi	Greenhouse	↓36%	↑ H/G	n.d.	n.d.	↓	Wagner et al., 2009
Pinus taeda	4CL	Xylem-specific RNAi	Greenhouse	↓33%	n.d.	↑	n.d.	n.d.	Edmunds et al., 2017
Pinus taeda	CAD	Mutant allele	Greenhouse	n.d.	↑ cinnamaldehydes, ↑ benzaldehyde, ↑ dihydroconiferyl alcohol, ↓ G units	n.d.	n.d.	n.d.	Ralph et al., 1997
				↓9%	↑ coniferaldehyde	n.d.	n.d.	WT	MacKay et al., 1997
				n.d.	↑ dihydroconiferyl alcohol, ↑ vanillin, ↑ coniferaldehyde, ↑ H/G	n.d.	n.d.	n.d.	Lapierre et al., 2000
Pinus taeda	F5H + COMT + SAD/CAD	Heterologous (over) expression	Greenhouse	WT	Incorporation of S units	WT	n.d.	WT	Edmunds et al., 2017
Picea abies	CCR	Antisense	Greenhouse	↓8%	↓ H/G	n.d.	↑	↓ diameter	Wadenbäck et al., 2008
Eucalyptus urophylla × *E. grandis*	C3'H	Antisense	Greenhouse	↓27%	↓ S/G, ↑ H units	↑	n.d.	↓ height	Sykes et al., 2015
Eucalyptus urophylla × *E. grandis*	C4H	Antisense	Greenhouse	↓36%	↓ S/G	↑	n.d.	↓ height	Sykes et al., 2015
P. tremula × *P.tremuloides*	C4H	Antisense	Greenhouse	↓33%	WT	n.d.	n.d.	↓ height	Bjurhager et al., 2010
P. tremuloides Michx.	4CL	Antisense	Greenhouse	↓45%	WT S/G	n.d.	n.d.	↑	Hu et al., 1999*
P. tremuloides	4CL	Downregulation	Greenhouse	Up to ↓40%	WT S/G	n.d.	n.d.	WT	Li L. et al., 2003*
	4CL + F5H	Downregulation + overexpression	Greenhouse	↓52%	↑ S/G	n.d.	n.d.	WT	
P. tremuloides	4CL	Antisense	Greenhouse	↓40%	WT S/G	n.d.	n.d.	WT	Li L. et al., 2003
P. tomentosa	4CL	Antisense	Greenhouse	↓42%	n.d.	n.d.	n.d.	WT	Jia et al., 2004
P. trichocarpa	4CL	Downregulation	Greenhouse	↓30%	WT S/V	↑	n.d.	n.d.	Min et al., 2012
P. nigra L. × *P. maximowiczii*	4CL	Antisense	Greenhouse	Up to ↓55%	↓ S/V/↑ S/V	n.d.	n.d.	n.d.	Min et al., 2014
P. nigra L. × *P. maximowiczii*	4CL + F5H	Antisense + antisense	Greenhouse	WT/↓	WT	n.d.	n.d.	n.d.	

(Continued)

TABLE 1 | Continued

Species	Gene	Method	Growth conditions	Lignin amount	Lignin composition	Saccharification efficiency	Pulping efficiency	Biomass yield	References
P. nigra L. × P. maximowiczii	4CL + F5H	Antisense + overexpression	Greenhouse	↓	↑S/V	n.d.	n.d.	n.d.	
P. tremula × P. alba	4CL1	CRISPR/Cas9 mutants	Greenhouse	↓23%	↓S/G	n.d.	n.d.	WT	Zhou et al., 2015
P. tremula × P. alba	4CL2	CRISPR/Cas9 mutants	Greenhouse	WT	WT S/G	n.d.	n.d.	WT	Zhou et al., 2015
P. nigra L. × P. maximowiczii	4CL	Antisense	Greenhouse	↓47%	↑S/V	n.d.	n.d.	n.d.	Xiang et al., 2017
P. alba × P. grandidentata	C3'H	RNAi	Greenhouse	↓56%	↑H units, ↑S/G	n.d.	n.d.	n.d.	Coleman et al., 2008a
				n.d.	n.d.	n.d.	n.d.	↓	Coleman et al., 2008b
				WT/↓	n.d.	↑	↑	↓	Mansfield et al., 2012
P. alba × P. grandidentata	C3'H	Hairpin	Greenhouse	↓50%	↑H units, ↑ S/G	n.d.	n.d.	n.d.	Ralph et al., 2012
P. alba × P. glandulosa	C3'H	Downregulation	Greenhouse	↓30%	n.d.	n.d.	n.d.	↓ diameter	Zhou et al., 2018
P. nigra	HCT	Mutant allele	Greenhouse	WT	↑H units, ↑S/G	n.d.	n.d.	n.d.	Vanholme et al., 2013a
P. alba × P. glandulosa	HCT	Downregulation	Greenhouse	↓20%	n.d.	n.d.	n.d.	↓ diameter	Zhou et al., 2018
P. tremula × P. alba	CSE	Hairpin	Greenhouse	up to ↓25%	↑H units, ↓S/G	↑	n.d.	WT	Saleme et al., 2017
P. tremula × P. alba	CCoAOMT	Sense	Greenhouse	↓12%	↑ S/G, incorporation of *p*-hydroxybenzoic acid	n.d.	n.d.	WT	Meyermans et al., 2000
P. tremula × P. alba	CCoAOMT	Antisense	Greenhouse	Up to ↓40%	WT, incorporation of *p*-hydroxybenzoic acid	n.d.	n.d.	WT	Zhong et al., 2000
P. tormentosa	CCoAOMT	Antisense	Greenhouse	Up to ↓26%	n.d.	n.d.	n.d.	WT	Lu et al., 2004
P. tremula × P. alba	CCR	Antisense and co-suppression	Greenhouse	up to ↓30%	↓S/G, ferulic acid incorporation	n.d.	n.d.	WT	Leplé et al., 2007*
				↓19%	WT/↓	↑	n.d.	n.d.	Van Acker et al., 2014
P. nigra L. × P. maximowiczii	F5H	Antisense	Greenhouse	WT/↑	↓S/G	n.d.	n.d.	n.d.	Min et al., 2014
P. tremuloides	F5H	Heterologous OE	Greenhouse	WT	↑S/G	n.d.	n.d.	WT	Li L. et al., 2003
P. tremula × P. alba	F5H	Heterologous OE	Greenhouse	n.d.	↑ S/G	n.d.	n.d.	n.d.	Franke et al., 2000
				WT	↑ S/G	n.d.	↑	WT	Huntley et al., 2003
				↓	↑ S/G, ↓ *p*-hydroxybenzoic acid	n.d.	n.d.	WT	Stewart et al., 2009
				WT	↑ S/G	WT	↑	WT	Mansfield et al., 2012
P. tremula × P. alba	COMT	Antisense	Greenhouse	WT	↓ S/G, incorporation of 5-OH-G	n.d.	n.d.	WT	Van Doorsselaere et al., 1995*
				WT	↓ S/G	n.d.	↓	WT	Lapierre et al., 1999

(Continued)

TABLE 1 | Continued

Species	Gene	Method	Growth conditions	Lignin amount	Lignin composition	Saccharification efficiency	Pulping efficiency	Biomass yield	References
P. tremula × P. alba	COMT	Co-suppression	Greenhouse	↓17%	↓ S/G, incorporation of 5-OH-G	n.d.	↑	WT	Jouanin et al., 2000
P. tremula × P. alba	CAD	Antisense and co-suppression	Greenhouse	WT	WT S/G, ↑ conjugated aldehyde moieties	n.d.	↑	WT	Baucher et al., 1996*
				↓	WT S/G	n.d.	↑	WT	Lapierre et al., 1999
P. tremula × P. alba	CAD	Hairpin	Greenhouse	↓10%	↓ canonical S/G, ↑ sinapaldehyde	↑	n.d.	WT	Van Acker et al., 2017
P. tremula × P. alba	MOMT4	Heterologous OE	Greenhouse	↓15%	↓ S/G, ↓ *p*-hydroxybenzoic acid	↑	n.d.	WT	Cai et al., 2016
P. alba × P. grandidentata	PMT	Heterologous OE	Greenhouse	WT	WT S/G, incorporation of *p*-coumarate conjugates	n.d.	n.d.	WT	Smith et al., 2015
P. alba × P. grandidentata	FMT	Heterologous OE	Greenhouse	WT	↑ S/G, incorporation of acylated monolignols	↑	n.d.	WT	Wilkerson et al., 2014
				n.d.	n.d.	↑	n.d.	n.d.	Kim et al., 2017
				WT	WT S/G	↑	n.d.	n.d.	Bhalla et al., 2018
P. tremula × P. alba	4CL	Antisense	Field trial	WT	Low S/G in brown wood	WT	n.d.	↓	Voelker et al., 2010
P. tomentosa Carr.	4CL	Sense	Field trial	Up to ↓30%	↑ S/G	n.d.	n.d.	↑	Tian X.M. et al., 2013
P. nigra L. × P. maximowiczii A	4CL	Antisense	Field trial	WT/ ↓	WT	↑	n.d.	n.d.	Xiang et al., 2017
P. trichocarpa	4CL	Antisense	Field trial	↓	↓ S/G	n.d.	n.d.	↓	Stout et al., 2014
P. tomentosa	4CL	Antisense	Field trial	Up to ↓10%	n.d.	WT	n.d.	WT	Wang et al., 2012
P. tomentosa	CCoAOMT	Antisense	Field trial	↓6-10%	n.d.	↑	n.d.	WT	Wang et al., 2012
P. tremula × P. alba	CCoAOMT	Antisense	Field trial	↓13%	Slight increment in S/G	n.d.	↑	WT	Wei et al., 2008
P. tremula × P. alba	CCR	Antisense and co-suppression	Field trial	Up to ↓47%	↓ S/G, Incorporation of ferulic acid	n.d.	↑	↓	Leplé et al., 2007
				Up to ↓24%	Incorporation of ferulic acid	↑	n.d.	↓	Van Acker et al., 2014
P. tremula × P. alba	COMT	Antisense	Field trial	WT	↓ S/G, incorporation of 5-OH-G	n.d.	↓	WT	Lapierre et al., 1999*
P. tremula × P. alba	COMT	Antisense	Field trial	WT	↓ S/G, incorporation of 5-OH-G	n.d.	↓	WT	Pilate et al., 2002*
P. tremula × P. alba	CAD	Sense and antisense	Field trial	Slightly ↓	WT S/G, ↑ free phenolic units	n.d.	↑	WT	Lapierre et al., 1999*
P. tremula × P. alba	CAD	Antisense	Field trial	Slightly ↓	WT S/G, ↑ free phenolic units	n.d.	↑	WT	Pilate et al., 2002*

*n.d., not determined; n/a, not applicable; S/G, syringyl/guaiacyl ratio; S/V, syringaldehyde/vanillin ratio; H/G, p-hydroxyphenyl/guaiacyl ratio. Papers reporting plants that have been used for independent studies reporting on biotic or abiotic stress tolerance are shown with a *. For abbreviations of gene names (see legend **Figure 1**). Lignin amount was determined by various methods, see the corresponding reference for specific information. Readers are referred to Wang et al. (2018) for additional raw data on downregulated lines for monolignol biosynthesis genes in P. trichocarpa.*

Confirming the potential of modified lignocellulosic biomass as a substrate for applications, several lignin-engineered field-trial grown trees showed improvements in wood processing. Poplars downregulated for *CCoAOMT* grown for 5 years in a field trial in Beijing (China), showed an increased glucose and xylose release upon saccharification (Wang et al., 2012). Poplars downregulated for *CCR* and grown in a field trial in France, proved to be more amenable to chemical kraft pulping (Leplé et al., 2007). Two additional field trials conducted in France and Belgium with *CCR*-downregulated poplars resulted in up to 160% improvement in ethanol production in simultaneous saccharification and fermentation (SSF) assays; however the plants had up to 50% biomass reduction (Van Acker et al., 2014). Field trials with *CAD*-downregulated poplar also showed promising results. These trees showed slightly less lignin than wild type and proved more amenable to kraft delignification (Lapierre et al., 1999). Consistently, the same lines grown in larger-scale field trials in France and the United Kingdom showed a mild decrease in lignin amount and an improvement in kraft pulping deemed commercially relevant, since the plants needed 6% less alkali to achieve a delignification similar to that of wild-type trees (Pilate et al., 2002).

However, conflicting reports on both biomass yield and downstream processing efficiency suggest that these parameters are highly influenced by environmental factors. While a field trial conducted in China using *4CL* downregulated poplars found that, even with a 28% decrease in lignin content compared to wild type, the trees had about 8% increased height (Tian X.M. et al., 2013), consistent with greenhouse studies (Hu et al., 1999), other field trials found that *4CL*-downregulated poplars had decreased biomass and were sometimes even dwarfed (Voelker et al., 2010; Stout et al., 2014; Marchin et al., 2017). Reports also diverge regarding downstream processing efficiencies of wood derived from these *4CL*-downregulated field-grown poplars. While up to 100% increase in sugar recovery was found for *4CL1*-downregulated trees (35S-driven antisense *4CL* construct) grown in a mountain site in the United States (Xiang et al., 2017), data obtained from field studies conducted in Oregon (United States) found that *Pt4CL1* promoter-driven antisense silenced *4CL1* poplars had no improvement in saccharification efficiency compared to wild type (Voelker et al., 2010). Likewise, a long term study in Wenling (China), found that *4CL*-downregulated poplars did not show a significant improvement in sugar yield compared to wild type (Wang et al., 2012). In both latter cases, the trees showed mild decreases in lignin amount which did not translate into higher processing efficiency, potentially because of the higher concentration of extractives that could interfere with enzymatic activity (Voelker et al., 2010).

Field trial studies have shown that environmental factors can influence lignification and restore traits to wild-type levels as compared to the levels achieved when the same plants were grown in the greenhouse. While *4CL*-downregulated trees had decreased lignin content when grown under greenhouse conditions, analysis of the same *4CL* antisense poplars, but grown in the field, has often shown that the lignin content was increased as compared to the greenhouse-grown trees and sometimes even restored to wild-type levels (Stout et al., 2014; Xiang et al.,
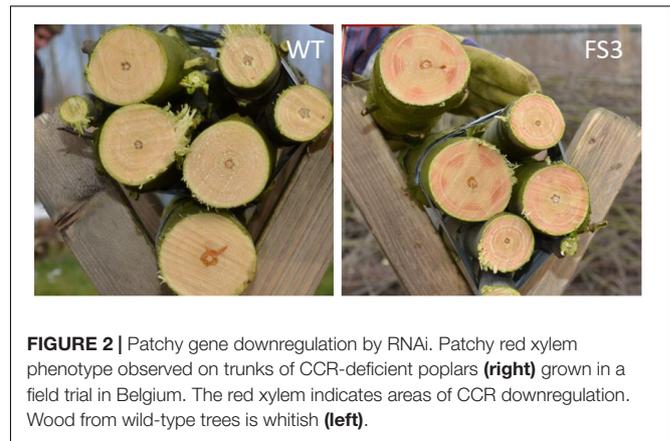
FIGURE 2 | Patchy gene downregulation by RNAi. Patchy red xylem phenotype observed on trunks of CCR-deficient poplars **(right)** grown in a field trial in Belgium. The red xylem indicates areas of CCR downregulation. Wood from wild-type trees is whitish **(left)**.

2017). Similarly, lignin levels were much less reduced in CCR-deficient poplars when they were grown in the field as compared to when they were grown in the greenhouse (Van Acker et al., 2014). At least for the CCR-deficient poplars, it is possible that the higher lignification level of field-grown trees is due to the fact that the wood samples were taken during winter. When tree growth ceases in autumn, the trees still have time to fully lignify their cell walls by the time the tree enters dormancy, as compared to greenhouse-grown trees that develop new xylem continuously. Lignin composition has also been shown to differ between greenhouse- and field-grown low-lignin trees. *4CL*-downregulated poplars grown in a field in North Carolina had lignin with a lower S/G than when the same lines were grown in the greenhouse (Stout et al., 2014).

Taken together, these results show that data obtained from greenhouse-grown trees cannot easily be extrapolated to field-grown trees, underpinning the need for field trial experiments at different locations. Some lines presented a yield penalty rendering them less interesting for applications, highlighting the need for a better understanding of the molecular basis of the yield penalty and the development of strategies to overcome this problem.

Lignin has been shown to play an important role in pathogen resistance (Miedes et al., 2014; Zhao and Dixon, 2014), and it plays a pivotal role in allowing the plant to transport water. This suggests that lignin modifications could have an impact on plant stress tolerance. While further investigation is needed to fully address this possibility, the downregulation of *4CL*, *COMT,* and *CAD* in poplar did not dramatically alter the feeding performance of leaf-feeding herbivores (Tiimonen et al., 2005; Brodeur-Campbell et al., 2006; Hjalten et al., 2013). The effect of the downregulation of *COMT* and *CAD* in poplar on plant-insect interactions has also been assessed on field-grown trees, and it was shown that the lignin-modified trees had normal incidence of visiting and feeding insects, as well as normal responses to microbial pathogens (Pilate et al., 2002; Halpin et al., 2007). These results indicate that trees with modified lignin do not necessarily suffer more than wild-type plants from pests and diseases. Nevertheless, profiling of the endosphere bacterial microbiome of wood harvested from field-grown, *CCR*-downregulated poplars demonstrated shifts in the bacterial

FIGURE 3 | Genetic improvement of forest trees through a combination of breeding tools. To accelerate the genetic improvement of forest trees for pulp and biorefinery applications, classical and new breeding tools need to be smartly combined. Classical breeding involves phenotypic selection of trees for controlled crosses, followed by phenotypic selection. With the advent of genome sequence information of many forest trees, new strategies such as Genomic Selection, Genome Wide Association Studies (GWAS) and Breeding with Rare Defective Alleles (BRDA) have been developed to speed up the capture and enrichment of DNA polymorphisms associated with beneficial traits. CRISPR-based genome editing allows to modify the genome in a way that mimics natural polymorphisms. Genetic modification involves the stable integration of foreign DNA into the tree to overproduce (an) enzyme(s) or downregulate (a) gene(s). Combining the classical and New Breeding Techniques is needed to provide sufficient highly quality wood for society.

community, presumably because of the altered abundance of particular phenolic metabolites in the xylem (Beckers et al., 2017).

Considering the role of lignin in xylem function and structure, the water relations of a few low lignin-modified poplars have been assessed. *4CL*-downregulated poplars were found to have reduced hydraulic conductivity, potentially interfering with plant growth (Marchin et al., 2017). Hydraulic stress experiments with poplars downregulated for *CCR*, *COMT* or *CAD* showed that these plants had a lower resistance to cavitation, while maintaining normal xylem hydraulic conductivity and water transport (Awad et al., 2012). These results suggest that the growth of low-lignin mutants might be influenced by water availability. As for any new hybrid obtained from classical breeding, field tests are needed to evaluate field performance and stress tolerance of lignin-engineered trees.

PROSPECTS FOR LIGNIN ENGINEERING IN FOREST TREES

The performance of lignin-engineered plants appears to be highly influenced by the environmental conditions. It is unclear, however, whether the differences observed between greenhouse-grown and field-grown trees, or between trees grown in different field locations, result from different levels of gene suppression or from interaction of the engineered trait with the environment (GxE). Indeed, unstable downregulation is a shortcoming of gene silencing techniques that are based on RNAi. This is witnessed by observing variation in the red xylem phenotype that is observed when particular lignin biosynthesis genes, such as *CAD*, *COMT*, or *CCR*, are downregulated. The red xylem coloration is often not uniform throughout the xylem, but rather appears in patches that reflect variable levels of gene silencing (Leplé et al., 2007; Voelker et al., 2010; Van Acker et al., 2014; **Figure 2**). In

addition, the use of gene silencing methods can potentially result in concomitant silencing of closely related gene family members – perhaps to various degrees – clouding the interpretations and camouflaging the effects of downregulation of individual genes.

These issues can now be easily overcome by the use of CRISPR-based gene editing technologies that enable stable loss-of-function mutations (knock-outs) in specific target genes, allowing the dissection of the function of individual genes within families. For example, the targeting of individual *4CL* gene family members in poplar showed that 4CL1 is related to lignification, whereas 4CL2 is involved in proanthocyanidin production (Zhou et al., 2015). In addition to knock-out alleles, CRISPR-based gene editing also allows to create new alleles that confer partial reduction in enzyme activity. This opens the possibility to fine-tune the level of residual enzyme activity and to bypass the yield penalty that is often observed when lignin amount drops below a threshold level. Another promising avenue for lignin engineering in forest trees made possible through CRISPR-based genome engineering is the simultaneous editing of multiple genes (allele stacking) to optimize biomass processing efficiency, as exemplified in Arabidopsis where stacking of the *transaldolase* (*tra*) and *comt* mutations, the *c4h* and *comt* mutations, or the *4cl* and *comt* mutations resulted in additive and synergistic improvements in saccharification efficiency (de Vries et al., 2018). Indeed, a systems approach in *P. trichocarpa* predicts that the concomitant downregulation of *PAL* and *CCoAOMT*, or *PAL*, *C3'H* and *CCOAOMT* will substantially improve wood properties and sugar release (Wang et al., 2018).

The use of CRISPR-based genome editing in tree improvement for the pulp and paper and the bio-refinery industries, as well as for the production of platform aromatics from the hydrogenolytic breakdown of lignin, will be most valuable when this technology is strategically combined with other breeding techniques (**Figure 3**). Indeed, large variation in lignin amount and S/G composition already exists in natural populations of forest trees (Studer et al., 2011). Given that both traits affect the glucose release upon saccharification (Yoo et al., 2018), exploiting this genetic diversity by conventional breeding,

aided by Genome Wide Association Studies (GWAS) (Porth et al., 2013; Fahrenkrog et al., 2017; Liu et al., 2018), Breeding with Rare Defective Alleles (BRDA) (Vanholme et al., 2013a) or genomic selection (Yin et al., 2010; Muchero et al., 2015; Pawar et al., 2018; Xie et al., 2018), is a valuable strategy to obtain lines that have improved wood processing efficiency. Once elite trees are obtained by these breeding methods, genetic engineering and CRISPR-based gene editing of specific genes is a very promising avenue to further improve these elite genotypes without breaking up their genetic constitution and without going through lengthy breeding cycles. Given the imminent climate crash, we have no more time to lose in adopting these new breeding techniques in our race to the biobased economy.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Awad, H., Herbette, S., Brunel, N., Tixier, A., Pilate, G., Cochard, H., et al. (2012). No trade-off between hydraulic and mechanical properties in several transgenic poplars modified for lignins metabolism. *Environ. Exp. Bot.* 77, 185–195. doi: 10.1016/j.envexpbot.2011.11.023

Baucher, M., Chabbert, B., Pilate, G., Vandoorsselaere, J., Tollier, M. T., Petitconil, M., et al. (1996). Red xylem and higher lignin extractability by down-regulating a cinnamyl alcohol dehydrogenase in poplar. *Plant Physiol.* 112, 1479–1490. doi: 10.1104/pp.112.4.1479

Beckers, B., Op De Beeck, M., Weyens, N., Boerkan, W., and Vangronsveld, J. (2017). Structural variability and niche differentiation in the rhizosphere and endosphere bacterial microbiome of field-grown poplar trees. *Microbiome* 5:25. doi: 10.1186/s40168-017-0241-2

Berthet, S., Demont-Caulet, N., Pollet, B., Bidzinski, P., Cézard, L., Le Bris, P., et al. (2011). Disruption of *LACCASE4* and *17* results in tissue-specific alterations to lignification of *Arabidopsis thaliana* stems. *Plant Cell* 23, 1124–1137. doi: 10.1105/tpc.110.082792

Bhalla, A., Bansal, N., Pattathil, S., Li, M. Y., Shen, W., Particka, C. A., et al. (2018). Engineered lignin in poplar biomass facilitates Cu-catalyzed alkaline-oxidative pretreatment. *ACS Sustain. Chem. Eng.* 6, 2932–2941. doi: 10.1021/acssuschemeng.7b02067

Bhuiya, M.-W., and Liu, C.-J. (2010). Engineering monolignol 4-O-methyltransferases to modulate lignin biosynthesis. *J. Biol. Chem.* 285, 277–285. doi: 10.1074/jbc.M109.036673

Bjurhager, I., Olsson, A.-M., Zhang, B., Gerber, L., Kumar, M., Berglund, L. A., et al. (2010). Ultrastructure and mechanical properties of *Populus* wood with reduced lignin content caused by transgenic down-regulation of cinnamate 4-hydroxylase. *Biomacromolecules* 11, 2359–2365. doi: 10.1021/bm100487e

Boerjan, W., Ralph, J., and Baucher, M. (2003). Lignin biosynthesis. *Annu. Rev. Plant Biol.* 54, 519–546.

Bonawitz, N. D., and Chapple, C. (2010). The genetics of lignin biosynthesis: connecting genotype to phenotype. *Annu. Rev. Genet.* 44, 337–363. doi: 10.1146/annurev-genet-102209-163508

Bonawitz, N. D., Kim, J. I., Tobimatsu, Y., Ciesielski, P. N., Anderson, N. A., Ximenes, E., et al. (2014). Disruption of Mediator rescues the stunted growth of a lignin-deficient *Arabidopsis* mutant. *Nature* 509, 376–380. doi: 10.1038/nature13084

Brodeur-Campbell, S. E., Vucetich, J. A., Richter, D. L., Waite, T. A., Rosemier, J. N., and Tsai, C. J. (2006). Insect herbivory on low-lignin transgenic aspen. *Environ. Entomol.* 35, 1696–1701. doi: 10.1603/0046-225x(2006)35%5B1696:iholta%5D2.0.co;2

Cai, Y., Zhang, K., Kim, H., Hou, G., Zhang, X., Yang, H., et al. (2016). Enhancing digestibility and ethanol yield of Populus wood via expression of

an engineered monolignol 4-O-methyltransferase. *Nat. Commun.* 7:11989. doi: 10.1038/ncomms11989

Cao, L., Yu, I. K. M., Liu, Y., Ruan, X., Tsang, D. C. W., Hunt, A. J., et al. (2018). Lignin valorization for the production of renewable chemicals: state-of-the-art review and future prospects. *Bioresour. Technol.* 269, 465–475. doi: 10.1016/j.biortech.2018.08.065

Coleman, H. D., Park, J.-Y., Nair, R., Chapple, C., and Mansfield, S. D. (2008a). Rnai-mediated suppression of *p*-coumaroyl-CoA 3′-hydroxylase in hybrid poplar impacts lignin deposition and soluble secondary metabolism. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4501–4506. doi: 10.1073/pnas.0706537105

Coleman, H. D., Samuels, A. L., Guy, R. D., and Mansfield, S. D. (2008b). Perturbed lignification impacts tree growth in hybrid poplar – A function of sink strength, vascular integrity, and photosynthetic assimilation. *Plant Physiol.* 148, 1229–1237. doi: 10.1104/pp.108.125500

Davis, R., Tao, L., Tan, E., Biddy, M., Beckham, G., Scarlata, C., et al. (2013). *Process Design and Economics for the Conversion of Lignocellulosic Biomass to Hydrocarbons: Dilute-Acid and Enzymatic Deconstruction of Biomass to Sugars and Biological Conversion of Sugars to Hydrocarbons*. Golden, CO: National Renewable Energy Lab.

De Meester, B., De Vries, L., Özparpucu, M., Gierlinger, N., Corneillie, S., Pallidis, A., et al. (2018). Vessel-specific reintroduction of cinnamoyl-coa reductase1 (CCR1) in dwarfed ccr1 mutants restores vessel and xylary fiber integrity and increases biomass. *Plant Physiol.* 176:611. doi: 10.1104/pp.17.01462

de Vries, L., Vanholme, R., Van Acker, R., De Meester, B., Sundin, L., and Boerjan, W. (2018). Stacking of a low-lignin trait with an increased guaiacyl and 5-hydroxyguaiacyl unit trait leads to additive and synergistic effects on saccharification efficiency in *Arabidopsis thaliana*. *Biotechnol. Biofuels* 11:257. doi: 10.1186/s13068-018-1257-y

Edmunds, C. W., Peralta, P., Kelley, S. S., Chiang, V. L., Sharma-Shivappa, R. R., Davis, M. F., et al. (2017). Characterization and enzymatic hydrolysis of wood from transgenic *Pinus taeda* engineered with syringyl lignin or reduced lignin content. *Cellulose* 24, 1901–1914. doi: 10.1007/s10570-017-1231-z

Eriksson, M. E., Israelsson, M., Olsson, O., and Moritz, T. (2000). Increased gibberellin biosynthesis in transgenic trees promotes growth, biomass production and xylem fiber length. *Nat. Biotechnol.* 18, 784–788. doi: 10.1038/77355

Fahrenkrog, A. M., Neves, L. G., Resende, M. F. R. Jr., Vazquez, A. I., De Los Campos, G., Dervinis, C., et al. (2017). Genome-wide association study reveals putative regulators of bioenergy traits in Populus deltoides. *New Phytol.* 213, 799–811. doi: 10.1111/nph.14154

Franke, R., Mcmichael, C. M., Meyer, K., Shirley, A. M., Cusumano, J. C., and Chapple, C. (2000). Modified lignin in tobacco and poplar plants over-expressing the *Arabidopsis* gene encoding ferulate 5-hydroxylase. *Plant J.* 22, 223–234. doi: 10.1046/j.1365-313x.2000.00727.x

Gallego-Giraldo, L., Escamilla-Trevino, L., Jackson, L. A., and Dixon, R. A. (2011). Salicylic acid mediates the reduced growth of lignin down-regulated plants. *Proc. Natl. Acad. Sci. U.S.A.* 108, 20814–20819. doi: 10.1073/pnas.1117873108

Halpin, C., Thain, S. C., Tilston, E. L., Guiney, E., Lapierre, C., and Hopkins, D. W. (2007). Ecological impacts of trees with modified lignin. *Tree Genet. Genomes* 3, 101–110. doi: 10.1007/s11295-006-0060-2

Hjalten, J., Axelsson, E. P., Julkunen-Tiitto, R., Wennstrom, A., and Pilate, G. (2013). Innate and introduced resistance traits in genetically modified aspen trees and their effect on leaf beetle feeding. *PLoS One* 8:8. doi: 10.1371/journal.pone.0073819

Holladay, J. E., White, J. F., Bozell, J. J., and Johnson, D. (2007). *Top Value-Added Chemicals From Biomass-Volume II—Results of Screening for Potential Candidates From Biorefinery Lignin*. Richland, WA: Pacific Northwest National Lab

Hu, W.-J., Harding, S. A., Lung, J., Popko, J. L., Ralph, J., Stokke, D. D., et al. (1999). Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees. *Nat. Biotechnol.* 17, 808–812. doi: 10.1038/11758

Huntley, S. K., Ellis, D., Gilbert, M., Chapple, C., and Mansfield, S. D. (2003). Significant increases in pulping efficiency in C4H-F5H-transformed poplars: improved chemical savings and reduced environmental toxins. *J. Agric. Food Chem.* 51, 6178–6183. doi: 10.1021/jf034320o

Jia, C., Zhao, H., Wang, H., Xing, Z., Du, K., Song, Y., et al. (2004). Obtaining the transgenic poplars with low lignin content through down-regulation OF4CL. *Chin. Sci. Bull.* 49, 905–909. doi: 10.1007/bf03184009

Jouanin, L., Goujon, T., De Nadaï, V., Martin, M.-T., Mila, I., Vallet, C., et al. (2000). Lignification in transgenic poplars with extremely reduced caffeic acid *O*-methyltransferase activity. *Plant Physiol.* 123, 1363–1373.

Kim, K. H., Dutta, T., Ralph, J., Mansfield, S. D., Simmons, B. A., and Singh, S. (2017). Impact of lignin polymer backbone esters on ionic liquid pretreatment of poplar. *Biotechnol. Biofuels* 10:101. doi: 10.1186/s13068-017-0784-2

Lapierre, C., Jouin, D., and Monties, B. (1989). On the molecular origin of the alkali solubility of Gramineae lignins. *Phytochemistry* 28, 1401–1403. doi: 10.1016/s0031-9422(00)97755-0

Lapierre, C., Pollet, B., Petit-Conil, M., Toval, G., Romero, J., Pilate, G., et al. (1999). Structural alterations of lignins in transgenic poplars with depressed cinnamyl alcohol dehydrogenase or caffeic acid *O*-methyltransferase activity have an opposite impact on the efficiency of industrial kraft pulping. *Plant Physiol.* 119, 153–163.

Lapierre, C., Pollet, B., MacKay, J. J., and Sederoff, R. R. (2000). Lignin structure in a mutant pine deficient in cinnamyl alcohol dehydrogenase. *J. Agric. Food Chem.* 48, 2326–2331. doi: 10.1021/jf991398p

Leplé, J.-C., Dauwe, R., Morreel, K., Storme, V., Lapierre, C., Pollet, B., et al. (2007). Downregulation of cinnamoyl-coenzyme A reductase in poplar: multiple-level phenotyping reveals effects on cell wall polymer metabolism and structure. *Plant Cell* 19, 3669–3691. doi: 10.1105/tpc.107.054148

Li, C., Ma, X., Yu, H., Fu, Y., and Luo, K. (2018). Ectopic expression of Ptomyb74 in poplar and arabidopsis promotes secondary cell wall formation. *Front. Plant Sci.* 9:1262. doi: 10.3389/fpls.2018.01262

Li, C., Zhao, X., Wang, A., Huber, G. W., and Zhang, T. (2015). Catalytic transformation of lignin for the production of chemicals and fuels. *Chem. Rev.* 115, 11559–11624. doi: 10.1021/acs.chemrev.5b00155

Li, C. F., Wang, X. Q., Ran, L. Y., Tian, Q. Y., Fan, D., and Luo, K. M. (2015). Ptomyb92 is a transcriptional activator of the lignin biosynthetic pathway during secondary cell wall formation in *Populus tomentosa*. *Plant Cell Physiol.* 56, 2436–2446. doi: 10.1093/pcp/pcv157

Li, L., Zhou, Y., Cheng, X., Sun, J., Marita, J. M., Ralph, J., et al. (2003). Combinatorial modification of multiple lignin traits in trees through multigene cotransformation. *Proc. Natl. Acad. Sci. U.S.A.* 100, 4939–4944. doi: 10.1073/pnas.0831166100

Li, Y. H., Kajita, S., Kawai, S., Katayama, Y., and Morohoshi, N. (2003). Down-regulation of an anionic peroxidase in transgenic aspen and its effect on lignin characteristics. *J. Plant Res.* 116, 175–182. doi: 10.1007/s10265-003-0087-5

Li, M., Pu, Y., and Ragauskas, A. J. (2016). Current understanding of the correlation of lignin structure with biomass recalcitrance. *Front. Chem.* 4:45.

Liang, H. Y., Frost, C. J., Wei, X. P., Brown, N. R., Carlson, J. E., and Tien, M. (2008). Improved sugar release from lignocellulosic material by introducing a tyrosine-rich cell wall peptide gene in poplar. *Clean-Soil Air Water* 36, 662–668. doi: 10.1002/clen.200800079

Lin, C. Y., Li, Q. Z., Tunlaya-Anukit, S., Shi, R., Sun, Y. H., Wang, J. P., et al. (2016). A cell wall-bound anionic peroxidase, Ptrpo21, is involved in lignin polymerization in Populus trichocarpa. *Tree Genet. Genomes* 12:18. doi: 10.1021/jf502560k

Liu, J., Ye, M., Zhu, S., Jiang, L., Sang, M., Gan, J., et al. (2018). Two-stage identification of Snp effects on dynamic poplar growth. *Plant J.* 93, 286–296. doi: 10.1111/tpj.13777

Lu, F. C., Marita, J. M., Lapierre, C., Jouanin, L., Morreel, K., Boerjan, W., et al. (2010). Sequencing around 5-hydroxyconiferyl alcohol-derived units in caffeic acid *O*-methyltransferase-deficient poplar lignins. *Plant Physiol.* 153, 569–579. doi: 10.1104/pp.110.154278

Lu, J., Zhao, H. Y., Wei, J. H., He, Y. K., Shi, C., Wang, H. Z., et al. (2004). Lignin reduction in transgenic poplars by expressing antisense Ccoaomt gene. *Progr. Natural Sci.* 14, 1060–1063. doi: 10.1080/10020070412331344801

Lu, S., Li, Q., Wei, H., Chang, M.-J., Tunlaya-Anukit, S., Kim, H., et al. (2013). Ptr-miR397a is a negative regulator of laccase genes affecting lignin content in *Populus trichocarpa*. *Proc. Natl. Acad. Sci. U.S.A.* 110, 10848–10853. doi: 10.1073/pnas.1308936110

MacKay, J. J., O'Malley, D. M., Presnell, T., Booker, F. L., Campbell, M. M., Whetten, R. W., et al. (1997). Inheritance, gene expression, and lignin characterization in a mutant pine deficient in cinnamyl alcohol-dehydrogenase. *Proc. Natl. Acad. Sci.* 94, 8255–8260. doi: 10.1073/pnas.94.15.8255

Mahon, E. L., and Mansfield, S. D. (2018). Tailor-made trees: engineering lignin for ease of processing and tomorrow's bioeconomy. *Curr. Opin. Biotechnol.* 56, 147–155. doi: 10.1016/j.copbio.2018.10.014

Mansfield, S. D., Kang, K.-Y., and Chapple, C. (2012). Designed for deconstruction - poplar trees altered in cell wall lignification improve the efficacy of bioethanol production. *New Phytol.* 194, 91–101. doi: 10.1111/j.1469-8137.2011.04031.x

Marchin, R. M., Stout, A. T., Davis, A. A., and King, J. S. (2017). Transgenically altered lignin biosynthesis affects photosynthesis and water relations of field-grown Populus trichocarpa. *Biomass Bioenergy* 98, 15–25. doi: 10.1016/j.biombioe.2017.01.013

McCann, M. C., and Carpita, N. C. (2015). Biomass recalcitrance: a multi-scale, multi-factor, and conversion-specific property. *J. Exp. Bot.* 66, 4109–4118. doi: 10.1093/jxb/erv267

Meyermans, H., Morreel, K., Lapierre, C., Pollet, B., De Bruyn, A., Busson, R., et al. (2000). Modifications in lignin and accumulation of phenolic glucosides in poplar xylem upon down-regulation of caffeoyl-coenzyme A O-methyltransferase, an enzyme involved in lignin biosynthesis. *J. Biol. Chem.* 275, 36899–36909. doi: 10.1074/jbc.m006915200

Miedes, E., Vanholme, R., Boerjan, W., and Molina, A. (2014). The role of the secondary cell wall in plant resistance to pathogens. *Front. Plant Sci.* 5:358. doi: 10.3389/fpls.2014.00358

Min, D., Li, Q., Jameel, H., Chiang, V., and Chang, H. M. (2012). The cellulase-mediated saccharification on wood derived from transgenic low-lignin lines of black cottonwood (Populus trichocarpa). *Appl. Biochem. Biotechnol.* 168, 947–955. doi: 10.1007/s12010-012-9833-2

Min, D. Y., Yang, C. M., Chiang, V., Jameel, H., and Chang, H. M. (2014). The influence of lignin-carbohydrate complexes on the cellulase-mediated saccharification II: transgenic hybrid poplars (Populus nigra L. and Populus maximowiczii A.). *Fuel* 116, 56–62. doi: 10.1016/j.fuel.2013.07.046

Morreel, K., Ralph, J., Lu, F., Goeminne, G., Busson, R., Herdewijn, P., et al. (2004). Phenolic profiling of caffeic acid O-methyltransferase-deficient poplar reveals novel benzodioxane oligolignols. *Plant Physiol.* 136, 4023–4036. doi: 10.1104/pp.104.049312

Mottiar, Y., Vanholme, R., Boerjan, W., Ralph, J., and Mansfield, S. D. (2016). Designer lignins: harnessing the plasticity of lignification. *Curr. Opin. Biotechnol.* 37, 190–200. doi: 10.1016/j.copbio.2015.10.009

Muchero, W., Guo, J., Difazio, S. P., Chen, J.-G., Ranjan, P., Slavov, G. T., et al. (2015). High-resolution genetic mapping of allelic variants associated with cell wall chemistry in Populus. *BMC Genomics* 16:24. doi: 10.1186/s12864-015-1215-z

Muro-Villanueva, F., Mao, X., and Chapple, C. (2019). Linking phenylpropanoid metabolism, lignin deposition, and plant growth inhibition. *Curr. Opin. Biotechnol.* 56, 202–208. doi: 10.1016/j.copbio.2018.12.008

Nishimura, H., Kamiya, A., Nagata, T., Katahira, M., and Watanabe, T. (2018). Direct evidence for alpha ether linkage between lignin and carbohydrates in wood cell walls. *Sci. Rep.* 8:6538. doi: 10.1038/s41598-018-24328-9

Obudulu, O., Mahler, N., Skotare, T., Bygdell, J., Abreu, I. N., Ahnlund, M., et al. (2018). A multi-omics approach reveals function of secretory carrier-associated membrane proteins in wood formation of populus trees. *BMC Genomics* 19:18. doi: 10.1186/s12864-017-4411-1

Pawar, P. M.-A., Schnürer, A., Mellerowicz, E. J., and Rönnberg-Wästljung, A. C. (2018). Qtl mapping of wood Ft-Ir chemotypes shows promise for improving biofuel potential in short rotation coppice willow (Salix spp.). *Bioenergy Res.* 11, 351–363. doi: 10.1007/s12155-018-9901-8

Pilate, G., Guiney, E., Holt, K., Petit-Conil, M., Lapierre, C., Leple, J. C., et al. (2002). Field and pulping performances of transgenic trees with altered lignification. *Nat. Biotechnol.* 20, 607–612. doi: 10.1038/nbt0602-607

Porth, I., Klapště, J., Skyba, O., Hannemann, J., Mckown, A. D., Guy, R. D., et al. (2013). Genome-wide association mapping for wood characteristics in Populus identifies an array of candidate single nucleotide polymorphisms. *New Phytol.* 200, 710–726. doi: 10.1111/nph.12422

Ragauskas, A. J., Beckham, G. T., Biddy, M. J., Chandra, R., Chen, F., Davis, M. F., et al. (2014). Lignin valorization: improving lignin processing in the biorefinery. *Science* 344:1246843. doi: 10.1126/science.1246843

Ralph, J. (2006). "What makes a good monolignol substitute?," in *The Science and Lore of the Plant Cell Wall – Biosynthesis, Structure and Function*, ed. T. Hayashi (Boca Raton, FL: BrownWalker Press).

Ralph, J., Akiyama, T., Coleman, H. D., and Mansfield, S. D. (2012). Effects on lignin structure of coumarate 3-hydroxylase downregulation in poplar. *Bioenergy Res.* 5, 1009–1019. doi: 10.1007/s12155-012-9218-y

Ralph, J., Kim, H., Lu, F., Grabber, J. H., Leplé, J.-C., Berrio-Sierra, J., et al. (2008). Identification of the structure and origin of a thioacidolysis marker compound for ferulic acid incorporation into angiosperm lignins (and an indicator for cinnamoyl CoA reductase deficiency). *Plant J.* 53, 368–379. doi: 10.1111/j.1365-313X.2007.03345.x

Ralph, J., Lapierre, C., and Boerjan, W. (2019). Lignin structure and its engineering. *Curr. Opin. Biotechnol.* 56, 240–249. doi: 10.1016/j.copbio.2019.02.019

Ralph, J., Lundquist, K., Brunow, G., Lu, F., Kim, H., Schatz, P. F., et al. (2004). Lignins: natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids. *Phytochem. Rev.* 3, 29–60. doi: 10.1023/b:phyt.0000047809.65444.a4

Ralph, J., MacKay, J. J., Hatfield, R. D., O'Malley, D. M., Whetten, R. W., and Sederoff, R. R. (1997). Abnormal lignin in a loblolly pine mutant. *Science* 277, 235–239. doi: 10.1126/science.277.5323.235

Rastogi, S., and Dwivedi, U. N. (2006). Down-regulation of lignin biosynthesis in transgenic *Leucaena leucocephala* harboring O-methyltransferase gene. *Biotechnol. Progr.* 22, 609–616. doi: 10.1021/bp050206%2B

Rinaldi, R., Jastrzebski, R., Clough, M. T., Ralph, J., Kennema, M., Bruijnincx, P. C., et al. (2016). Paving the way for lignin valorisation: recent advances in bioengineering, biorefining and catalysis. *Angew Chem. Int. Ed. Engl.* 55, 8164–8215. doi: 10.1002/anie.201510351

Saleme, M. L. S., Cesarino, I., Vargas, L., Kim, H., Vanholme, R., Goeminne, G., et al. (2017). Silencing caffeoyl Shikimate Esterase affects lignification and improves saccharification in poplar. *Plant Physiol.* 175, 1040–1057. doi: 10.1104/pp.17.00920

Schutyser, W., Renders, T., Van Den Bosch, S., Koelewijn, S. F., Beckham, G. T., and Sels, B. F. (2018). Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading. *Chem. Soc. Rev.* 47, 852–908. doi: 10.1039/c7cs00566k

Sibout, R., Le Bris, P., Legée, F., Cézard, L., Renault, H., and Lapierre, C. (2016). Structural redesigning arabidopsis lignins into alkali-soluble lignins through the expression of p-coumaroyl-CoA:monolignol transferase PMT. *Plant Physiol.* 170, 1358–1366. doi: 10.1104/pp.15.01877

Smith, R. A., Gonzales-Vigil, E., Karlen, S. D., Park, J. Y., Lu, F., Wilkerson, C. G., et al. (2015). Engineering monolignol p-Coumarate Conjugates into poplar and *Arabidopsis* lignins. *Plant Physiol.* 169, 2992–3001. doi: 10.1104/pp.15.00815

Stewart, J. J., Akiyama, T., Chapple, C., Ralph, J., and Mansfield, S. D. (2009). The effects on lignin structure of overexpression of ferulate 5-hydroxylase in hybrid poplar. *Plant Physiol.* 150, 621–635. doi: 10.1104/pp.109.137059

Stout, A. T., Davis, A. A., Domec, J. C., Yang, C. M., Shi, R., and King, J. S. (2014). Growth under field conditions affects lignin content and productivity in transgenic *Populus trichocarpa* with altered lignin biosynthesis. *Biomass Bioenergy* 68, 228–239. doi: 10.1016/j.biombioe.2014.06.008

Studer, M. H., Demartini, J. D., Davis, M. F., Sykes, R. W., Davison, B., Keller, M., et al. (2011). Lignin content in natural Populus variants affects sugar release. *Proc. Natl. Acad. Sci. U.S.A.* 108, 6300–6305. doi: 10.1073/pnas.1009252108

Sykes, R. W., Gjersing, E. L., Foutz, K., Rottmann, W. H., Kuhn, S. A., Foster, C. E., et al. (2015). Down-regulation of p-coumaroyl quinate/shikimate 3′-hydroxylase (C3'H) and cinnamate 4-hydroxylase (C4H) genes in the lignin biosynthetic pathway of Eucalyptus urophylla × E. grandis leads to improved sugar release. *Biotechnol. Biofuels* 8, 128–128. doi: 10.1186/s13068-015-0316-x

Tian, Q., Wang, X., Li, C., Lu, W., Yang, L., Jiang, Y., et al. (2013). Functional characterization of the poplar R2R3-Myb transcription factor Ptomyb216 involved in the regulation of lignin biosynthesis during wood formation. *PLoS One* 8:e76369. doi: 10.1371/journal.pone.0076369

Tian, X. M., Xie, J., Zhao, Y. L., Lu, H., Liu, S. C., Qu, L., et al. (2013). Sense-, antisense- and Rnai-4cl1 regulate soluble phenolic acids, cell wall components and growth in transgenic *Populus tomentosa* Carr. *Plant Physiol. Biochem.* 65, 111–119. doi: 10.1016/j.plaphy.2013.01.010

Tiimonen, H., Aronen, T., Laakso, T., Saranpää, P., Chiang, V., Ylioja, T., et al. (2005). Does lignin modification affect feeding preference or growth performance of insect herbivores in transgenic silver birch (Betula pendula Roth)? *Planta* 222, 699–708. doi: 10.1007/s00425-005-0002-5

Tuck, C. O., Perez, E., Horvath, I. T., Sheldon, R. A., and Poliakoff, M. (2012). Valorization of biomass: deriving more value from waste. *Science* 337, 695–699. doi: 10.1126/science.1218930

Upton, B. M., and Kasko, A. M. (2016). Strategies for the conversion of lignin to high-value polymeric materials: review and perspective. *Chem. Rev.* 116, 2275–2306. doi: 10.1021/acs.chemrev.5b00345

Van Acker, R., Déjardin, A., Desmet, S., Hoengenaert, L., Vanholme, R., Morreel, K., et al. (2017). Different routes for conifer- and sinapaldehyde and higher saccharification upon deficiency in the dehydrogenase CAD1. *Plant Physiol.* 175, 1018–1039. doi: 10.1104/pp.17.00834

Van Acker, R., Leplé, J.-C., Aerts, D., Storme, V., Goeminne, G., Ivens, B., et al. (2014). Improved saccharification and ethanol yield from field-grown transgenic poplar deficient in cinnamoyl-CoA reductase. *Proc. Natl. Acad. Sci. U.S.A.* 111:845. doi: 10.1073/pnas.1321673111

Van den Bosch, S., Schutyser, W., Vanholme, R., Driessen, T., Koelewijn, S.-F., Renders, T., et al. (2015). Reductive lignocellulose fractionation into soluble lignin-derived phenolic monomers and dimers and processable carbohydrate pulps. *Energy Environ. Sci.* 8, 1748–1763. doi: 10.1039/C5EE00204D

Van Doorsselaere, J., Baucher, M., Chognot, E., Chabbert, B., Tollier, M.-T., Petit-Conil, M., et al. (1995). A novel lignin in poplar trees with a reduced caffeic acid 5-hydroxyferulic acid O-methyltransferase activity. *Plant J.* 8, 855–864. doi: 10.1046/j.1365-313x.1995.8060855.x

Vanholme, B., Cesarino, I., Goeminne, G., Kim, H., Marroni, F., Van Acker, R., et al. (2013a). Breeding with rare defective alleles (BRDA): a natural *Populus nigra* Hct mutant with modified lignin as a case study. *New Phytol.* 198, 765–776. doi: 10.1111/nph.12179

Vanholme, B., Desmet, T., Ronsse, F., Rabaey, K., Van Breusegem, F., De Mey, M., et al. (2013b). Towards a carbon-negative sustainable bio-based economy. *Front. Plant Sci.* 4:174. doi: 10.3389/fpls.2013.00174

Vanholme, R., De Meester, B., Ralph, J., and Boerjan, W. (2019). Lignin biosynthesis and its integration into metabolism. *Curr. Opin. Biotechnol.* 56, 230–239. doi: 10.1016/j.copbio.2019.02.018

Vanholme, R., Demedts, B., Morreel, K., Ralph, J., and Boerjan, W. (2010). Lignin biosynthesis and structure. *Plant Physiol.* 153, 895–905.

Vanholme, R., Morreel, K., Darrah, C., Oyarce, P., Grabber, J. H., Ralph, J., et al. (2012). Metabolic engineering of novel lignin in biomass crops. *New Phytol.* 196, 978–1000. doi: 10.1111/j.1469-8137.2012.04337.x

Vargas, L., Cesarino, I., Vanholme, R., Voorend, W., Saleme, M. D. S., Morreel, K., et al. (2016). Improving total saccharification yield of Arabidopsis plants by vessel-specific complementation of caffeoyl shikimate esterase (cse) mutants. *Biotechnol. Biofuels* 9:16. doi: 10.1186/s13068-016-0551-9

Voelker, S. L., Lachenbruch, B., Meinzer, F. C., Jourdes, M., Ki, C. Y., Patten, A. M., et al. (2010). Antisense down-regulation of 4cl expression alters lignification, tree growth, and saccharification potential of field-grown poplar. *Plant Physiol.* 154, 874–886. doi: 10.1104/pp.110.159269

Wadenbäck, J., Von Arnold, S., Egertsdotter, U., Walter, M. H., Grima-Pettenati, J., Goffner, D., et al. (2008). Lignin biosynthesis in transgenic Norway spruce plants harboring an antisense construct for cinnamoyl CoA reductase (Ccr). *Transgenic Res.* 17, 379–392. doi: 10.1007/s11248-007-9113-z

Wagner, A., Donaldson, L., Kim, H., Phillips, L., Flint, H., Steward, D., et al. (2009). Suppression of 4-coumarate-CoA ligase in the coniferous gymnosperm *Pinus radiata. Plant Physiol.* 149, 370–383. doi: 10.1104/pp.108.125765

Wagner, A., Ralph, J., Akiyama, T., Flint, H., Phillips, L., Torr, K., et al. (2007). Exploring lignification in conifers by silencing hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyltransferase in *Pinus radiata. Proc. Natl. Acad. Sci.* 104, 11856–11861. doi: 10.1073/pnas.0701428104

Wagner, A., Tobimatsu, Y., Goeminne, G., Phillips, L., Flint, H., Steward, D., et al. (2013). Suppression of *CCR* impacts metabolite profile and cell wall composition in Pinus radiata tracheary elements. *Plant Mol. Biol.* 81, 105–117. doi: 10.1007/s11103-012-9985-z

Wagner, A., Tobimatsu, Y., Phillips, L., Flint, H., Geddes, B., Lu, F., et al. (2015). Syringyl lignin production in conifers: proof of concept in a Pine tracheary element system. *Proc. Natl. Acad. Sci. U.S.A.* 112:6218. doi: 10.1073/pnas.1411926112

Wagner, A., Tobimatsu, Y., Phillips, L., Flint, H., Torr, K., Donaldson, L., et al. (2011). CCoAOMT suppression modifies lignin composition in *Pinus radiata. Plant J.* 67, 119–129. doi: 10.1111/j.1365-313X.2011.04580.x

Wang, H. Z., Xue, Y. X., Chen, Y. J., Li, R. F., and Wei, J. H. (2012). Lignin modification improves the biofuel production potential in transgenic *Populus tomentosa. Indust. Crops Prod.* 37, 170–177. doi: 10.1016/j.indcrop.2011.12.014

Wang, J. P., Matthews, M. L., Williams, C. M., Shi, R., Yang, C., Tunlaya-Anukit, S., et al. (2018). Improving wood properties for wood utilization through multi-omics integration in lignin biosynthesis. *Nat. Commun.* 9:1579. doi: 10.1038/s41467-018-03863-z

Wei, J. H., Wang, Y., Wang, H., Li, R., Lin, N., Ma, R., et al. (2008). Pulping performance of transgenic poplar with depressed Caffeoyl-CoA O-methyltransferase. *Chin. Sci. Bull.* 53, 3553–3558. doi: 10.1007/s11434-008-0477-0

Weng, J.-K., Mo, H., and Chapple, C. (2010). Over-expression of F5H in COMT-deficient Arabidopsis leads to enrichment of an unusual lignin and disruption of pollen wall formation. *Plant J.* 64, 898–911. doi: 10.1111/j.1365-313X.2010.04391.x

Wilkerson, C. G., Mansfield, S. D., Lu, F., Withers, S., Park, J.-Y., Karlen, S. D., et al. (2014). Monolignol ferulate transferase introduces chemically labile linkages into the lignin backbone. *Science* 344, 90–93. doi: 10.1126/science.1250161

Xiang, Z. Y., Sen, S. K., Min, D. Y., Savithri, D., Lu, F. C., Jameel, H., et al. (2017). Field-grown transgenic hybrid poplar with modified lignin biosynthesis to improve enzymatic saccharification efficiency. *ACS Sustain. Chem. Eng.* 5, 2407–2414. doi: 10.1021/acssuschemeng.6b02740

Xie, M., Muchero, W., Bryan, A. C., Yee, K., Guo, H.-B., Zhang, J., et al. (2018). A 5-enolpyruvylshikimate 3-phosphate synthase functions as a transcriptional repressor in *Populus. Plant Cell* 30:1645. doi: 10.1105/tpc.18.00168

Xu, C., Fu, X., Liu, R., Guo, L., Ran, L., Li, C., et al. (2017). Ptomyb170 positively regulates lignin deposition during wood formation in poplar and confers drought tolerance in transgenic *Arabidopsis. Tree Physiol.* 37, 1713–1726. doi: 10.1093/treephys/tpx093

Yang, L., Zhao, X., Ran, L., Li, C., Fan, D., and Luo, K. (2017). Ptomyb156 is involved in negative regulation of phenylpropanoid metabolism and secondary cell wall biosynthesis during wood formation in poplar. *Sci. Rep.* 7:41209. doi: 10.1038/srep41209

Yin, T., Zhang, X., Gunter, L., Priya, R., Sykes, R., Davis, M., et al. (2010). Differential detection of genetic loci underlying stem and root lignin content in *Populus. PLoS One* 5:e14021. doi: 10.1371/journal.pone.0014021

Yoo, C. G., Dumitrache, A., Muchero, W., Natzke, J., Akinosho, H., Li, M., et al. (2018). Significance of lignin S/G ratio in biomass recalcitrance of *Populus trichocarpa* variants for bioethanol production. *ACS Sustain. Chem. Eng.* 6, 2162–2168. doi: 10.1021/acssuschemeng.7b03586

Zeng, Y., Zhao, S., Yang, S., and Ding, S.-Y. (2014). Lignin plays a negative role in the biochemical process for producing lignocellulosic biofuels. *Curr. Opin. Biotechnol.* 27, 38–45. doi: 10.1016/j.copbio.2013.09.008

Zhao, Q., and Dixon, R. A. (2014). "Altering the cell wall and its impact on plant disease: from forage to bioenergy," in *Annual Review of Phytopathology*, Vol. 52, ed. N. K. Vanalfen (Palo Alto: Annual Reviews).

Zhao, Q., Nakashima, J., Chen, F., Yin, Y., Fu, C., Yun, J., et al. (2013). Laccase is necessary and nonredundant with *Peroxidase* for lignin polymerization during vascular development in *Arabidopsis. Plant Cell* 25, 3976–3987. doi: 10.1105/tpc.113.117770

Zhong, R. Q., Morrison, W. H., Himmelsbach, D. S., Poole, F. L., and Ye, Z. H. (2000). Essential role of caffeoyl coenzyme A O-methyltransferase in lignin biosynthesis in woody poplar plants. *Plant Physiol.* 124, 563–577.

Zhou, S., Runge, T., Karlen, S. D., Ralph, J., Gonzales-Vigil, E., and Mansfield, S. D. (2017). Chemical pulping advantages of zip-lignin hybrid poplar. *ChemSusChem* 10, 3565–3573. doi: 10.1002/cssc.201701317

Zhou, X., Jacobs, T. B., Xue, L. J., Harding, S. A., and Tsai, C. J. (2015). Exploiting SNPS for biallelic CRISPR mutations in the outcrossing woody perennial *Populus* reveals 4-coumarate:CoA ligase specificity and redundancy. *New Phytol.* 208, 298–301. doi: 10.1111/nph.13470

Zhou, X., Ren, S., Lu, M., Zhao, S., Chen, Z., Zhao, R., et al. (2018). Preliminary study of cell wall structure and its mechanical properties of C3H and HCT RNAI transgenic poplar sapling. *Sci. Rep.* 8:10508. doi: 10.1038/s41598-018-28675-5

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Check for
updates

Engineering Non-cellulosic Polysaccharides of Wood for the Biorefinery

Evgeniy Donev[1], Madhavi Latha Gandla[2], Leif J. Jönsson[2] and Ewa J. Mellerowicz[1]*

[1] *Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, Umeå, Sweden,*
[2] *Department of Chemistry, Umeå University, Umeå, Sweden*

Non-cellulosic polysaccharides constitute approximately one third of usable woody biomass for human exploitation. In contrast to cellulose, these substances are composed of several different types of unit monosaccharides and their backbones are substituted by various groups. Their structural diversity and recent examples of their modification in transgenic plants and mutants suggest they can be targeted for improving wood-processing properties, thereby facilitating conversion of wood in a biorefinery setting. Critical knowledge on their structure-function relationship is slowly emerging, although our understanding of molecular interactions responsible for observed phenomena is still incomplete. This review: (1) provides an overview of structural features of major non-cellulosic polysaccharides of wood, (2) describes the fate of non-cellulosic polysaccharides during biorefinery processing, (3) shows how the non-cellulosic polysaccharides impact lignocellulose processing focused on yields of either sugars or polymers, and (4) discusses outlooks for the improvement of tree species for biorefinery by modifying the structure of non-cellulosic polysaccharides.

Keywords: non-cellulosic polysaccharides, woody biomass, secondary cell wall, hemicellulose, pectin, galactan, tree genetic improvement, wood biorefining

NON-CELLULOSIC POLYSACCHARIDES IN WOOD

Plant cell walls constitute the most abundant renewable resource on Earth (Pauly and Keegstra, 2008). Wood, which essentially consists of cell walls, is naturally degradable and renewable, and technologies are currently being developed aiming at utilization of all organic wood components, i.e., cellulose, lignin, non-cellulosic polysaccharides, and extractives. Non-cellulosic polysaccharides, which include hemicelluloses, pectins, type II arabinogalactan (AG-II), and callose, account for roughly one third of wood dry weight. Among these four groups, hemicelluloses are most abundant contributing to 26–33% of the dry weight in softwoods (conifer species) and 19–34% in hardwoods (dicotyledonous species) depending on species, cell type, developmental stage, and environmental conditions (Sjöström, 1993).

Non-cellulosic polysaccharides exhibit remarkable variability in different layers of wood cell walls, and chiefly define these layers (Mellerowicz and Gorshkova, 2012). Thus, the middle lamella

Abbreviations: AG-II, type II arabinogalactan; CML, compound middle lamella; GGM, galactoglucomannan; G-layer, gelatinous layer; GM, glucomannan; GT, glycosyltransferase family; GX, glucuronoxylan; HG, homogalacturonan; IRX, irregular xylem; KD, knock-down; LCCs, lignin carbohydrate complexes; PCW, primary cell wall; RES, reducing end sequence in xylan; RG-I, rhamnogalacturonan I; RG-II, rhamnogalacturonan II; S/G, syringyl/guaiacyl; SCW, secondary cell wall; Xyl*p*, xylopyranosyl residue.

is dominated by pectins, the primary cell wall (PCW) layer by pectins and xyloglucan, the secondary cell wall (SCW) layers by xylans and mannans, and the gelatinous layer (G-layer) present as a tertiary layer in tension wood fibers by galactans and AG-II. These different polysaccharides blend with the lignin matrix and cellulose microfibrils, and are involved in covalent, ionic, and hydrophobic interactions with other cell wall components and with themselves (Cosgrove, 2005; Scheller and Ulvskov, 2010). They are the main source of lignin carbohydrate complexes (LCCs) linking to lignin by ether, glycoside, or ester bonds (Lawoko and Henriksson, 2006; Balakshin et al., 2007; Giummarella and Lawoko, 2016; Giummarella et al., 2016). Thereby, the non-cellulosic polysaccharides affect cell wall architecture, wood traits, and properties of lignocellulosic biomass being favorite targets for improving biomass properties (reviewed in Tavares et al., 2015; Damm et al., 2016; Marriott et al., 2016; Wang et al., 2016; Smith et al., 2017). They are also a precious source of large amounts of assimilated carbon for which clever applications are being sought (e.g., Zhao et al., 2015; Oinonen et al., 2016).

Among hemicelluloses, xylan, which includes the glucuronoxylan (GX) of hardwoods and the arabinoglucuronoxylan of softwoods (**Figure 1**), is a ubiquitous component of wood SCWs (Donaldson and Knox, 2012; Kim and Daniel, 2012). Approx. 60% of the xylopyranosyl residue (Xyl$p$) of hardwood xylan are mono- or di-acetylated (Teleman et al., 2000, 2002). The acetyl groups compete with glucuronic acid for Xyl$p$ position 2, and a decrease of one of these substituents usually leads to an increase of the other (Chong et al., 2014; Lee et al., 2014). Mannans, which include water-soluble galactoglucomannan (GGM) and water-insoluble glucomannan (GM) (**Figure 1**), are the most abundant hemicelluloses in softwood SCWs, whereas hardwood SCWs contain lower fractions of GM (Teleman, 2009). Xyloglucan (**Figure 1**) is localized in PCWs of hardwoods and softwoods (Bourquin et al., 2002; Donaldson and Knox, 2012; Kim and Daniel, 2013), where it may associate with hydrophobic cellulose surfaces or become entrapped inside cellulose fibrils (Park and Cosgrove, 2015).

Pectins, which include homogalacturonan (HG), rhamnogalacturonan I (RG-I), and rhamnogalacturonan II (RG-II) (**Figure 1**), are acidic polysaccharides. They constitute a large part of the middle lamella and PCW layers, jointly referred to as the compound middle lamella (CML) (Kim and Daniel, 2013). Reaction wood, such as tension wood of hardwood and compression wood of softwood, typically contains high mass fractions of β-1,4-galactans (**Figure 1**) presumably associated with RG-I.

Water-soluble softwood arabinogalactan, a variant of AG-II (**Figure 1**), is highly abundant in larch (25%). Other softwoods and hardwoods contain small amounts of AG-II. AG-II may be covalently linked to xylan and pectin (Tan et al., 2013). Callose (**Figure 1**), or laricinan, accumulates in hardwoods and softwoods in response to damage and stress (Teleman, 2009). It is abundant in pits and between cavities of the inner S2 layer in compression wood (Hoffman and Timell, 1970; Chaffey and Barlow, 2002; Altaner et al., 2010; Zhang et al., 2016).

This review addresses the importance of the non-cellulosic polysaccharides in technological processes currently used in wood biorefining, and prospects of altering them in trees for obtaining either higher productivity or improved lignocellulose properties, like extractability or improved biochemical conversion to sugars.

FATE OF NON-CELLULOSIC POLYSACCHARIDES DURING WOOD BIOREFINING

Biorefining of wood includes the pulping (mechanical and chemical pulping, as well as combinations thereof), biochemical processes, and thermochemical processes. Mechanical pulping aims at high recovery of all major wood constituents, including non-cellulosic polysaccharides (Sjöström, 1993; Ek et al., 2009). In contrast, chemical pulping and subsequent bleaching steps are designed to target the lignin and preserve the cellulose, whereas the fate of the non-cellulosic polysaccharides is strongly dependent on the aim and the process technology, which can be Kraft (sulfate), sulfite, soda, or organosolv pulping (Sjöström, 1993; Ek et al., 2009). For example, dissolving pulp manufacture based on sulfite pulping, or, sometimes, Kraft pulping, aims at producing a cellulose of relatively high purity, which means that most of the non-cellulosic polysaccharides are degraded and removed together with the lignin. In many other processes, such as conventional Kraft pulping for manufacturing of paper products, preservation of hemicelluloses is beneficial, as the pulp yield will then be higher.

Biochemical conversion is typically based on saccharification of the polysaccharides using pretreatment and enzymatic hydrolysis. This creates monosaccharides, which are then refined further using microbial fermentation or chemical catalysis. The aim of the pretreatment is to make the cellulose susceptible to enzymatic hydrolysis, which otherwise would be too slow and provide too low sugar yields. Among many different pretreatment methods (Yang and Wyman, 2008; Sun et al., 2016; Jönsson and Martín, 2016), hydrothermal pretreatment under acidic conditions, with or without externally added acid and with or without the disruptive effect of steam explosion, is a common approach. Due to autohydrolysis and formation of carboxylic acids, which are derived mainly from the non-cellulosic polysaccharides, the process will be acidic even without externally added acid (Jönsson and Martín, 2016). The main target of hydrothermal pretreatment under acidic conditions is the hemicellulose. Cellulose and lignin are also affected, but typically to much lesser degree than hemicelluloses, which can be degraded almost quantitatively in well-optimized pretreatment (Wang et al., 2018). The severity of the hydrothermal pretreatment (time, temperature, and acidity) needs to be adapted to the feedstock. Softwoods require higher severity, whereas hardwoods can be processed using lower severity. Nevertheless, pressurized reactors and temperatures in the range 160–240°C (Sun et al., 2016) are typically used to create a pretreated material that is suitable for subsequent enzymatic saccharification.

Thermochemical conversion processes, such as combustion, gasification, and pyrolysis, will degrade all organic wood constituents and are less relevant within the context of the current review.
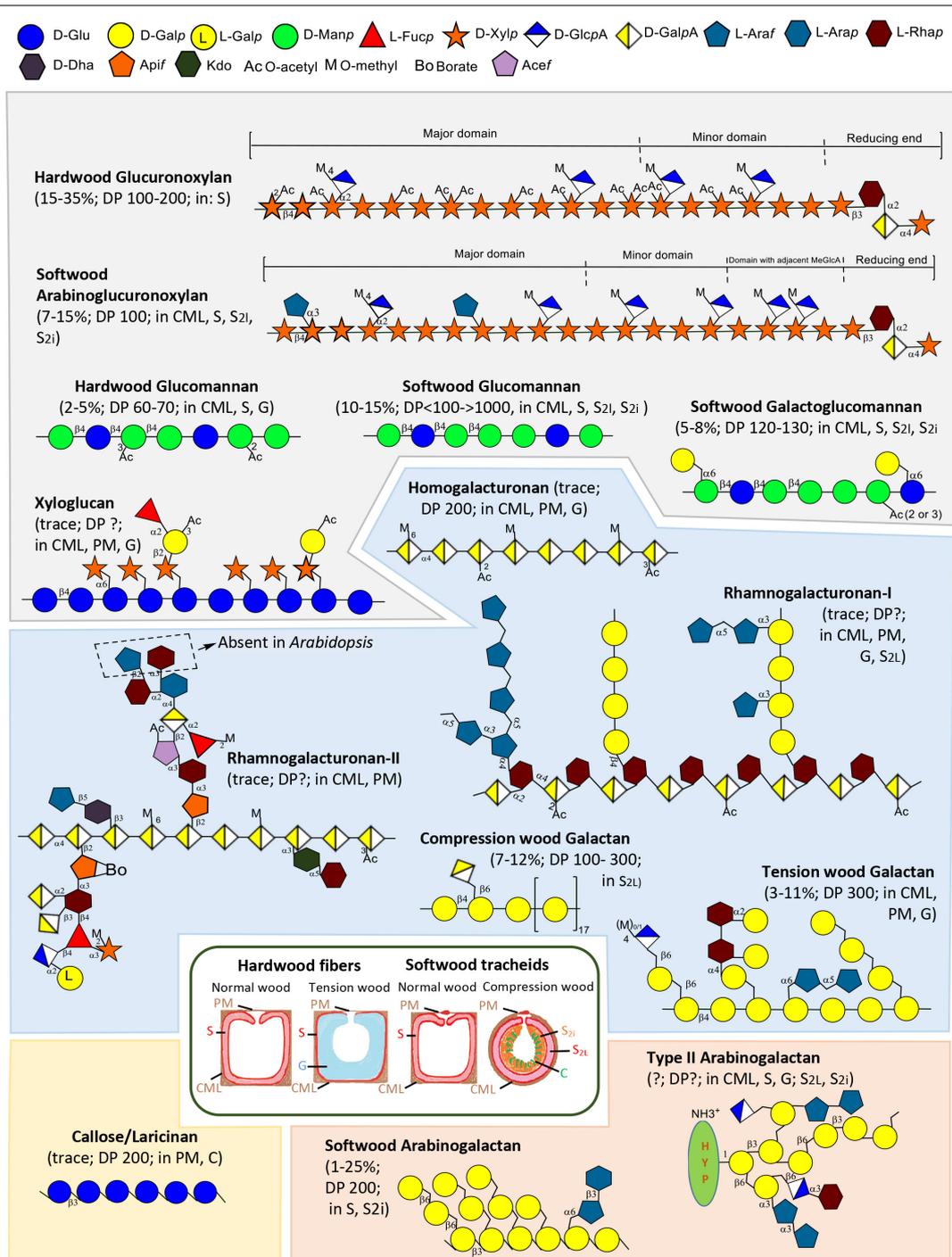
FIGURE 1 | Schematic illustration of types of non-cellulosic polysaccharides of wood, including hemicelluloses (gray background), pectins (blue background), callose (yellow background) and AGs-II (orange background), and hardwood fibers and softwood tracheids (inset). Polymer structures were based on different sources: hardwood GX (Teleman, 2009; Smith et al., 2017), softwood arabinoglucuronoxylan (Teleman, 2009; Martínez-Abad et al., 2017; Smith et al., 2017), hardwood and softwood glucomannan (GM), softwood GGM, tension and compression wood galactans, callose (Teleman, 2009), xyloglucan (Carpita and McCann, 2000; Teleman, 2009), HG (Atmodjo et al., 2013), RG-I and -II (Edashige and Ishii, 1996, 1997, 1998; Atmodjo et al., 2013), AG-II (Carpita and McCann, 2000; Hijazi et al., 2014), softwood arabinogalactan (Ponder and Richards, 1997; Teleman, 2009). Polymer localization is based on the following sources: hardwood GX and mannans (Kim and Daniel, 2012; Gorshkova et al., 2015; Guedes et al., 2017), softwood arabinoglucuronoxylan (Altaner et al., 2010; Donaldson and Knox, 2012), callose (Altaner et al., 2010; Zhang et al., 2016), xyloglucan (Bourquin et al., 2002; Sandquist et al., 2010; Nishikubo et al., 2011; Donaldson and Knox, 2012; Kim and Daniel, 2013; Guedes et al., 2017), HG (Kim and Daniel, 2013), RG-I /compression wood galactan/tension wood galactan (Gorshkova et al., 2015; Zhang et al., 2016; Guedes et al., 2017), AG-II/softwood arabinogalactan (Altaner et al., 2010; Guedes et al., 2017). PM, pit membrane; CML, compound middle lamella; S, secondary wall layer (S-layer), G, gelatinous layer (G-layer); C, cavities; S_{2i}, inner S_2 layer; S_{2L}, outer lignified S_2 layer.

ROLE OF NON-CELLULOSIC POLYSACCHARIDES IN RECALCITRANCE AND ATTEMPTS TO IMPROVE CONVERTIBILITY

Improvement of Xylan Structure

Xylan Content and Length Affect Saccharification and Plant Growth

Xylan is a key factor of recalcitrance, mainly by reducing cellulose accessibility (Bura et al., 2009; De Martini et al., 2013), prompting efforts to reduce its content in hardwoods. Attempts have been made using *Arabidopsis* as a model, either by targeting the xylan synthase complex or the enzymes synthesizing the reducing end sequence (RES) (Smith et al., 2017). However, strong reductions in xylan content led to mechanical failure of vessels [the so-called irregular xylem phenotype (IRX)] and stem weakening. Subsequent work with *Populus* indicated that xylan content can be reduced by 5–50% by knocking down (KD) different xylan biosynthetic genes including *GT47C* (Lee et al., 2009), two *GAUT12/GT8D* paralogs (Lee et al., 2011; Li et al., 2011; Biswal et al., 2015), *GT43A* and *GT43B* (Lee et al., 2011), and *GT43B* along with *GT43C* clade genes (Ratke et al., 2018) or by expressing fungal xylanase *HvXyl1* and targeting it to cell walls (Kaida et al., 2009; **Table 1**). Such reductions either did not affect (Lee et al., 2009) or stimulated growth (Biswal et al., 2015; Ratke et al., 2018), decreased xylem cell wall thickness (Li et al., 2011; Ratke et al., 2018), which sometimes (Lee et al., 2009, 2011) was coupled to a mild IRX phenotype. Beside xylan, the cellulose content decreased in case of *GT47C* KD, coupled with increased lignification (Lee et al., 2009). A similar increase in lignin coupled with brittleness of stems was observed in strong KD *GAUT12-1* and *-2* (Li et al., 2011), but not when only *GAUT12-1* was reduced (Biswal et al., 2015). Lignin syringyl/guaiacyl (S/G) ratio was increased in *GAUT12-1* KD poplar (Biswal et al., 2015), and in severe KD *GT43B* poplar (Lee et al., 2011). However, the S/G ratio was reduced without change in lignin content in mildly affected *GT43B* and *GT43C* KD (Ratke et al., 2018). Thus, it is difficult to predict how lignin might be affected in transgenic lines with reduced xylan content, and these changes should be carefully monitored, since they affect saccharification.

For non-pretreated wood, downregulation of *GT43* genes resulted in a 30% increase in glucose yield in enzymatic saccharification (Lee et al., 2011; Ratke et al., 2018), but the benefits were less apparent after acid pretreatment (Ratke et al., 2018; **Table 1**). Reductions in xylan content by downregulation of *GAUT12/GT8D* did not improve saccharification without pretreatment (Lee et al., 2011) or did so only slightly (after steam pretreatment) (Biswal et al., 2015) whereas post-synthetic xylan reduction resulted in approx. 50% increased glucose yield in saccharification after steam pretreatment (Kaida et al., 2009; **Table 1**). There might be several and possibly opposing factors at play affecting saccharification. For example, cell wall thickness, lignin content and composition, and content of tension wood can all affect glucose yields (Escamez et al., 2017). KD *GAUT12/GT8D* poplar had less LCCs, which contribute to recalcitrance (Min et al., 2014a). Clearly, manipulation of xylan induces indirect

effects, some of which, such as increased growth (Biswal et al., 2015; Derba-Maceluch et al., 2015; Yang et al., 2017; Ratke et al., 2018), or increased drought tolerance (Keppler and Showalter, 2010) are interesting for biotechnological applications.

Xylan Acetylation Affects Cell Wall Architecture

Deacetylation of lignocellulosic biomass prior to enzymatic saccharification results in improved sugar yields (reviewed by Pawar et al., 2013). For lignocellulosic biomass with high acetyl content such as hardwoods, reduction of acetylation might have an added benefit for ethanolic fermentation processes, as high concentrations of acetic acid are inhibitory to fermenting microorganisms (reviewed by Jönsson and Martín, 2016).

Modest reductions in acetylation in KD *RWA* aspen (Pawar et al., 2017b) and in aspen expressing fungal xylan acetyl esterase *An*AXE1 targeted to cell walls (Pawar et al., 2017a; **Table 1**) were well tolerated by plants. These plants yielded 20–25% more glucose in enzymatic saccharification without pretreatment. Results with *Arabidopsis* (Pawar et al., 2016) suggested that deacetylation *in planta* reduces recalcitrance by other mechanism than reducing the inhibition of xylan enzymatic hydrolysis by acetyl groups. Indeed, aspen expressing *An*AXE1 exhibited increased lignin solubility and reduced xylan content, xylan chain length, and lignin S/G ratio (Pawar et al., 2017a). Increased extractability of lignin and xylan agrees with the suggested xylan models (Ruel et al., 2006; Busse-Wicher et al., 2014), where the minor xylan domain (**Figure 1**) interacts with lignin. This domain, characterized by consecutive Xyl*p* acetylation, would become (after deacetylation) more susceptible to hydrolysis by wall-residing enzymes, such as XYN10A (Derba-Maceluch et al., 2015), leading to loosening of xylan fraction interacting with lignin.

Overexpression of *PdDUF231A*, similar to *AtPMR5*, in *P. deltoides* resulted in increased xylan acetylation and surprisingly in improved saccharification without pretreatment (Yang et al., 2017; **Table 1**). Decreased lignin content and increased cellulose content in transgenic lines might have affected the sugar yield.

Significance of Glucuronosylation

Glucuronosylation of xylan makes its backbone less prone to hydrolysis by GH10 and GH11 xylanases, and requires α-glucuronidases for saccharification (Mortimer et al., 2010). It is also associated with LCCs of hardwoods (Min et al., 2014a,b; Bååth et al., 2016). The majority of glucuronate in SCW xylan is methylated (Teleman, 2009) and KD of a GX methyl transferases *DUF579-3/GXM3* in poplar reduced not only methylation but also resulted in reduced xylan glucuronosylation, and reduced growth (Song et al., 2016). Xylose yield of acid pretreatment increased as well as glucose yield of enzymatic saccharification. Thus, methylation of glucuronate is a promising target, but some means of avoiding growth penalty need to be designed.

To reduce ester links between GX and lignin in aspen, a fungal glucuronoyl esterase was expressed and targeted to cell walls (Gandla et al., 2015; **Table 1**). Increased cellulose-to-glucose conversion was observed, but plants exhibited premature leaf senescence and impaired growth (Gandla et al., 2015). There

TABLE 1 | Effects of engineering expression of genes encoding proteins affecting non-cellulosic polysaccharides on enzymatic saccharification and cell wall composition of woody species.[a]

Gene	Approach	Species	Protein	Protein family	Promoter	Role of protein	Saccharification efficiency	Phenotypic effect	Effect on cell wall composition	Reference
PtxtPME1	Down- or up-regulation	Hybrid aspen	Pectin methyl esterase (PME)	CE8	35S	Removal of methyl ester groups from HG	EH(+), AEH(−) in down-regulated lines	In down-regulated lines: Height(ND) diameter(+)	Up-regulation resulted in (−) in DM while downregulation resulted in (+) DM and (−) in the amount of galacturonan.	Siedlecka et al., 2008; Escamez et al., 2017
PoGT47C	RNAi down-regulation	Hybrid poplar	Glycosyl transferase	GT47	35S	Probably biosynthesis of RES in X	EH(+)	Growth(ND)	S/G(+), Xyl content(−), Glu content(−), GX(−), thickness of cell walls(−), RES of GX (−)	Lee et al., 2009
AaXEG2, HvXYL1, Tv6GAL, AtCel1	Over-expression	Poplar (P. alba)	Xyloglucanase, xylanase, galactanase, and glucanase	GH5, GH10, GH35, and GH9	35S	Hydrolysis of XG, X, Ga, and GM	EH(+) for AaXEG2, HvXYL1 and AtCel1, EH(−) for Tv6GAL	Growth(+) for AaXEG2	L content (ND) for HvXYL1 and AtCel1, (−) for AaXEG2.	Park et al., 2004; Kaida et al., 2009
AaXEG2	Over-expression	Acacia mangium	Xyloglucanase	GH5	35S	Hydrolysis of XG	EH(+)	Growth(+)	Cellulose(+), L(−),wall bound XG in the xylem(−)	Hartati et al., 2011; Kaku et al., 2011
PtxXET16A/ PtxXTH34	Up-regulation	Hybrid aspen	Xyloglucan endotransglucosylase (XET)	GH16	35S	Hydrolysis and re-ligation of XG	EH(ND), AEH(ND)	Growth(ND)	Loosely bound XG fraction(−), tightly bound XG fraction(+), PCW XG(+)	Nishikubo et al., 2011; Escamez et al., 2017
PoGT43B, PtxtGT43B and PtxtGT43C	RNAi down-regulation	Poplar (P. trichocarpa), hybrid aspen	Glycosyl transferases	GT43	35S and GT43B	Scaffolding in X synthase complex, X backbone biosynthesis	EH(+), AEH(ND)	Growth(+) for GT43B promoter	Xyl content relative to RES (Xyl:RES) in X(−), Cell wall thickness(−), wood density(−), MFA(−), S/G(−).	Lee et al., 2011; Ratke et al., 2018
PtxtPL1-27	Up-regulation	Hybrid aspen	Pectate lyase (PEL)	PL1	35S	Cleavage of HG backbone	EH(+) (for release of Xyl), AEH(+)	Growth(−)	L(+), M(−), TFA soluble Glu (+), solubility of pectin, X, and XG(+)	Biswal et al., 2014
PtGAUT12-1/PtGT8D-1, PtGAUT12-2/PtGT8D-2	RNAi down-regulation	Poplar (P. trichocarpa, P.deltoides)	Galacturonosyl transferase	GT8	35S	Biosynthesis of pectin and RES in X	EH(+)	Growth(ND or +)	L (+), mechanical strength (−), S/G(+), Xyl(−), GalA(−), pectin(−), X(−), Rha(−)	Lee et al., 2011; Li et al., 2011; Biswal et al., 2015

(Continued)

TABLE 1 | Continued

Gene	Approach	Species	Protein	Protein family	Promoter	Role of protein	Saccharification efficiency	Phenotypic effect	Effect on cell wall composition	Reference
PcGCE	Over-expression	Hybrid aspen	Glucuronoyl esterase	CE15	35S	Hydrolysis of ester linkages between MeGlcA and L	EH(−), AEH(−), Cellulose conversion (+)	Growth(-)	L(+), C(−), S/G(+), Ara(−), Rha(−), MeGlcA(−), Compositional changes of pectin.	Gandla et al., 2015
PtxXyn10A	Antisense down-regulation	Hybrid aspen	Xylan endotransglycosylase	GH10	35S	Hydrolysis and re-ligation of X chains	EH(ND), AEH(ND)	Growth(+)	Ga(−), X(+), TW content in stem(−), MFA(−)	Derba-Maceluch et al., 2015; Escamez et al., 2017
PtrDUF579-3	Antisense down-regulation	Poplar (P. trichocarpa)	GX methyl transferase (GXM3)	GT8	35S	Methylation of GlcA in GX biosynthesis	AEH(+)	Growth(ND)	Mechanical strength of stem(−), methylation of GlcA in GX(−)	Song et al., 2016
PdDUF231A	Up-regulation	Poplar (P. deltoides)	Acyl transferase/esterase	DUF231	UBI3	Probably acetylation of X	EH(+)	Growth(+)	L(−), Cellulose(+), acetylated X at O-3(+)	Yang et al., 2017
PtxRWA (AB and CD clades)	RNAi suppression of RWA family	Hybrid aspen	RWA-A, RWA-B, RWA-C, and RWA-D	RWA	GT43B	Probably transport of acetyl-CoA to Golgi	EH(+), AEH(+)	Growth(ND)	Total acetylation and X acetylation at position O2(−), MeGlcA(+)	Pawar et al., 2017b
AnAXE1	Over-expression	Hybrid aspen	Acetylxylan esterase (AX)	CE1	35S	Deacetylation of X and possibly other polymers	EH(+), AEH(+)	Growth(ND)	Total acetylation and acetylation of X at position O2 (-). X chain length and size of L polymer(−). L solubility (+), S/G(−)	Pawar et al., 2017a
PdGAUT4	RNAi down-regulation	Poplar (P. deltoides)	α-1,4-Galacturonosyl transferase	GT8	35S	Probably synthesis of HG	HWEH(+) (Glu yield)	Growth(+)	S/G(+), HG(−), RGII(−), GalA(−), Calcium and boron in the cell wall(−). HG and RGII cross-link in cell wall(−)	Biswal et al., 2018

[a]AEH, saccharification efficiency of enzymatic hydrolysis after acid pretreatment; Ara, arabinose; C, carbohydrate; DM, degree of methylation; EH, saccharification efficiency of enzymatic hydrolysis without pretreatment; Ga, galactan; GalA, galacturonic acid; Glu, glucose; GlcA, glucuronic acid; GM, glucomannan; GX, glucuronoxylan; HG, homogalacturonan; Hybrid poplar, Populus alba × tremula; HWEH, saccharification efficiency of enzymatic hydrolysis after hot water pretreatment; L, lignin; M, mannan; MeGlcA, 4-O-methylglucuronic acid; MFA, microfibril angle; ND, No detectable effect compared to WT; RGII, Rhamnogalacturonan II; Rha, rhamnose; RES, reducing end sequence of xylan; S/G, syringyl:guaiacyl ratio; TFA, trifluoroaceticacid; TW, tension wood; X, xylan; Xyl, xylose; XG, xyloglucan; (+), increase compared to WT; (−), decrease compare to WT.

was a decrease in extractives and an increase in lignin (Gandla et al., 2015) and mechanisms underlying these responses are not understood.

Prospects for Mannan Structure Improvement

GGM is tightly associated with cellulose microfibrils (Schröder et al., 2009; Salmén, 2015) and increasing its extractability could be beneficial for saccharification. Mannan hydrolysis *in planta* by overexpression of plasma-membrane-bound mannanase MAN6 induced production of active galactoglucomanno-oligosaccharides that modified growth of poplar and inhibited SCW formation (Zhao et al., 2013), making this approach problematic. However, overexpression of extracellular GM-active endoglucanase *At*Cel1 in poplar lead to approx. 30% increase in sugar yield and in cellulose conversion in saccharification after steam pretreatment (Kaida et al., 2009).

Mannan biosynthetic genes identified to date include a GM synthase (CSLA clade) (Suzuki et al., 2006), a mannan galactosyltransferase (GT34) (Edwards et al., 2004), and unknown function GT65 members *At*MSR1 and 2 (Wang et al., 2013), and could be employed for biomass modification. For example, increasing the degree of galactosylation of mannan by increasing expression of mannan galactosyltransferase might lead to higher solubility of GGM (Edwards et al., 2004).

Xyloglucan and Pectins – Minor Wood Components With High Impact on Properties

Several lines of evidence indicate that xyloglucan and pectins, despite being minor wood components (below 5% dry weight), have significant effects on wood properties, including recalcitrance. Increased xyloglucan content in aspen obtained by overexpressing *XTH34/XET16A* did not affect growth (Nishikubo et al., 2011) nor saccharification (Escamez et al., 2017), but its reduction in poplar, achieved by expression of fungal xyloglucanase *AaXEG2*, highly stimulated growth, wood cellulose content, density, and mechanical strength in the greenhouse-grown poplar (Park et al., 2004; **Table 1**). Cellulose microfibrils were larger in transgenic plants (Yamamoto et al., 2011) and their lignocellulose yielded almost 50% more glucose per wood weight and per cellulose weight in saccharification after steam pretreatment (Kaida et al., 2009). Similar effects on growth and saccharification were observed in *Acacia mangium* (Kaku et al., 2011; Hartati et al., 2011). However, in a 4-year field trial transgenic poplars expressing *AaXEG2* displayed substantially reduced biomass (Taniguchi et al., 2012). Furthermore, the plants were unable to bend upward when placed horizontally despite forming tension wood as in wild type, suggesting that xyloglucan is essential for generation of tensile stress (Baba et al., 2009).

Suppression of aspen pectin methyl esterase PME1 resulted in highly methylesterified HG in developing wood (Siedlecka et al., 2008; **Table 1**). The glucose yields in saccharification without pretreatment increased, but there was no improvement after acid pretreatment (Escamez et al., 2017).

Overexpression of aspen wood-expressed pectate lyase PL1-27 increased the solubility of xylan and xyloglucan suggesting that HG constrains the solubility of main non-cellulosic polysaccharides (Biswal et al., 2014; **Table 1**). There was a positive effect on the glucose yield for the transgenic lines, but only after acid pretreatment. Interestingly, decreased HG biosynthesis in poplar by KD of the *GAUT* genes involved in biosynthesis of pectin and xylan (Mohnen, 2008), led to substantial growth stimulation and a small increase in glucose yields in saccharification after acid or hot-water pretreatment (Biswal et al., 2015; Biswal et al., 2018; **Table 1**).

FUTURE PROSPECTS FOR IMPROVING NON-CELLULOSIC POLYSACCHARIDES FOR BIOREFINING OF WOOD

Two decades of research on modifying non-cellulosic polysaccharides have provided some insight on the role of these polysaccharides in cell wall architecture, on their importance for the efficiency of pretreatment and enzymatic saccharification, and have identified some off-target effects. They have also identified the most promising targets for achieving better growth and saccharification. Most research has been focused on xylan modification identifying xylan chain length, degree of acetylation, glucuronosylation, and glucuronosyl methylation as possible targets. The discovery of microbial enzymes hydrolyzing ester links between glucuronosyl units and lignin opened new prospects for directly reducing LCCs in cell walls, and should be further explored. The modification of minor wood components, HG and xyloglucan, had some of the highest impacts on saccharification, pointing to the crucial role of these polymers in wood integrity, but their modification sometimes led to growth inhibition. Using the wood-specific promoters, such as the *GT43B* promoter (Ratke et al., 2015), for expressing transgenes, can prevent off-target modifications in meristems, root hairs and other primary walled tissues, and possibly avoid growth penalty. Overall, there were few attempts in trees to employ different promoters for expressing the transgenes. Utilization of heat inducible promoters and heat-active enzymes in trees for modifying properties during post-harvest heat treatment has not yet been explored. Modification of other non-cellulosic components, for example mannans, RG-I and RG-II has not been so far investigated in trees and will be interesting to reveal the role of these polymers in woody biomass organization.

Currently, there is little understanding of the molecular mechanisms responsible for the observed phenotypes. In most cases, the distinction between primary and secondary effects of transgene expression is not possible. Interestingly, some types of xylan modification lead to increased growth (Biswal et al., 2015; Derba-Maceluch et al., 2015; Ratke et al., 2018), which might be mediated by the SCW integrity sensing mechanism (Ratke et al., 2018). On the other hand, the transcriptome analyses in *Arabidopsis* mutants impaired in xylan biosynthesis did not reveal any changes indicative of SCW integrity sensing (Faria-Blanc et al., 2018). It would be important to elucidate if such signaling

exists and if so, what triggers it, for successful modification of SCW.

Almost all results reviewed here are based on greenhouse experiments. Experience with xyloglucanase-modified poplar (Taniguchi et al., 2012) points to a need for early field experiments to pinpoint possible problems of transgenic lines. Field-grown trees will also provide sufficient biomass for testing pulping properties.

Finally, the tension wood of hardwoods appears to be particularly suitable for saccharification (Brereton et al., 2012; Sawada et al., 2018). Progress in identification of pathways involved in the induction of tension wood (Felten et al., 2018) will make it possible to design strategies to stimulate tension wood formation in normal growth conditions for dedicated biorefinery feedstocks.

AUTHOR CONTRIBUTIONS

ED, LJJ, and EJM wrote the paper. MLG prepared the Figure and the Table. All authors read and approved the final submission.

FUNDING

REFERENCES

Altaner, C. M., Tokareva, E. N., Jarvis, M. C., and Harris, P. J. (2010). Distribution of (1→4)-β-galactans, arabinogalactan proteins, xylans and (1→3)-β-glucans in tracheid cell walls of softwoods. *Tree Physiol.* 30, 782–793. doi: 10.1093/treephys/tpq021

Atmodjo, M. A., Hao, Z., and Mohnen, D. (2013). Evolving views of pectin biosynthesis. *Annu. Rev. Plant Biol.* 64, 747–779. doi: 10.1146/annurev-arplant-042811-105534

Bååth, J., Giummarella, N., Klaubauf, S., Lawoko, M., and Olsson, L. (2016). A glucuronoyl esterase from *Acremonium alcalophilum* cleaves native lignin-carbohydrate ester bonds. *FEBS Lett.* 590, 2611–2618. doi: 10.1002/1873-3468

Baba, K., Park, Y. W., Kaku, T., Kaida, R., Takeuchi, M., Yoshida, M., et al. (2009). Xyloglucan for generating tensile stress to bend tree stem. *Mol. Plant* 2, 893–903. doi: 10.1093/mp/ssp054

Balakshin, M., Capanema, E. A., and Chang, H. (2007). MWL fraction with a high concentration of lignin-carbohydrate linkages: isolation and 2D NMR spectroscopic analysis. *Holzforschung* 61, 1–7. doi: 10.1515/HF.2007.001

Biswal, A. K., Atmodjo, M. A., Li, M., Baxter, H. L., Yoo, C. G., Pu, Y., et al. (2018). Sugar release and growth of biofuel crops are improved by downregulation of pectin biosynthesis. *Nat. Biotechnol.* 36, 249–257. doi: 10.1038/nbt.4067

Biswal, A. K., Hao, Z., Pattathil, S., Yang, X., Winkeler, K., Collins, C., et al. (2015). Downregulation of *GAUT12* in *Populus deltoides* by RNA silencing results in reduced recalcitrance, increased growth and reduced xylan and pectin in a woody biofuel feedstock. *Biotechnol. Biofuels* 18:41. doi: 10.1186/s13068-015-0218-y

Biswal, A. K., Soeno, K., Gandla, M. L., Immerzeel, P., Pattathil, S., Lucenius, J., et al. (2014). Aspen pectate lyase *PtxtPL1-27* mobilizes matrix polysaccharides from woody tissues and improves saccharification yield. *Biotechnol. Biofuels* 7:11. doi: 10.1186/1754-6834-7-11

Bourquin, V., Nishikubo, N., Abe, H., Brumer, H., Denman, S., Eklund, M., et al. (2002). Xyloglucan endotransglycosylases have a function during the formation of secondary cell walls of vascular tissues. *Plant Cell* 14, 3073–3088. doi: 10.1105/tpc.007773

Brereton, N. J., Ray, M. J., Shield, I., Martin, P., Karp, A., and Murphy, R. J. (2012). Reaction wood - A key cause of variation in cell wall recalcitrance in willow. *Biotechnol. Biofuels* 5:83. doi: 10.1186/1754-6834-5-83

Bura, R., Chandra, R., and Saddler, J. (2009). Influence of xylan on the enzymatic hydrolysis of steam-pretreated corn stover and hybrid poplar. *Biotechnol. Prog.* 25, 315–322. doi: 10.1002/btpr.98

Busse-Wicher, M., Gomes, T. C. F., Tryfona, T., Nikolovski, N., Stott, K., Grantham, N. J., et al. (2014). The pattern of xylan acetylation suggests xylan may interact with cellulose microfibrils as a twofold helical screw in the secondary plant cell wall of *Arabidopsis thaliana*. *Plant J.* 79, 492–506. doi: 10.1111/tpj.12575

Carpita, N., and McCann, M. (2000). "The Cell Wall," in *Biochemistry and Molecular Biology of Plants*, eds B. B. Buchanan, W. Gruissem, and R. L. Jones (Rockville, MD: American Society of Plant Biologists), 52–108.

Chaffey, N., and Barlow, P. (2002). Myosin, microtubules, and microfilaments: co-operation between cytoskeletal components during cambial cell division and secondary vascular differentiation in trees. *Planta* 214, 526–536. doi: 10.1007/s004250100652

Chong, S. L., Virkki, L., Maaheimo, H., Juvonen, M., Derba-Maceluch, M., Koutaniemi, S., et al. (2014). O-Acetylation of glucuronoxylan in *Arabidopsis thaliana* wild type and its change in xylan biosynthesis mutants. *Glycobiology* 24, 494–506. doi: 10.1093/glycob/cwu017

Cosgrove, D. J. (2005). Growth of the plant cell wall. *Nat. Rev. Mol. Cell Biol.* 6, 850–861. doi: 10.1038/nrm1746

Damm, T., Commandeur, U., Fischer, R., and Usadel, B. (2016). Improving the utilization of lignocellulosic biomass by polysaccharide modification. *Process Biochem.* 51, 288–296. doi: 10.1016/j.procbio.2015.12.003

De Martini, J. D., Pattathil, S., Miller, J. S., Li, H., Hahn, M. G., and Wyman, C. E. (2013). Investigating plant cell wall components that affect biomass recalcitrance in poplar and switchgrass. *Energy Environ. Sci.* 6, 898–909. doi: 10.1039/c3ee23801f

Derba-Maceluch, M., Awano, T., Takahashi, J., Lucenius, J., Ratke, C., Kontro, I., et al. (2015). Suppression of xylan endotransglycosylase PtxtXyn10A affects cellulose microfibril angle in secondary wall in aspen wood. *New Phytol.* 205, 666–681. doi: 10.1111/nph.13099

Donaldson, L. A., and Knox, J. P. (2012). Localization of cell wall polysaccharides in normal and compression wood of radiata pine: relationships with lignification and microfibril orientation. *Plant Physiol.* 158, 642–653. doi: 10.1104/pp.111.184036

Edashige, Y., and Ishii, T. (1996). Pectic polysaccharides from xylem-differentiating zone of *Cryptomeria japonica*. *Phytochemistry* 42, 611–616. doi: 10.1016/0031-9422(95)00958-2

Edashige, Y., and Ishii, T. (1997). Rhamnogalacturonan I from xylem differentiating zones of *Cryptomeria japonica*. *Carbohydr. Res.* 304, 357–365. doi: 10.1016/S0008-6215(97)10042-8

Edashige, Y., and Ishii, T. (1998). Rhamnogalacturonan II from cell walls of *Cryptomeria japonica*. *Phytochemistry* 49, 681–690. doi: 10.1016/S0031-9422(98)00238-6

Edwards, M. E., Choo, T., Dickson, C., Scott, C., Gidley, M. J., and Reid, J. S. G. (2004). The seeds of *Lotus japonicus* lines transformed with sense, antisense, and sense/antisense galactomannan galactosyltransferase constructs have structurally altered galactomannans in their endosperm cell walls. *Plant Physiol.* 134, 1153–1162. doi: 10.1104/pp.103.029967

Ek, M., Gellerstedt, G., and Henriksson, G. (2009). *Pulping Chemistry and Technology*. Berlin: De Gruyter. doi: 10.1515/9783110213423

Escamez, S., Gandla, L. M., Derba-Maceluch, M., Lundqvist, S. O., Mellerowicz, E. J., Jönsson, L. J., et al. (2017). A collection of genetically engineered *Populus* trees reveals wood biomass traits that predict glucose yield from enzymatic hydrolysis. *Sci. Rep.* 7:15798. doi: 10.1038/s41598-017-16013-0

Faria-Blanc, N., Mortimer, J. C., and Dupree, P. (2018). A transcriptomic analysis of xylan mutants does not support the existence of a secondary cell wall integrity system in *Arabidopsis*. *Front. Plant Sci.* 9:384. doi: 10.3389/fpls.2018.00384

Felten, J., Vahala, J., Love, J., Gorzsás, A., Rüggeberg, M., Delhomme, N., et al. (2018). Ethylene signaling induces gelatinous layers with typical features of tension wood in hybrid aspen. *New Phytol.* 218, 999–1014. doi: 10.1111/nph.15078

Gandla, L. M., Derba-Maceluch, M., Liu, X., Gerber, L., Master, E. R., Mellerowicz, E. J., et al. (2015). Expression of a fungal glucuronoyl esterase in *Populus*: effects on wood properties and saccharification efficiency. *Phytochemistry* 112, 210–220. doi: 10.1016/j.phytochem.2014.06.002

Giummarella, N., and Lawoko, M. (2016). Structural basis for the formation and regulation of lignin–xylan bonds in birch. *ACS Sustain. Chem. Eng.* 4, 5319–5326. doi: 10.1021/acssuschemeng.6b00911

Giummarella, N., Zhang, L., Henriksson, G., and Lawoko, M. (2016). Structural features of mildly fractionated lignin carbohydrate complexes (LCC) from spruce. *RSC Adv.* 6, 42120–42131. doi: 10.1039/c6ra02399a

Gorshkova, T., Mokshina, N., Chernova, T., Ibragimova, N., Salnikov, V., Mikshina, P., et al. (2015). Aspen tension wood fibers contain β-(1→4)-galactans and acidic arabinogalactans retained by cellulose microfibrils in gelatinous walls. *Plant Physiol.* 169, 2048–2063. doi: 10.1104/pp.15.00690

Guedes, F. T. P., Laurans, F., Quemener, B., Assor, C., Lainé-Prade, V., Boizot, N., et al. (2017). Non-cellulosic polysaccharide distribution during G-layer formation in poplar tension wood fibers: abundance of rhamnogalacturonan I and arabinogalactan proteins but no evidence of xyloglucan. *Planta* 246, 857–878. doi: 10.1007/s00425-017-2737-1

Hartati, S., Sudarmonowati, E., Kaku, T., Ikegaya, H., Kaida, R., Baba, K., et al. (2011). Overexpression of xyloglucanase (AaXEG2) accelerates heteroblastic development in mangium leaves. *J. Wood Sci.* 57, 463–469. doi: 10.1007/s10086-011-1211-0

Hijazi, M., Velasquez, S. M., Jamet, E., Estevez, J. M., and Albenne, C. (2014). An update on post-translational modifications of hydroxyproline-rich glycoproteins: toward a model highlighting their contribution to plant cell wall architecture. *Front. Plant Sci.* 5:395. doi: 10.3389/fpls.2014.00395

Hoffman, G. C., and Timell, T. E. (1970). Isolation of a β-1,3-glucan (laricinan) from compression wood of *Larix laricina*. *Wood Sci. Technol.* 4, 159–162. doi: 10.1007/BF00365301

Jönsson, L. J., and Martín, C. (2016). Pretreatment of lignocellulose: formation of inhibitory by-products and strategies for minimizing their effects. *Bioresour. Technol.* 199, 103–112. doi: 10.1016/j.biortech.2015.10.009

Kaida, R., Kaku, T., Baba, K., Oyadomari, M., Watanabe, T., Nishida, K., et al. (2009). Loosening xyloglucan accelerates the enzymatic degradation of cellulose in wood. *Mol. Plant* 2, 904–909. doi: 10.1093/mp/ssp060

Kaku, T., Kaida, R., Baba, K., Hartati, S., Sudarmonowati, E., and Hayashi, T. (2011). Improvement of fermentable sugar yields of mangium through transgenic overexpression of xyloglucanase. *J. Wood Sci.* 57, 545–548. doi: 10.1007/s10086-011-1180-3

Keppler, B. D., and Showalter, A. M. (2010). IRX14 and IRX14-LIKE, two glycosyl transferases involved in glucuronoxylan biosynthesis and drought tolerance in *Arabidopsis*. *Mol. Plant* 3, 834–841. doi: 10.1093/mp/ssq028

Kim, J. S., and Daniel, G. (2012). Distribution of glucomannans and xylans in poplar xylem and their changes under tension stress. *Planta* 236, 35–50. doi: 10.1007/s00425-012-1588-z

Kim, J. S., and Daniel, G. (2013). Developmental localization of homogalacturonan and xyloglucan epitopes in pit membranes varies between pit types in two poplar species. *IAWA J.* 34, 245–262. doi: 10.1163/22941932-00000021

Lawoko, M., and Henriksson, G. (2006). Characterisation of lignin-carbohydrate complexes (LCCs) of spruce wood (*Picea abies* L.) isolated with two methods. *Holzforschung* 60, 156–161. doi: 10.1515/HF.2006.025

Lee, C., Teng, Q., Huang, W., Zhong, R., and Ye, Z.-H. (2009). Downregulation of PoGT47C expression in poplar results in a reduced glucuronoxylan content and an increased wood digestibility by cellulase. *Plant Cell Physiol.* 50, 1075–1089. doi: 10.1093/pcp/pcp060

Lee, C., Teng, Q., Huang, W., Zhong, R., and Ye, Z. H. (2011). Molecular dissection of xylan biosynthesis during wood formation in poplar. *Mol. Plant* 4, 730–747. doi: 10.1093/mp/ssr035

Lee, C., Teng, Q., Zhong, R., and Ye, Z. H. (2014). Alterations of the degree of xylan acetylation in *Arabidopsis* xylan mutants. *Plant Signal. Behav.* 9:e27797. doi: 10.4161/psb.27797

Li, Q., Min, D., Wang, J. P., Peszlen, I., Horvath, L., Horvath, B., et al. (2011). Down-regulation of glycosyltransferase 8D genes in *Populus* trichocarpa caused reduced mechanical strength and xylan content in wood. *Tree Physiol.* 31, 226–236. doi: 10.1093/treephys/tpr008

Marriott, P. E., Gómez, L. D., and McQueen-Mason, S. J. (2016). Unlocking the potential of lignocellulosic biomass through plant science. *New Phytol.* 209, 1366–1381. doi: 10.1111/nph.13684

Martínez-Abad, A., Berglund, J., Toriz, G., Gatenholm, P., Henriksson, G., Lindström, M., et al. (2017). Regular motifs in xylan modulate molecular flexibility and interactions with cellulose surfaces. *Plant Physiol.* 175, 1579–1592. doi: 10.1104/pp.17.01184

Mellerowicz, E. J., and Gorshkova, T. A. (2012). Tensional stress generation in gelatinous fibres: a review and possible mechanism based on cell-wall structure and composition. *J. Exp. Bot.* 63, 551–565. doi: 10.1093/jxb/err339

Min, D. Y., Li, Q., Chiang, V., Jameel, H., Chang, H. M., and Lucia, L. (2014a). The influence of lignin–carbohydrate complexes on the cellulase-mediated saccharification I: Transgenic black cottonwood (western balsam poplar, California poplar) *P. trichocarpa* including the xylan down-regulated and the lignin down-regulated lines. *Fuel* 119, 207–213. doi: 10.1016/j.fuel.2013.11.047

Min, D. Y., Yang, C., Chiang, V., Jameel, H., and Chang, H. M. (2014b). The influence of lignin-carbohydrate complexes on the cellulase-mediated saccharification II: Transgenic hybrid poplars (*Populus nigra* L. and *Populus maximowiczii* A.). *Fuel* 116, 56–62. doi: 10.1016/j.fuel.2013.07.046

Mohnen, D. (2008). Pectin structure and biosynthesis. *Curr. Opin. Plant Biol.* 11, 266–277. doi: 10.1016/j.pbi.2008.03.006

Mortimer, J. C., Miles, G. P., Brown, D. M., Zhang, Z., Segura, M. P., Weimar, T., et al. (2010). Absence of branches from xylan in *Arabidopsis* gux mutants reveals potential for simplification of lignocellulosic biomass. *Proc. Natl. Acad. Sci. U.S.A.* 107, 17409–17414. doi: 10.1073/pnas.1005456107

Nishikubo, N., Takahashi, J., Roos, A. A., Derba-Maceluch, M., Piens, K., Brumer, H., et al. (2011). Xyloglucan endo-transglycosylase-mediated xyloglucan rearrangements in developing wood of hybrid aspen. *Plant Physiol.* 155, 399–413. doi: 10.1104/pp.110.166934

Oinonen, P., Krawczyk, H., Ek, M., Henriksson, G., and Moriana, R. (2016). Bioinspired composites from cross-linked galactoglucomannan and microfibrillated cellulose: thermal, mechanical and oxygen barrier properties. *Carboh. Pol.* 136, 146–153. doi: 10.1016/j.carbpol.2015.09.038

Park, Y. B., and Cosgrove, D. J. (2015). Xyloglucan and its interactions with other components of the growing cell wall. *Plant Cell Physiol.* 56, 180–194. doi: 10.1093/pcp/pcu204

Park, Y. W., Baba, K., Furuta, Y., Iida, I., Sameshima, K., Arai, M., et al. (2004). Enhancement of growth and cellulose accumulation by overexpression of xyloglucanase in poplar. *FEBS Lett.* 564, 183–187. doi: 10.1016/S0014-5793(04)00346-1

Pauly, M., and Keegstra, K. (2008). Cell-wall carbohydrates and their modification as a resource for biofuels. *Plant J.* 54, 559–568. doi: 10.1111/j.1365-313X.2008.03463.x

Pawar, P. M. A., Derba-Maceluch, M., Chong, S. L., Gandla, M. L., Bashar, S. S., Sparrman, T., et al. (2017a). *In muro* deacetylation of xylan affects lignin properties and improves saccharification of aspen wood. *Biotechnol. Biofuels* 10:98. doi: 10.1186/s13068-017-0782-4

Pawar, P. M. A., Derba-Maceluch, M., Chong, S. L., Gómez, L. D., Miedes, E., Banasiak, A., et al. (2016). Expression of fungal acetyl xylan esterase in *Arabidopsis thaliana* improves saccharification of stem lignocellulose. *Plant Biotechnol. J.* 14, 387–397. doi: 10.1111/pbi.12393

Pawar, P. M. A., Koutaniemi, S., Tenkanen, M., and Mellerowicz, E. J. (2013). Acetylation of woody lignocellulose: significance and regulation. *Front. Plant. Sci.* 4:118. doi: 10.3389/fpls.2013.00118

Pawar, P. M. A., Ratke, C., Balasubramanian, V. K., Chong, S. L., Gandla, M. L., Adriasola, M., et al. (2017b). Downregulation of RWA genes in hybrid aspen affects xylan acetylation and wood saccharification. *New Phytol.* 214, 1491–1505. doi: 10.1111/nph.14489

Ponder, G. R., and Richards, G. N. (1997). Arabinogalactan from Western larch, Part III: alkaline degradation revisited, with novel conclusions on molecular structure. *Carbohydr. Polym.* 34, 251–261. doi: 10.1016/S0144-8617(97)00099-4

Ratke, C., Terebieniec, B. K., Winestrand, S., Derba-Maceluch, M., Grahn, T., Schiffthaler, B., et al. (2015). *Populus* GT43 family members group into distinct sets required for primary and secondary wall xylan biosynthesis and

include useful promoters for wood modification. *Plant Biotechnol. J.* 13, 26–37. doi: 10.1111/pbi.12232

Ratke, C., Terebieniec, B. K., Winestrand, S., Derba-maceluch, M., Grahn, T., Schiffthaler, B., et al. (2018). Downregulating aspen xylan biosynthetic GT43 genes in developing wood stimulates growth via reprograming of the transcriptome. *New Phytol.* 219, 230–245. doi: 10.1111/nph.15160

Ruel, K., Chevalier-Billosta, V., Guillemin, F., Sierra, J. B., and Joseleau, J.-P. (2006). The wood cell wall at the ultrastructural scale-formation and topochemical organization. *Maderas Ciencia Tecnol.* 8, 107–116. doi: 10.4067/S0718-221X2006000200004

Salmén, L. (2015). Wood morphology and properties from molecular perspectives. *Ann. For. Sci.* 72, 679–684. doi: 10.1007/s13595-014-0403-3

Sandquist, D., Filonova, L., von Schantz, L., Ohlin, M., and Daniel, G. (2010). Microdistribution of xyloglucan in differentiating poplar cells. *BioResources* 5, 796–807.

Sawada, D., Kalluri, U. C., O'Neill, H., Urban, V., Langan, P., Davison, B., et al. (2018). Tension wood structure and morphology conducive for better enzymatic digestion. *Biotechnol. Biofuels* 11:44. doi: 10.1186/s13068-018-1043-x

Scheller, H. V., and Ulvskov, P. (2010). Hemicelluloses. *Annu. Rev. Plant Biol.* 61, 263–289. doi: 10.1146/annurev-arplant-042809-112315

Schröder, R., Atkinson, R. G., and Robert, J. (2009). Re-interpreting the role of endo-β-mannanases as mannan endotransglycosylase/hydrolases in the plant cell wall. *Ann. Bot.* 104, 197–204. doi: 10.1093/aob/mcp120

Siedlecka, A., Wiklund, S., Péronne, M.-A., Micheli, F., Leśniewska, J., Sethson, I., et al. (2008). Pectin methyl esterase inhibits intrusive and symplastic cell growth in developing wood cells of *Populus*. *Plant Physiol.* 146, 554–565. doi: 10.1104/pp.107.111963

Sjöström, E. (1993). *Wood Chemistry - Fundamentals and Applications*. San Diego, CA: Academic Press.

Smith, P. J., Wang, H. T., York, W. S., Peña, M. J., and Urbanowicz, B. R. (2017). Designer biomass for next-generation biorefineries: leveraging recent insights into xylan structure and biosynthesis. *Biotechnol. Biofuels* 10:286. doi: 10.1186/s13068-017-0973-z

Song, D., Gui, J., Liu, C., Sun, J., and Li, L. (2016). Suppression of PtrDUF579-3 expression causes structural changes of the glucuronoxylan in *Populus*. *Front. Plant Sci.* 7:493. doi: 10.3389/fpls.2016.00493

Sun, S., Sun, S., Cao, X., and Su, R. (2016). The role of pretreatment in improving the enzymatic hydrolysis of lignocellulosic materials. *Bioresour. Technol.* 199, 49–58. doi: 10.1016/j.biortech.2015.08.061

Suzuki, S., Li, L., Sun, Y. H., and Chiang, V. L. (2006). The cellulose synthase gene superfamily and biochemical functions of xylem-specific cellulose synthase-like genes in *Populus* trichocarpa. *Plant Physiol.* 142, 1233–1245. doi: 10.1104/pp.106.086678

Tan, L., Eberhard, S., Pattathil, S., Warder, C., Glushka, J., Yuan, C., et al. (2013). An *Arabidopsis* cell wall proteoglycan consists of pectin and arabinoxylan covalently linked to an arabinogalactan protein. *Plant Cell* 25, 270–287. doi: 10.1105/tpc.112.107334

Taniguchi, T., Konagaya, K., Kurita, M., Takata, N., Ishii, K., Kondo, T., et al. (2012). Growth and root sucker ability of field-grown transgenic poplars overexpressing xyloglucanase. *J. Wood Sci.* 58, 550–556. doi: 10.1007/s10086-012-1281-7

Tavares, E. Q., De Souza, A. P., and Buckeridge, M. S. (2015). How endogenous plant cell-wall degradation mechanisms can help achieve higher efficiency in saccharification of biomass. *J. Exp. Bot.* 66, 4133–4143. doi: 10.1093/jxb/erv171

Teleman, A. (2009). "Hemicelluloses and pectins," in *Wood Chemistry and Wood Biotechnology*, Vol. 1, eds G. Gellerstedt, M. Ek, and G. Henriksson (Berlin: De Gruyter), 101–120.

Teleman, A., Lundqvist, J., Tjerneld, F., Stålbrand, H., and Dahlman, O. (2000). Characterization of acetylated 4-O-methylglucuronoxylan isolated from aspen employing 1H and 13C NMR spectroscopy. *Carbohydr. Res.* 329, 807–815. doi: 10.1016/S0008-6215(00)00249-4

Teleman, A., Tenkanen, M., Jacobs, A., and Dahlman, O. (2002). Characterization of O -acetyl- (4- O -methylglucurono) xylan isolated from birch and beech. *Carbohydr. Res.* 337, 373–377. doi: 10.1016/S0008-6215(01)00327-5

Wang, Y., Fan, C., Hu, H., Li, Y., Sun, D., Wang, Y., et al. (2016). Genetic modification of plant cell walls to enhance biomass yield and biofuel production in bioenergy crops. *Biotechnol. Adv.* 34, 997–1017. doi: 10.1016/j.biotechadv.2016.06.001

Wang, Y., Mortimer, J. C., Davis, J., Dupree, P., and Keegstra, K. (2013). Identification of an additional protein involved in mannan Biosynthesis. *Plant J.* 73, 105–117. doi: 10.1111/tpj.12019

Wang, Z., Wu, G., and Jönsson, L. J. (2018). Effects of impregnation of softwood with sulfuric acid and sulfur dioxide on chemical and physical characteristics, enzymatic digestibility, and fermentability. *Bioresour. Technol.* 247, 200–208. doi: 10.1016/j.biortech.2017.09.081

Yamamoto, M., Saito, T., Isogai, A., Kurita, M., Kondo, T., Taniguchi, T., et al. (2011). Enlargement of individual cellulose microfibrils in transgenic poplars overexpressing xyloglucanase. *J. Wood Sci.* 57, 71–75. doi: 10.1007/s10086-010-1140-3

Yang, B., and Wyman, C. E. (2008). Pretreatment: the key to unlocking low-cost cellulosic ethanol. *Biofuel Bioprod. Biorefin.* 2, 26–40. doi: 10.1002/bbb.49

Yang, Y., Yoo, C. G., Winkeler, K. A., Collins, C. M., Hinchee, M. A. W., Jawdy, S. S., et al. (2017). Overexpression of a domain of unknown function 231-containing protein increases O-xylan acetylation and cellulose biosynthesis in *Populus*. *Biotechnol. Biofuels* 10:311. doi: 10.1186/s13068-017-0998-3

Zhang, M., Chavan, R. R., Smith, B. G., McArdle, B. H., and Harris, P. J. (2016). Tracheid cell-wall structures and locations glucans in compression woods of radiata pine (*Pinus radiata* D. Don). *BMC Plant Biol.* 16:194. doi: 10.1186/s12870-016-0884-3

Zhao, W., Odelius, K., Edlund, U., Zhao, C., and Albertsson, A. C. (2015). In situ synthesis of field-responsive hemicellulose hydrogels for drug delivery. *Biomacromolecules* 16, 2522–2528. doi: 10.1021/acs.biomac.5b00801

Zhao, Y., Zhang, Q., Yuan, L., Zhang, R., and Li, L. (2013). N-glycosylation and dimerization regulate the PtrMAN6 enzyme activity that may modulate generation of oligosaccharide signals. *Plant Signal. Behav.* 8:e26956. doi: 10.4161/psb.26956

Check for updates

Breeding and Engineering Trees to Accumulate High Levels of Terpene Metabolites for Plant Defense and Renewable Chemicals

Gary F. Peter*

School of Forest Resources and Conservation, Genetics Institute, University of Florida, Gainesville, FL, United States

Plants evolved the capacity to synthesize highly diverse sets of secondary metabolites which are important for plant adaptation and health. In forest trees, many classes of compounds, particularly ones related to defense against insects, fungi, and bacteria accumulate to levels that enable their recovery and commercial use. One of the oldest examples is conifer terpenes, but terpenes are important secondary products from other tree species including eucalypts. Because terpenes, latex, and natural gums are synthesized and stored in specialized secretory glands, ducts, and laticifers in mostly pure forms they can be collected from live trees in addition to being extracted during industrial processing of wood. This minireview describes the potential of breeding and genetic engineering approaches to increase the quantities of terpene secondary metabolites to increase the amount of secondary products and thereby increasing the value of planted and managed forest trees. I advance the view that breeding and genetic engineering of metabolic pathways and specialized cell secretory structures can dramatically increase tissue terpene content.

Keywords: resin duct, terpene, conifer, oil gland, *Eucalyptus*, defense, biosynthesis, cell specialization

INTRODUCTION

Plants evolved the ability to synthesize a large diversity of carbon rich compounds, including fatty acids, lipids, phenolics, sterols, tannins, and terpenes. Some of these hydrocarbons can accumulate to high percentages of dry mass. For example, oil seeds accumulate 10–75% triacylglycerol by dry mass, depending on the species. With the goal of producing hydrocarbons directly in plants for renewable chemicals and biofuels, tobacco was engineered using a push-pull-protect strategy to accumulate 15% of leaf dry mass as triacylglycerol (Vanhercke et al., 2014). In addition to hydrocarbons synthesized via primary metabolic pathways, there are many synthesized via secondary metabolic pathways. One of the most diverse sets of plant hydrocarbons are terpenes. Plant terpenes are synthesized via the plastid localized 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway and the cytosolic mevalonate pathway (MVA) (Kirby and Keasling, 2009; Vranova et al., 2013). Terpenes are critical components of plant defense against herbivores, insects, pathogenic fungi, and bacteria (Franceschi et al., 2005). For terpene defenses to be effective, sufficient amounts need to accumulate prior to invasion; however, at these levels they are phytotoxic. To minimize this toxicity, plants evolved the ability to sequester them in extracellular spaces formed during

specialized differentiation of secretory tissues (Fahn, 1979, 1988). In a wide variety of plant species terpenes accumulate in extracellular cavities, or lumen that form during differentiation of glandular trichomes (Glas et al., 2012; Huchelmann et al., 2017), secretory cavities, or oil glands (OG) (Mewalal et al., 2017), and resin ducts (Zulak and Bohlmann, 2010). These tissues are specialized for the biosynthesis and storage of terpenes.

In many forest tree species, terpenes are vital chemical, and physical defense compounds, and specialized tissues are crucial to their biosynthesis and storage at high concentrations. For example, in *Eucalyptus* species mono- and sesquiterpenes are important for defense against insects, and herbivores, and are synthesized, and accumulate in secretory cavities, or OG within leaves (King et al., 2006; Naidoo et al., 2014). Eucalypt leaf terpene content ranges from 0.5 to 8% of dry mass and they are commercially extracted and used in the fragrance and pharmaceutical industries and have potential as biofuels (Mewalal et al., 2017). Conifer terpenes are a primary component of a durable, quantitative defense mechanism against stem boring insects and fungal pathogens (Trapp and Croteau, 2001; Franceschi et al., 2005). Constitutive and induced terpenes are critical for resistance (Hodges et al., 1985; Strom et al., 2002; Klepzig et al., 2005). Upon wounding, preformed constitutively produced terpenes are released at the wound site and the insects, and fungi are contained and growth is inhibited. The toxic terpenes cannot only kill the attacking organisms but they also flush and seal the wound site forming a physical barrier protecting the stem from further infection (Trapp and Croteau, 2001). Conifer terpenes are synthesized and accumulate in primary and secondary resin ducts, and form a viscous oleoresin secretion consisting of a complex mixture of volatile monoterpene olefins, and non-volatile diterpenoid resin acids that naturally accumulate to between 2 and 40% of the dry mass of wood (Trapp and Croteau, 2001).

The native levels of terpenes in eucalypt leaves and wood of conifer species is adequate for commercial extraction and use as renewable chemicals, and biofuels. In conifers, commercial scale collection of terpenes occurs today from live pine trees by tapping, solvent extraction of stumps, and during chemical pulping of wood, with global production from all sources about 3.5–4.0 millon tons y^{-1}. Southeastern United States pulp mills recover ~900,000 tons y^{-1} of pine terpenes and fatty acids as crude tall oil (CTO) supporting specialty chemical biorefineries that compete in markets with petroleum derived feedstocks justifying the concept that biofuels from pine hydrocarbons could be profitable without subsidy if supply was increased (American, 2011). In the European Union, CTO is converted to renewable diesel fuel at two industrial scale biorefineries: one with a capacity of 120 million liters y^{-1} (UPM-Kymmene, 2018). Pine monoterpenes can be efficiently dimerized to make jet fuel (Harvey et al., 2010; Mewalal et al., 2017). Globally, pine chemicals are the largest non-food based renewable hydrocarbon supply. However, the pine terpene supply is limited by the relatively low 2–4% average wood terpene content of dry wood.

Here, I discuss three synergistic approaches to increase tissue terpene content that can lead to forest trees with improved resistance to herbivores, insects, and pathogens as well as greater supplies of terpenes for renewable chemicals, and biofuels. Three main approaches to increase tissue terpene content considered are: (1) increase the number of specialized secretory glands and duct in plant organs to create greater biosynthetic, and storage capacity; (2) increase luminal storage volume of specialized secretory tissues; and (3) enhance carbon flux into terpene biosynthetic pathways.

INCREASE SECRETORY TISSUES IN PLANT ORGANS

In conifers, primary resin ducts (RD) form in needles and non-woody shoots, and secondary RD form in the secondary xylem, or wood (Larson, 1994). Although there is limited knowledge of primary RD formation, secondary RD formation, and function are better studied (Larson, 1994; Zulak and Bohlmann, 2010). Within a network of secondary RDs, terpenes are synthesized by heterotrophic epithelial cells that line the duct and are secreted into the lumen where they accumulate (Fahn, 1979; Trapp and Croteau, 2001; Zulak and Bohlmann, 2010; **Figure 1**). The lumens of axial and radial RDs are interconnected such that upon wounding the terpene oleoresin can flow from within the wood toward the surface of the stem (Larson, 1994; **Figure 1A**). The interconnectedness of the network is also important for the RD to obtain sucrose and nutrients from the phloem at the periphery of the stem through the live ray cell files. During stem diameter growth, axial, and radial RDs differentiate in the cambial zone from derivatives of fusiform and ray initials (Larson, 1994). Two adjacent derivatives begin dividing anticlinally to form a multicellular pre-duct structure which subsequently form the lumen schizogenously by separation of the cells at the middle lamella (Larson, 1994; **Figure 2**).

The genetic mechanisms that control constitutive RD cell fate and differentiation are not known. In loblolly pine, we showed that constitutive axial resin duct number (Westbrook et al., 2015) is under moderate genetic control, inherited as a polygenic trait, positively genetically correlated with xylem growth, and has sufficient variation for breeding to increase constitutive axial RD number about 10% in a single generation. *Picea* species (spruce) have emerged as the best-studied experimental system for dissecting the developmental signals and biochemical basis of induced terpene resin formation (Franceschi et al., 2005; Keeling and Bohlmann, 2006; Zulak and Bohlmann, 2010). Insect feeding, wounding, ethylene, and methyl jasmonate (MeJ) treatments induce development of traumatic resin ducts (TRD), which develop as a single contiguous row of elements derived from xylem mother cells in the cambial zone (Franceschi et al., 2005). Anatomical time course studies show that axial TRD take about 9–36 days to differentiate depending on the species (Larson, 1994; Nagy et al., 2000). In Norway spruce, a detailed time course of TRD formation and terpene synthesis in wood after MeJ application shows that geranylgeranyl-pyrophosphate synthase, monoterpene, and diterpene synthase activity began increasing after 5 days and peaked ~15 days and increases in wood monoterpene and diterpenoid content
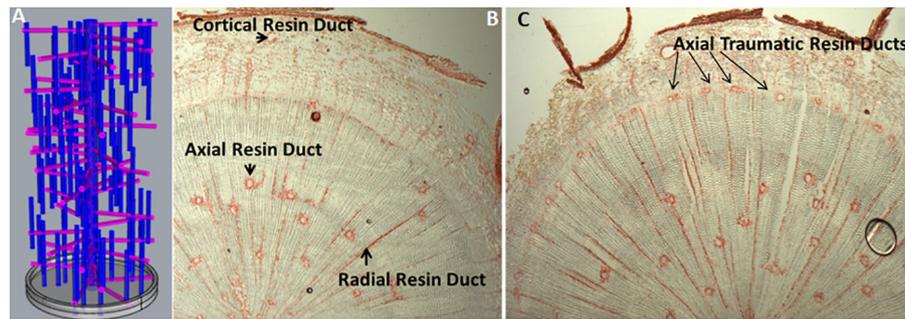
FIGURE 1 | Resin duct organization and induction. **(A)** A theoretical three dimensional model depicting the interconnectivity of constitutive axial and radial resin ducts in woody stems of Pinaceae; **(B)** Transverse cryosection of an untreated 1-year-old loblolly pine stem, stained with safranin red, showing constitutive axial and radial resin ducts in the wood and cortical resin duct in the bark (Peter, 2018, Unpublished); **(C)** Transverse cryosection of 1-year-old loblolly pine stem 10 days after treatment with 10 mM methyl jasmonate (Martin et al., 2002), stained with safranin red showing newly formed axial traumatic resin ducts in the cambial zone/young xylem (Peter, 2018, Unpublished).
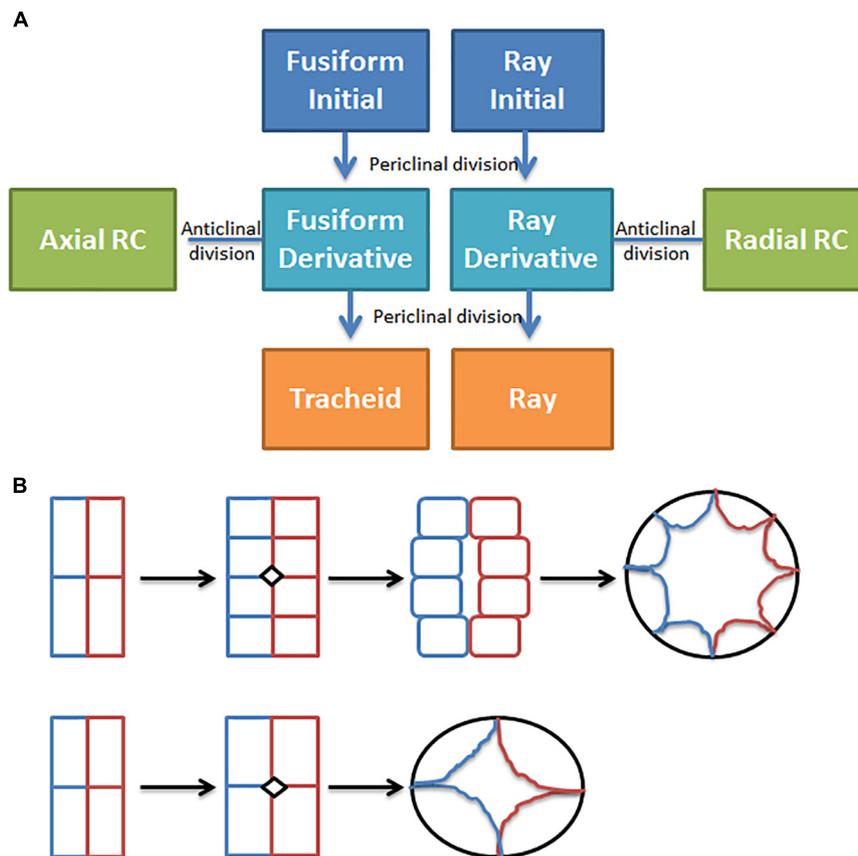
FIGURE 2 | Resin duct formation in the cambial zone. **(A)** Simple conceptual model for secondary resin duct formation; **(B)** Transverse plane drawing of two ways cambial zone cells may divide to form epithelial cells in more narrow or wider axial resin ducts, adapted from Sato and Ishida (1983), the colors indicate cell lineage.

were observed after 10 days and peaked between 15 and 25 days (Martin et al., 2002). New TRD are observed at 10 days and supports the conclusion that in spruce wood newly synthesized terpene is only present in the newly formed TRD and not in undifferentiated ray cells. Thus, increasing RD numbers should be an effective strategy to increase wood terpene content. Exogenous application of MeJ induces axial TRD in the cambial region (Martin et al., 2002; **Figure 1C**), and application of inhibitors of ethylene biosynthesis suppresses jasmonate TRD induction, suggesting that jasmonate induces ethylene

synthesis which then induces TRD formation (Hudgins and Franceschi, 2004). Consistent with this model, MeJ application to sitka spruce bark rapidly upregulates genes that code for the rate limiting enzyme for ethylene biosynthesis, 1-aminocyclopropane carboxylic acid synthase (Ralph et al., 2007). These observations support that genetically increasing ethylene, or jasmonate biosynthesis, or signaling in the cambial zone should increase RD number and lead to greater wood terpene content.

To identify RD specific and jasmonate regulated genes in RDs, the transcriptomes of isolated cortical RDs from bark of Norway spruce were compared before and after treatment with MeJ (Celedon et al., 2017). In untreated bark, 863 cortical RD specific transcripts (Tau score > 0.75) were identified, and in MeJ treated bark, 766 cortical RD specific transcripts were identified. Of these 766 transcripts, about 648 were also identified as being specific to untreated cortical RD, indicating that 85% of the cortical RD specific genes remain RD specific 8 days after MeJ treatment (Celedon et al., 2017). As expected in the untreated cortical RD, specific transcripts included genes that code for all MEP biosynthetic enzymes, and a large number of cytochrome P450s, and terpene synthases. In addition, they included 34 transcripts that are annotated as transcription factors or encode putative DNA binding domains. There is potential that these regulators mediate RD specification and terpene pathway gene expression. In bark, ∼3,700 genes were differentially expressed 8 days after MeJ treatment; however, the number of jasmonate responsive genes in cortical RD was not reported. The terpene pathway genes in cortical RD showed that they were reduced 8 days after MeJ treatment relative to controls (Celedon et al., 2017). Overall, this data shows that in control and jasmonate treated bark, the genes coding for terpene biosynthesis are specifically expressed in cortical RDs.

Manipulating ethylene or jasmonate may induce more secondary RDs but could also have unintended effects on tracheids, so targeted genetic manipulation may be preferred. Targeted manipulation is limited by our lack of knowledge of the molecular mechanisms that regulate resin duct cell fate, differentiation, and function. **Figure 2** shows a simple conceptual model for secondary axial and radial RD formation. During normal stem diameter growth, constitutively formed secondary axial RDs appear to occur randomly throughout the wood volume but are typically adjacent to one or more ray cell files (**Figure 1**). While, derivatives of fusiform initials typically differentiate into tracheids, those adjacent to ray initials can also differentiate into axial resin ducts. Because the vast majority of fusiform derivatives differentiate into tracheids, for axial RDs to form, a first step is to inhibit regulators that specify tracheid cell fate. Although the molecular control of tracheid cell specification is presently unknown in conifers, vascular related NAC domain transcription factors function as master switches in angiosperms xylem (Zhang et al., 2014; Jokipii-Lukkari et al., 2017). In analogy to xylem cell specification, it is reasonable to hypothesize that a master transcriptional regulator specifies RD identity. Based on parsimony, this regulator also will be responsive to ethylene and/or jasmonate, and therefore these hormones could be important for constitutive

RD formation in addition to TRD formation. However, the overlap between the traumatic RD and constitutive RD pathways is not known. It is also unknown whether the fusiform derivatives need to first have ray cell identity to differentiate into RDs or whether derivatives can directly acquire RD cell fate. Constitutive radial RDs also form in the cambial zone; however, it is unclear if they can differentiate within existing ray cell files deeper in the stem. Once a radial RD forms, typically next to an axial RD, the radial RD fate extends within that ray cell file out to the cambial meristem (**Figure 1**). To discover genes involved in RD formation, we are using association genetics in loblolly pine. With single nucleotide polymorphisms (SNP) in 4027 genes we identified 16 well supported candidate genes associated with axial RD number, some of which were also significant for oleoresin flow (Westbrook et al., 2015). More recently, we genotyped ∼70,000 SNPs in this population and identified significant SNPs in an additional 18 genes for constitutive axial RD number (Peter, 2018, Unpublished). We are investigating the importance of some of these candidate genes in RD specification through genetic engineering.

INCREASE LUMINAL VOLUME OF SECRETORY TISSUES

Conifer species differ in the organization of terpene resin producing cells and tissues. Abies species form both resin cells and multicellular resin blisters, while some Pinaceae species form constricted resin passages and other Pinaceae species form non-occluded resin ducts (Penhallow, 1907; Lewinsohn et al., 1991). Although no rigorous comparison amongst these species is available to separate the significance of resin structure number and extracellular volume, in general, wood terpene content is greater in species with resin ducts compared with species with resin cells (Lewinsohn et al., 1991).

Increasing luminal volume in addition to increasing RD number should lead to wood with greater terpene content. What controls RD luminal diameter is unknown, but lumen volume tends to be greater in RDs that form from a greater number of early cell divisions during differentiation (**Figure 2B**). Possibly four instead of two cambial derivatives could be induced to divide and commit to RD large diameter lumen would form. In loblolly pine, the length of axial RD range from 5 to 20 cm and average 10.4 cm, and in slash pine average RD length was more than double (49.8 cm vs. 20 cm) at 20 versus 10 years (Larson, 1994). In southern pines, axial RDs are 10–100 times longer than individual fusiform initials (3–5 mm) suggesting that a large number of contiguous fusiform initial derivatives within a single file need to commit to form axial RDs. The mechanism(s) which mediates axial RD length are unknown, but I hypothesize that cell to cell communication signals that specify RD identity are important for coordinating differentiation of the 10–100s of cambial meristem derivatives. If this hypothesis is correct, then discovery of this transmissible signal could enable engineering of larger RD and thereby increase wood terpene content.

In *Eucalyptus polybractea*, the number of OG is correlated well with leaf terpene content (King et al., 2006; Goodger and Woodrow, 2012). Analysis of individual OG shows they have very similar capacities, suggesting that the storage volume of the cavity is relatively tightly controlled, and it was hypothesized that gland capacity is controlled by leaf expansion (King et al., 2006). If King's hypothesis is correct, then increasing leaf thickness could increase cavity volume. Because leaf terpene content is limited not only by the number of OG but also by luminal volume of the gland it could be possible to increase this volume by engineering OG to be more duct like, similar to primary RD in conifers. I hypothesize that the difference between formation of spherical glands and elongated ducts is related to the existence of a cell to cell signaling mechanism which evolved in Pinaceae that mediates duct formation.

ENHANCE CARBON FLUX TO TERPENES

Plants synthesize monoterpenes and diterpenes via the plastid localized 2-C-MEP pathway, and sesquiterpenes, and triterpenes via the cytosolic MVA (Kirby and Keasling, 2009; Vranova et al., 2013; Banerjee and Sharkey, 2014). The MEP pathway is considered the high flux pathway in part because many species including *Eucalyptus* are rich in monoterpenes and conifers are rich in monoterpenes, and diterpenoids. The first committed step in the MEP pathway is catalyzed by 1-deoxyxylulose 5-phosphate synthase (DXS). Constitutive overexpression of DXS enzymes elevated terpene alkaloids in *C. roseus* (Peebles et al., 2011), taxadiene in *Arabidopsis thaliana* expressing a taxadiene synthase (Botella-Pavia et al., 2004), and terpenes 2–3.5 fold in leaves of *L. latifolia* (Munoz-Bertomeu et al., 2006). DXS is sensitive to feedback inhibition by isoprenoid precursor's isopentenyl diphosphate and dimethylallyl diphosphate which can control carbon flux through the pathway (Banerjee et al., 2013; Banerjee and Sharkey, 2014). A mutant form of poplar DXS was described that is less sensitive to feedback regulation; however, this has yet to be tested in plants (Banerjee et al., 2016). In spruce, the DXS gene family codes for type I and type II enzymes. Type II DXS genes are involved in secondary metabolism and in spruce are upregulated by wounding and fungal infection, suggesting that DXS overexpression should enhance wood terpene synthesis (Phillips et al., 2007). Upregulation of deoxyxylulose 5-phospate reductoisomerase (DXR) leads to a 50% increase in terpene based essential oils stored in globular trichomes of peppermint (Mahmoud and Croteau, 2001), and increased isoprene products, e.g., chlorophyll, β-carotene and taxadiene in *A. thaliana* expressing taxadiene synthase (Carretero-Paulet et al., 2006).

The plant MEP pathway is conserved with the deoxyxylulose 5-phosphate (DXP) pathway in microbes which has been improved through extensive genetic engineering (Ajikumar et al., 2008; Wang et al., 2009; Kirby et al., 2016) and synthetic biology approaches (Keasling, 2014). The results from many of these studies show that larger increases in isoprene products occur when multiple genes are manipulated in combination (Ajikumar et al., 2008; Wang et al., 2009). Changes to multiple genes in plants will require advances in genetic engineering efficiencies and more effective gene editing tools, particularly for conifers which have very large genomes and large gene families (Birol et al., 2013; Nystedt et al., 2013; Neale et al., 2014; Wegrzyn et al., 2014; Zimin et al., 2014). A particularly interesting finding was the identification of two *Escherichia coli* genes that can bypass the DXS catalyzed step (Kirby et al., 2015). By-passing DXS should improve carbon efficiency as DXS releases CO_2 when it condenses pyruvate and glyceraldehyde 3-phosphate into the five carbon 1-DXP product. The wild type yajO and mutant forms of ribB gene product were shown to convert ribulose 5-phosphate into DXP *in vitro* and upon expression in *E. coli* improve terpene synthesis (Kirby et al., 2015). Transgenic loblolly pine constitutively expressing plastid localized YajO and RibB(G108S) showed significantly increased terpenes in the wood of 1-year-old seedlings (Kirby et al., 2018, Unpublished).

CONCLUSION

There is great opportunity to increase terpene hydrocarbon content in eucalypts and conifers through breeding and genetic engineering. Continued discovery of the molecular mechanisms that control OG and RD specification and cell to cell signaling have the potential to enable genetic engineering and targeted breeding to increase the number of resin ducts and OG and their size leading to trees with greater levels of constitutively produced terpenes. Capitalizing on the advances in genetic engineering and synthetic biology with the MEP pathway in microbes should further enable greater carbon flux and carbon efficiency.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Ajikumar, P. K., Tyo, K., Carlsen, S., Mucha, O., Phon, T. H., and Stephanopoulos, G. (2008). Terpenoids: opportunities for biosynthesis of natural product drugs using engineered microorganisms. *Mol. Pharm.* 5, 167–190. doi: 10.1021/mp700151b

American Chemical Council (2011). *Economic Benefits of the American Pine Chemical Industry*. Available at: https://pinechemistry.americanchemistry.com/Pine-Chemistry-Basics/Learn-about-the-Economic-Benefits-of-Pine-Chemistry.pdf

Banerjee, A., Preiser, A. L., and Sharkey, T. D. (2016). Engineering of recombinant poplar deoxy-D-xylulose-5-phosphate synthase (PtDXS) by Site-directed mutagenesis improves its activity. *PLoS One* 11:e0161534. doi: 10.1371/journal.pone.0161534

Banerjee, A., and Sharkey, T. D. (2014). Methylerythritol 4-phosphate (MEP) pathway metabolic regulation. *Nat. Prod. Rep.* 31, 1043–1055. doi: 10.1039/c3np70124g

Banerjee, A., Wu, Y., Banerjee, R., Li, Y., Yan, H. G., and Sharkey, T. D. (2013). Feedback inhibition of deoxy-D-xylulose-5-phosphate synthase regulates the methylerythritol 4-phosphate pathway. *J. Biol. Chem.* 288, 16926–16936. doi: 10.1074/jbc.M113.464636

Birol, I., Raymond, A., Jackman, S. D., Pleasance, S., Coope, R., Taylor, G. A., et al. (2013). Assembling the 20 Gb white spruce (Picea glauca) genome from whole-genome shotgun sequencing data. *Bioinformatics* 29, 1492–1497. doi: 10.1093/bioinformatics/btt178

Botella-Pavia, P., Besumbes, O., Phillips, M. A., Carretero-Paulet, L., Boronat, A., and Rodriguez-Concepcion, M. (2004). Regulation of carotenoid biosynthesis in plants: evidence for a key role of hydroxymethylbutenyl diphosphate reductase in controlling the supply of plastidial isoprenoid precursors. *Plant J.* 40, 188–199. doi: 10.1111/j.1365-313X.2004.02198.x

Carretero-Paulet, L., Cairo, A., Botella-Pavia, P., Besumbes, O., Campos, N., Boronat, A., et al. (2006). Enhanced flux through the methylerythritol 4-phosphate pathway in Arabidopsis plants overexpressing deoxyxylulose 5-phosphate reductoisomerase. *Plant Mol. Biol.* 62, 683–695. doi: 10.1007/s11103-006-9051-9

Celedon, J. M., Yuen, M. M. S., Chiang, A., Henderson, H., Reid, K. E., and Bohlmann, J. (2017). Cell-type- and tissue-specific transcriptomes of the white spruce (*Picea glauca*) bark unmask fine-scale spatial patterns of constitutive and induced conifer defense. *Plant J.* 92, 710–726. doi: 10.1111/tpj.13673

Fahn, A. (1979). *Secretory Tissues in Plants*. New York, NY: Academic Press, 302.

Fahn, A. (1988). Secretory-tissues in vascular plants. *New Phytol.* 108, 229–257. doi: 10.1111/j.1469-8137.1988.tb04159.x

Franceschi, V. R., Krokene, P., Christiansen, E., and Krekling, T. (2005). Anatomical and chemical defenses of conifer bark against bark beetles and other pests. *New Phytol.* 167, 353–375. doi: 10.1111/j.1469-8137.2005.01436.x

Glas, J. J., Schimmel, B. C., Alba, J. M., Escobar-Bravo, R., Schuurink, R. C., and Kant, M. R. (2012). Plant glandular trichomes as targets for breeding or engineering of resistance to herbivores. *Int. J. Mol. Sci.* 13, 17077–17103. doi: 10.3390/ijms131217077

Goodger, J. Q. D., and Woodrow, I. E. (2012). Genetic determinants of oil yield in eucalyptus polybractea R.T. Baker. *Trees* 26, 1951–1956. doi: 10.1007/s00468-012-0744-1

Harvey, B. G., Wright, M. E., and Quintana, R. L. (2010). High-density renewable fuels based on the selective dimerization of pinenes. *Energy Fuels* 24, 267–273. doi: 10.1021/ef900799c

Hodges, J. D., Nebeker, T. E., Deangelis, J. D., Karr, B. L., and Blanche, C. A. (1985). Host resistance and mortality: a hypothesis based on southern pine beetle-microorganism-host interactions. *Bull. Entomol. Soc. Am.* 31, 31–35. doi: 10.1093/besa/31.1.31

Huchelmann, A., Boutry, M., and Hachez, C. (2017). Plant glandular trichomes: natural cell factories of high biotechnological interest. *Plant Physiol.* 175, 6–22. doi: 10.1104/pp.17.00727

Hudgins, J. W., and Franceschi, V. R. (2004). Methyl jasmonate-induced ethylene production is responsible for conifer phloem defense responses and reprogramming of stem cambial zone for traumatic resin duct formation. *Plant Physiol.* 135, 2134–2149. doi: 10.1104/pp.103.037929

Jokipii-Lukkari, S., Sundell, D., Nilsson, O., Hvidsten, T. R., Street, N. R., and Tuominen, H. (2017). NorWood: a gene expression resource for evo-devo studies of conifer wood development. *New Phytol.* 216, 482–494. doi: 10.1111/nph.14458

Keasling, J. D. (2014). Synthetic biology: a global approach. *Nature* 510, 218. doi: 10.1038/510218c

Keeling, C. I., and Bohlmann, J. (2006). Diterpene resin acids in conifers. *Phytochemistry* 67, 2415–2423. doi: 10.1016/j.phytochem.2006.08.019

King, D. J., Gleadow, R. M., and Woodrow, I. E. (2006). Regulation of oil accumulation in single glands of *Eucalyptus polybractea*. *New Phytol.* 172, 440–451. doi: 10.1111/j.1469-8137.2006.01842.x

Kirby, J., Dietzel, K. L., Wichmann, G., Chan, R., Antipov, E., Moss, N., et al. (2016). Engineering a functional 1-deoxy-D-xylulose 5-phosphate (DXP) pathway in *Saccharomyces cerevisiae*. *Metab. Eng.* 38, 494–503. doi: 10.1016/j.ymben.2016.10.017

Kirby, J., and Keasling, J. D. (2009). Biosynthesis of plant isoprenoids: perspectives for microbial engineering. *Annu. Rev. Plant Biol.* 60, 335–355. doi: 10.1146/annurev.arplant.043008.091955

Kirby, J., Nishimoto, M., Chow, R. W. N., Baidoo, E. E., Wang, G., Martin, J., et al. (2015). Enhancing terpene yield from sugars via novel routes to 1-deoxy-D-xylulose 5-phosphate. *Appl. Environ. Microbiol.* 81, 130–138. doi: 10.1128/AEM.02920-14

Klepzig, K. D., Robison, D. J., Fowler, G., Minchin, P. R., Hain, F. P., and Allen, H. L. (2005). Effects of mass inoculation on induced oleoresin response in intensively managed loblolly pine. *Tree Physiol.* 25, 681–688. doi: 10.1093/treephys/25.6.681

UPM-Kymmene Corp (2018). Available at: https://www.upm.com/Businesses/upm-biofuels/Pages/default.aspx

Larson, P. (1994). *The Vascular Cambium –Development and Structure*. Berlin: Springer-Verlag, 725. doi: 10.1007/978-3-642-78466-8

Lewinsohn, E., Gijzen, M., Savage, T. J., and Croteau, R. (1991). Defense-mechanisms of conifers - relationship of monoterpene cyclase activity to anatomical specialization and oleoresin monoterpene content. *Plant Physiol.* 96, 38–43. doi: 10.1104/pp.96.1.38

Mahmoud, S. S., and Croteau, R. B. (2001). Metabolic engineering of essential oil yield and composition in mint by altering expression of deoxyxylulose phosphate reductoisomerase and menthofuran synthase. *Proc. Natl. Acad. Sci. U.S.A.* 98, 8915–8920. doi: 10.1073/pnas.141237298

Martin, D., Tholl, D., Gershenzon, J., and Bohlmann, J. (2002). Methyl jasmonate induces traumatic resin ducts, terpenoid resin biosynthesis, and terpenoid accumulation in developing xylem of Norway spruce stems. *Plant Physiol.* 129, 1003–1018. doi: 10.1104/pp.011001

Mewalal, R., Rai, D. K., Kainer, D., Chen, F., Külheim, C., Peter, G. F., et al. (2017). Plant-derived terpenes: a feedstock for specialty biofuels. *Trends Biotechnol.* 35, 227–240. doi: 10.1016/j.tibtech.2016.08.003

Munoz-Bertomeu, J., Arrillaga, I., Ros, R., and Segura, J. (2006). Up-regulation of 1-deoxy-D-xylulose-5-phosphate synthase enhances production of essential oils in transgenic spike lavender. *Plant Physiol.* 142, 890–900. doi: 10.1104/pp.106.086355

Nagy, N. E., Franceschi, V. R., Solheim, H., Krekling, T., and Christiansen, E. (2000). Wound-induced traumatic resin duct development in stems of Norway spruce (Pinaceae): anatomy and cytochemical traits. *Am. J. Bot.* 87, 302–313. doi: 10.2307/2656626

Naidoo, S., Kulheim, C., Zwart, L., Mangwanda, R., Oates, C. N., Visser, E. A., et al. (2014). Uncovering the defence responses of eucalyptus to pests and pathogens in the genomics age. *Tree Physiol.* 34, 931–943. doi: 10.1093/treephys/tpu075

Neale, D. B., Wegrzyn, J. L., Stevens, K. A., Zimin, A. V., Puiu, D., Crepeau, M. W., et al. (2014). Decoding the massive genome of loblolly pine using haploid DNA and novel assembly strategies. *Genome Biol.* 15:R59. doi: 10.1186/gb-2014-15-3-r59

Nystedt, B., Street, N. R., Wetterbom, A., Zuccolo, A., Lin, Y.-C., Scofield, D. G., et al. (2013). The Norway spruce genome sequence and conifer genome evolution. *Nature* 497, 579–584. doi: 10.1038/nature12211

Peebles, C. A., Sander, G. W., Hughes, E. H., Peacock, R., Shanks, J. V., and San, K. Y. (2011). The expression of 1-deoxy-D-xylulose synthase and geraniol-10-hydroxylase or anthranilate synthase increases terpenoid indole alkaloid accumulation in *Catharanthus roseus* hairy roots. *Metab. Eng.* 13, 234–240. doi: 10.1016/j.ymben.2010.11.005

Penhallow, D. P. (1907). *A Manual of the North American Gymnosperms*. Boston: Athenaeum Press.

Phillips, M. A., Walter, M. H., Ralph, S. G., Dabrowska, P., Luck, K., Urós, E. M., et al. (2007). Functional identification and differential expression of 1-deoxy-D-xylulose 5-phosphate synthase in induced terpenoid resin formation of Norway spruce (Picea abies). *Plant Mol. Biol.* 65, 243–257. doi: 10.1007/s11103-007-9212-5

Ralph, S. G., Hudgins, J. W., Jancsik, S., Franceschi, V. R., and Bohlmann, J. (2007). Aminocyclopropane carboxylic acid synthase is a regulated step in ethylene-dependent induced conifer defense. Full-length cDNA cloning of a multigene family, differential constitutive, and wound- and insect-induced expression, and cellular and subcellular localization in spruce and Douglas fir. *Plant Physiol.* 143, 410–424. doi: 10.1104/pp.106.089425

Sato, K., and Ishida, S. (1983). Formation of vertical resin canals. *Res. Bull. Coll. Exp. For. Hokaido Univ.* 40, 723–740.

Strom, B. L., Goyer, R. A., Ingram, L. L., Boyd, G. D. L., and Lott, L. H. (2002). Oleoresin characteristics of progeny of loblolly pines that escaped attack by the southern pine beetle. *For. Ecol. Manag.* 158, 169–178. doi: 10.1016/S0378-1127(00)00710-6

Trapp, S., and Croteau, R. (2001). Defensive resin biosynthesis in conifers. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 52, 689–724. doi: 10.1146/annurev.arplant.52.1.689

Vanhercke, T., El Tahchy, A., Liu, Q., Zhou, X. R., Shrestha, P., Divi, U. K., et al. (2014). Metabolic engineering of biomass for high energy density: oilseed-like triacylglycerol yields from plant leaves. *Plant Biotechnol. J.* 12, 231–239. doi: 10.1111/pbi.12131

Vranova, E., Coman, D., and Gruissem, W. (2013). Network analysis of the MVA and MEP pathways for isoprenoid synthesis. *Annu. Rev. Plant Biol.* 64, 665–700. doi: 10.1146/annurev-arplant-050312-120116

Wang, H. H., Isaacs, F. J., Carr, P. A., Sun, Z. Z., Xu, G., Forest, C. R., et al. (2009). Programming cells by multiplex genome engineering and accelerated evolution. *Nature* 460, 894–898. doi: 10.1038/nature08187

Wegrzyn, J. L., Liechty, J. D., Stevens, K. A., Wu, L. S., Loopstra, C. A., Vasquez-Gross, H. A., et al. (2014). Unique features of the loblolly pine (*Pinus taeda* L.) megagenome revealed through sequence annotation. *Genetics* 196, 891–909. doi: 10.1534/genetics.113.159996

Westbrook, J. W., Walker, A. R., Neves, L. G., Munoz, P., Resende, M. F. Jr., Neale, D. B., et al. (2015). Discovering candidate genes that regulate resin canal number in Pinus stems by integrating association genetics and QTL analysis across environments, ages, and populations. *New Phytol.* 205, 627–641. doi: 10.1111/nph.13074

Zhang, J., Nieminen, K., Serra, J. A., and Helariutta, Y. (2014). The formation of wood and its control. *Curr. Opin. Plant Biol.* 17, 56–63. doi: 10.1016/j.pbi.2013.11.003

Zimin, A., Stevens, K. A., Crepeau, M. W., Holtz-Morris, A., Koriabine, M., Marçais, G., et al. (2014). Sequencing and assembly of the 22-gb loblolly pine genome. *Genetics* 196, 875–890. doi: 10.1534/genetics.113.159715

Zulak, K. G., and Bohlmann, J. (2010). Terpenoid biosynthesis and specialized vascular cells of conifer defense. *J. Integr. Plant Biol.* 52, 86–97. doi: 10.1111/j.1744-7909.2010.00910.x

frontiers
in Plant Science

Systems and Synthetic Biology of Forest Trees: A Bioengineering Paradigm for Woody Biomass Feedstocks

Alexander A. Myburg[1]*, Steven G. Hussey[1], Jack P. Wang[2], Nathaniel R. Street[3] and Eshchar Mizrachi[1]

[1]Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Hatfield, South Africa, [2]Forest Biotechnology Group, Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC, United States, [3]Umeå Plant Science Center, Department of Plant Physiology, Umeå University, Umeå, Sweden

Fast-growing forest plantations are sustainable feedstocks of plant biomass that can serve as alternatives to fossil carbon resources for materials, chemicals, and energy. Their ability to efficiently harvest light energy and carbon from the atmosphere and sequester this into metabolic precursors for lignocellulosic biopolymers and a wide range of plant specialized metabolites make them excellent biochemical production platforms and living biorefineries. Their large sizes have facilitated multi-omics analyses and systems modeling of key biological processes such as lignin biosynthesis in trees. High-throughput 'omics' approaches have also been applied in segregating tree populations where genetic variation creates abundant genetic perturbations of system components allowing construction of systems genetics models linking genes and pathways to complex trait variation. With this information in hand, it is now possible to start using synthetic biology and genome editing techniques in a bioengineering approach based on a deeper understanding and rational design of biological parts, devices, and integrated systems. However, the complexity of the biology and interacting components will require investment in big data informatics, machine learning, and intuitive visualization to fully explore multi-dimensional patterns and identify emergent properties of biological systems. Predictive systems models could be tested rapidly through high-throughput synthetic biology approaches and multigene editing. Such a bioengineering paradigm, together with accelerated genomic breeding, will be crucial for the development of a new generation of woody biorefinery crops.

Keywords: synthetic biology (synbio), systems biology, systems genetics, woody biomass, biorefinery, bioeconomy, lignin biosynthesis, wood formation

INTRODUCTION

Compared to herbaceous plants, forest trees afford numerous of advantages to plant biologists interested in studying growth and development. Most obviously, trees produce vast quantities of wood comprising multiple cells types produced from the meristematic cambial initials of the vascular cambium. In the context of developmental studies of secondary growth, greater size

represents greater spatial resolution to profile stages of development during the formation of secondary phloem and xylem as recently demonstrated for aspen (Obudulu et al., 2016; Sundell et al., 2017) and Norway spruce (Jokipii-Lukkari et al., 2017). Such studies have been instrumental in identifying the individual genes, proteins, metabolites, and pathways comprising the molecular components of biological processes such as wood formation. Many wood formation genes have been targeted using transgenic approaches to demonstrate the ability to modify wood properties (Chang et al., 2018). Unfortunately, compared to greenhouse results, some of these single gene modifications have had adverse effects on growth or altered transgene effect when tested in field trials (Leplé et al., 2007; Voelker et al., 2010). Such outcomes suggest a need for greater understanding of the interaction of systems components with intrinsic and extrinsic factors that differ between the greenhouse and field (Beckers et al., 2016).

In the context of industrial application, woody biomass is a renewable, carbon-neutral source of lignocellulosic materials for construction, pulp, bioenergy and, increasingly, for advanced biomaterials such as nanocellulose (Kunaver et al., 2016; Thomas et al., 2018). These and other biorefinery products, including novel biopolymers and biochemicals, are key to the emerging plant-based bio (Vanholme et al., 2013a; Van de Wouwer et al., 2018). Most woody biomass traits of commercial value are genetically complex. As such, it is challenging to devise strategies determining which genes contribute to these complex phenotypes and to associate these with causal mechanisms. In contrast to genomic selection that can be used as a "black box" route to improve complex traits, bioengineering approaches require systems-level understanding to elucidate the molecular basis of complex traits. "Multi-omics" analyses, together with extensive transgenic perturbation of a biological pathway and mathematical modeling of pathway dynamics, have recently been successfully applied to generate the most detailed systems biology model yet, of lignin biosynthesis, in a forest tree (Wang et al., 2018). The large amount of genetic diversity retained in tree breeding programs has also facilitated the use of natural variation to dissect complex traits (Mizrachi and Myburg, 2016; Mizrachi et al., 2017). Such systems genetics approaches pose opportunities for rapid advances in genomic breeding and provide diverse genetic backgrounds for genetic engineering of biomass traits.

Synthetic biology ("SynBio") aims to reduce biologically complex systems into discrete functional components that can be combined in numerous ways to create efficient and novel biological products or properties (for a synopsis of SynBio development, see Cameron et al., 2014). This transdisciplinary field draws on engineering principles and standardization of the fabrication process (i.e., gene circuit assembly) (Heinemann and Panke, 2006). Not only can synthetic biology incorporate existing genetic components (biological "parts") across the three Domains of life, but chemical gene and whole-genome synthesis, directed evolution, model-informed protein engineering, and synthetic regulatory machinery design enable the creation of xenobiological systems. The long-lived nature of trees and their vast reserves of carbon-rich sink tissues present both a challenge and opportunity for accumulating and storing synthetically produced compounds (Wilkerson et al., 2014; Mottiar et al., 2016).

Synthetically modified trees could serve as renewable factories for the production of large quantities of custom-designed biomass and biochemicals that do not exist naturally in trees.

We present a brief review of recent progress and our perspective on the application of systems biology approaches to understand complex tree biology, from molecular to organismal level crucial for sustainable production of woody biomass. We propose a bioengineering paradigm based on rational design, drawing on systems-level understanding of biological processes – and the application of synthetic biology and genome editing technologies – to engineer such processes in combination with accelerated, genome-assisted breeding of forest trees.

SYSTEMS BIOLOGY: INTEGRATIVE UNDERSTANDING OF BIOLOGICAL PROCESSES IN FOREST TREES

Systems modeling of biological processes requires extensive perturbation of systems components to produce the comprehensive experimental data required for modeling (**Figure 1**). The lack of extensive mutant collections in forest trees has necessitated the use of transgenic approaches to perturb genes. This has limited such studies to a relatively small number of genes, typically in the same biological pathway. Lignin biosynthesis has emerged as one of the most well-studied biological processes in forest trees, largely due to the interest in reducing lignin content or altering lignin composition to reduce the recalcitrance of woody biomass during industrial processing (Pilate et al., 2002; Leplé et al., 2007; Voelker et al., 2010; Mansfield et al., 2012; Bryan et al., 2016; Saleme et al., 2017). The first comprehensive, systems-level analysis of the lignin biosynthetic pathway was performed in *Arabidopsis* (Vanholme et al., 2012). This resulted in the discovery of additional components of the biological process, in particular, a novel component in the form of a caffeoyl shikimate esterase (CSE) that hydrolyzes caffeoyl shikimate into caffeate and, when mutated, leads to a fourfold increase in glucose release from cellulose during saccharification without pretreatment in *Arabidopsis* (Vanholme et al., 2013b). In poplar downregulated for CSE, 60% more glucose was released from wood without pretreatment, as a consequence of the lower lignin amount and the higher cellulose content (Saleme et al., 2017).

The most extensive analysis of systems biology in forest trees has been done on the lignin biosynthetic pathway in differentiating xylem of *Populus trichocarpa* (Wang et al., 2018). Transgenic trees systematically perturbed in the expression of lignin genes were produced, and quantitative data on genomics, transcriptomics, fluxomics, biochemical, chemical, and cellular analyses were integrated to construct a mathematical description of the pathway. The lignin model calculates how changing expression of any pathway gene or gene combination affects protein abundance, metabolic-flux, and phenotypic traits, including lignin content and composition, tree growth, wood density, and saccharification, for sugar release. The model predicts improvements in any of these traits individually or in combinations, through modifying

the expression of specific lignin genes. The lignin pathway is amenable to modeling because major pathway components have been identified and methodologies for their quantification have been developed (Shi et al., 2010; Vanholme et al., 2010, 2019; Wang et al., 2014, 2018). Genomic sequence information (Goujon et al., 2003; Raes et al., 2003; Tuskan et al., 2006; Myburg et al., 2014; Carocha et al., 2015) defined the gene families of the known proteins that are components of the lignin biosynthetic pathway. Some enzymes are encoded by single genes and others are members of large gene families (Goujon et al., 2003; Shi et al., 2010), while other members of the families are not associated with lignification in time or place. This highlights the importance of identifying those family members that are *bona fide* components of the systems model of wood lignification. The abundance of lignin enzymes *in vivo* has been measured in *P. trichocarpa* to determine the quantitative relationship between transcripts and proteins, and the enzymes in the pathway have been purified and characterized biochemically and genetically so reliable kinetics data could be obtained for substrates and inhibitors (Wang et al., 2014).

The systems analysis of lignin biosynthesis in *P. trichocarpa* has provided valuable information on the formation and function of lignin in forest trees, such as where genetic regulation is significantly different from non-woody species (e.g., *Arabidopsis*). For example, in the case of 4-coumarate CoA ligase (4CL), *Arabidopsis* has a single gene and enzyme for 4CL, while *P. trichocarpa* has two distinct *4CL* genes, one encoding a protein with a regulatory function (Chen et al., 2014). Similar regulatory functions have been demonstrated for genes upstream of lignin biosynthesis in the shikimate pathway (Xie et al., 2018) supporting the 4CL finding. *In silico* simulations furthermore predicted the consequence on flux and wood traits for every possible combination of multigene perturbations where expression of lignin genes was either upregulated, downregulated, or remained at the wild-type level. Systems modeling therefore provides a unique opportunity to guide multi-trait engineering and breeding strategies to create novel tree feedstocks optimized for fast growth and conversion to energy and materials. The models also inform identification of alleles or combinations of alleles to breed superior forest tree varieties (Wang et al., 2018).

SYSTEMS GENETICS: DISSECTING SYSTEMS COMPONENTS AND INTERACTIONS UNDERLYING COMPLEX TRAITS

While systems biology studies have advanced our understanding of developmental and stress response pathways (especially in identifying key regulators within these processes), they frequently are based on severe phenotypes that represent an aggregate of a cascade of cellular and physiological reactions to the adverse effect of the perturbation. Furthermore, such studies do not resolve the molecular basis or mechanisms of trait variation among individuals within populations. Revealing how complex trait variation arises from genetic polymorphisms

affords opportunities to discover novel approaches to inducing trait variation in addition to advancing understanding of the evolutionary processes and the mechanistic basis of variation in complex traits (**Figure 1**).

Complex traits are the emergent outcome of actions and interactions of hundreds to thousands of genes, functioning within pathways in combination with environmental effects. Variation at any biological layer from genome to trait can contribute to the final outcome. Full understanding of complex phenotypes requires approaches that can model these individual actions and interactions at the population level. Systems genetics provides a clearer path linking the genome to complex phenotypes by deducing the mechanisms by which genetic variation impacts variation in these intermediate biological layers (e.g., transcript, protein, and/or metabolite) to ultimately impact the complex trait (Civelek and Lusis, 2014; Feltus, 2014). Systems genetics has benefited substantially from recent sequencing technology improvements that have enabled population-wide genome resequencing and global transcript profiling *via* mRNA and small RNA sequencing. While SNP genotyping typically provides adequate genetic resolution for controlled crosses (QTL analysis using pedigrees), genome-wide association studies (GWAS) using natural populations often mandate whole-genome re-sequencing due to the rapid decay of LD in many tree species (McKown et al., 2014; Zhang et al., 2018). There have now been a number of studies in trees mapping the genetic architecture of gene expression variation (eQTLs) in tree pedigrees (Kirst et al., 2004, 2005; Drost et al., 2010) and natural populations (Porth et al., 2013; Mähler et al., 2017; Zhang et al., 2018); however, there is only one reported true systems genetics study (Mizrachi et al., 2017) in addition to genetical genomics approaches (Street et al., 2006; Drost et al., 2015).

These studies have reported contrasting results between pedigrees and natural populations, suggesting that the genetic background is an important consideration, especially for extrapolating findings to a wider context. Due to developments in sequencing technologies, there has been greater focus on eQTL studies; however, variation in protein (allelic) structural variation, epigenomic variation (Cheung et al., 2017), protein abundance (Consoli et al., 2002), and post translational modifications (Cesnik et al., 2016) or metabolome variation (Morreel et al., 2006; Joseph et al., 2014; Matsuda et al., 2015) fit equally well and indeed would be as responsible for variation, within the systems genetics framework. While systems genetics studies can reveal, for example, genomic loci linked to gene expression variation contributing to trait variation, it can be problematic to place these within a biological framework to understand how these novel mechanisms influence the trait. One approach to overcoming this barrier can be integration of systems biology studies, such as developmental studies, to provide context to these novel genes. Combining systems genetics and systems biology co-expression networks is also a powerful approach as biological inference can be aided by considering the neighborhood of genes connected to a novel candidate (Mähler et al., 2017).

To date, one level of complexity that has not been considered in eQTL and systems genetics studies is transcriptional plasticity.

FIGURE 1 | Systems approaches to study biological processes in trees. Axes represent different dimensions or types of perturbation that can be used to dissect biological processes. Network components (colored nodes and shapes) represent the study of different biological layers (DNA, RNA, protein, metabolite, etc.) using integrative approaches that take into account development/condition (*Systems biology*), genetic diversity (*Systems genetics*), and evolution (*Evo-devo*). Yellow nodes in "systems genetics" indicate the connection to complex traits that can be established in population-wide studies. Networks on the left represent intermediate integration steps of components that vary along these axes. Networks on the right represent integrated networks that can be constructed from each type of perturbation. Boxes show target traits and processes associated with different tissues or organs where systems approaches may provide new opportunities for research.

There is now extensive knowledge of alternative splicing and transcript usage, which can vary among tissues, during development or among genotypes (Bao et al., 2013; Xu et al., 2014; Zhao et al., 2014). Mapping expression at the transcript level will provide greater resolution to the link between genome and phenotype. However, fine-scale expression analysis at the transcript level remains challenging, mainly because most tissue samples contain a complex mix of transcripts and splicing variants from different cell types or cells at different stages of development. There is equally a need for improvement in network inference methods, with an understanding that no single current method is adequate to capture the range of interactions present in biological networks (Marbach et al., 2012) and with few tools available to facilitate aggregate inference approaches (Schiffthaler et al., 2018). Similarly, no studies have yet integrated genome-wide assays of DNA modifications or accessibility despite increasing evidence of the additional insight such information brings. There is also a paucity of large-scale transcription factor binding or protein-protein interaction data for plants in general, which further limits comprehensive understanding.

SYNTHETIC BIOLOGY: A NEW BIOENGINEERING PARADIGM FOR FOREST TREES

SynBio has made its greatest advancements in prokaryotes and single-celled eukaryotes such as yeast, but plant synthetic biology is catching up (Patron et al., 2015; Schaumberg et al., 2016;

de Lange et al., 2018; Hanson and Jez, 2018; Pouvreau et al., 2018). In addition to a large number of modifications made by conventional transgenic approaches (Chang et al., 2018), there have been some notable successes in the synthetic modification of trees using single-gene strategies such as the introduction of chemically labile ester linkages into the lignin backbone of poplar trees through the xylem-specific expression of an exogenous feruloyl-coenzyme A monolignol transferase (FMT) from *Angelica sinensis* (Wilkerson et al., 2014). Future strategies will attempt to evaluate far more complex designs, relying to a large extent on the ability to assemble DNA fragments idempotently (that is, the flexibility to assemble basic parts with increasing complexity using a universal method and without having to re-modify each intermediate). There is currently a scarcity of freely available standardized biological parts suitable for plant biology akin to the International Genetically Engineered Machines (iGEM) BioBrick parts collection[1]. Encouragingly, the Phytobrick synthetic biology standard with a universal lexicon for plant gene elements coupled to powerful Type IIS idempotent assembly methods (Engler et al., 2014; Patron et al., 2015) is fast gaining traction, with an increasing number of Phytobricks now included in the iGEM Standard Registry[2]. Tree biologists and biotechnologists should adapt to this conceptual framework as soon as possible to keep up with synthetic biology development in annual crops. We recently developed an open access synthetic panel of 221

---

[1]http://parts.igem.org
[2]http://parts.igem.org/Collections/Plants

FIGURE 2 | A proposed bioengineering paradigm based on forest tree synthetic biology. **(i)** Systems biology models inform the initial design of synthetic multigene constructs. Individual components (DNA parts) are sourced from existing or novel genetic resources (e.g., bioprospecting), synthesized, and submitted to biological parts collections as standardized Phytobricks **(ii)**. The fabrication and testing phase **(iii)** involves high-throughput idempotent construct assembly followed by transformation and testing in a simplified chassis such as a cell culture derived from the target tree species **(iv)**. Construct expression and phenotypic data are then integrated into a machine learning model to optimize the construct design, an iterative process that produces a reduced number of semi-optimized constructs for validation in a genotype of the target species selected to perform favorably in laboratory conditions and conducive to transformation ("lab rat"). Such validation may include greenhouse trials involving juvenile trees or mature trees in field trials. **(v)** Successful constructs may be introduced into a preferred elite parental genotype for intraspecific or interspecific hybrid breeding. A possible avenue to rapidly mobilize synthetic gene constructs into more diverse genetic backgrounds would be to introduce an early flowering construct (such as overexpression of the *FLOWERING LOCUS T* gene; FT-OX) into a number of elite parental genotypes. Such genotypes can then be transformed with the optimized synthetic gene constructs and used as female parents in crosses with non-transgenic (wild-type, WT) parents to produce F1 progeny segregating for both constructs. If the two constructs are on different chromosomes, approximately 25% of the progeny should be WT for growth and flowering but contain the synthetic gene construct. Early flowering parental genotypes can be propagated *in vitro* to be transformed with various synthetic constructs and, if unrelated, different parental genotypes could be crossed for transgene stacking.

secondary cell wall-related *Eucalyptus grandis* transcription factors and 65 promoter sequences in partnership with the Department of Energy Joint Genome Institute, most of which were designed as Phytobricks (Hussey et al., in preparation). Accessibility of high-throughput DNA synthesis services will ensure that a growing number of standardized parts become available under open material transfer agreements.

One considerable challenge to tree synthetic biology is precise spatiotemporal control of complex multigene constructs, especially in woody tissues where inducing gene expression with external agents is impractical. Such constructs must function optimally in a tissue of interest, be resistant to eukaryotic silencing mechanisms such as RNAi or epigenetic silencing, be somatically stable such that somatic mutations that disrupt a synthetic construct should not be selectively favored, take into account compartmentalized plant cell biology, and have built-in biosafety mechanisms preventing transgene escape. Furthermore, synthetic gene circuits should consist of *composable* parts (de Lange et al., 2018) that individually encode defined and transferrable functions (a property known as *modularity*) and that function independently of endogenous processes to avoid unwanted interference, a property known as *orthogonality*. Designer transcription factors based on zinc finger, TALE, and dCas9 technologies targeting endogenous or synthetic promoters (Liu and Stewart, 2016) are ideal orthogonal synthetic tools

that allow considerable transgene regulation flexibility but may require extensive testing and optimization.

Currently, thousands of iterations of multigene constructs can be produced by robotics-assisted DNA Foundry services. However, it is not feasible to transform and phenotype thousands of transgenic trees for iterative design-build-test-learn cycles envisioned for plant synthetic biology (Pouvreau et al., 2018). Early synthetic designs will therefore have to be tested in a simplified system (or "chassis" in synthetic biology terminology) until optimized constructs can be evaluated in target tree species. This will necessitate the development of experimental chassis derived from the target tree species that are easier to transform and phenotype *en masse*, such as protoplasts, cell cultures, or agroinfiltrated leaves, or "lab rat" genotypes that perform well in tissue culture and have high transformation rates (**Figure 2**). *In vitro* tracheary element transdifferentiation approaches relying on hormonal induction (Fukuda and Komamine, 1980; Kubo et al., 2005; Saito et al., 2017), or the VND7 inducible system (Yamaguchi et al., 2010; Goué et al., 2013), for example, could be used to induce secondary cell wall formation in suspension culture cells or explant tissues of a target tree species and thus evaluate the phenotypes of many cell wall-modifying constructs before semi-optimized constructs are tested in a multicellular tree model. Successful constructs may either be directly introduced into an elite

genotype of the target species or rapidly crossed into elite breeding material after co-transformation of parental genotypes with an early flowering construct (**Figure 2**).

FUTURE PERSPECTIVE

Most fundamental discoveries, including proof-of-concept cell wall and growth modifications (and even extrapolations on biomass processing efficiency), are still derived from the analysis of *Arabidopsis* inflorescence stems, which remains a poor representation of large tree stems comprised mainly of wood. Of priority in the short term is testing genetic perturbations as much as possible in a woody model such as *Populus* and, if possible, directly in target species of interest. Several priorities must be met here, such as enhancing the transformation efficiency of commercial species or genotypes, capacity for large-scale transformation experiments, as well as (crucially) field trials confirming greenhouse phenotypes in mature trees. Large consortia and industry collaborations, as well as engagement and an improvement in the regulatory landscape, must be met for this to be truly realized.

Also in the short to medium term, the convergence of high-resolution technologies that capture genomics, epigenomics, and other cell 'omics', phenomics, and environment (including microbiome) data, as well as computational modeling of the interactions of these, requires transdisciplinary innovations and probably the application of artificial intelligence methodology. Combined with genome editing (with broader scale synthetic biology applications), this makes the field of forest biotechnology ripe for a new wave of creativity, especially in thinking of the tree itself as a living biorefinery and as a stable and continuous producer of specialized high-value compounds or polymers in sustainably harvestable tissues and organs such as leaves, secondary phloem, and bark. Higher resolution knowledge of metabolite precursors, tissue-specific pathway engineering, and knowledge of novel high-value derivatives that can be discovered using bioprospecting methods and produced in trees has the potential for a new generation of relatively low volume, but high value, forest products.

How far does the application of these technologies go? Given the long rotation times of forest trees as harvestable biomass crops, it is unlikely (and indeed not essential) that we will see movement toward a "bottom up" approach that builds on a synthetic minimal tree genome. It is much more important to optimize the precise introduction of complex regulatory circuits and metabolic pathways that remain stable through breeding generations, a nascent field of research in itself. Such a bioengineering paradigm, combined with advanced genomic breeding approaches and accelerated flowering technologies, may empower rapid development of woody biomass crops tailored for diverse biorefinery, biomaterials, and timber construction products. In many forest-growing countries, an advanced forest products industry will be one of the cornerstones of the bioeconomy and key to achieving global sustainable development goals.

AUTHOR CONTRIBUTIONS

All authors contributed to the drafting and editing of the manuscript.

FUNDING

REFERENCES

Bao, H., Li, E., Mansfield, S. D., Cronk, Q. C., El-Kassaby, Y. A., and Douglas, C. J. (2013). The developing xylem transcriptome and genome-wide analysis of alternative splicing in *Populus trichocarpa* (black cottonwood) populations. *BMC Genomics* 14:359. doi: 10.1186/1471-2164-14-359

Beckers, B., Op De Beeck, M., Weyens, N., Van Acker, R., Van Montagu, M., Boerjan, W., et al. (2016). Lignin engineering in field-grown poplar trees affects the endosphere bacterial microbiome. *Proc. Natl. Acad. Sci. USA* 113, 2312–2317. doi: 10.1073/pnas.1523264113

Bryan, A. C., Jawdy, S., Gunter, L., Gjersing, E., Sykes, R., Hinchee, M. A., et al. (2016). Knockdown of a laccase in *Populus deltoides* confers altered cell wall chemistry and increased sugar release. *Plant Biotechnol. J.* 14, 2010–2020. doi: 10.1111/pbi.12560

Cameron, D. E., Bashor, C. J., and Collins, J. J. (2014). A brief history of synthetic biology. *Nat. Rev. Microbiol.* 12, 381–390. doi: 10.1038/nrmicro3239

Carocha, V., Soler, M., Hefer, C., Cassan-Wang, H., Fevereiro, P., Myburg, A. A., et al. (2015). Genome-wide analysis of the lignin toolbox of *Eucalyptus grandis*. *New Phytol.* 206, 1297–1313. doi: 10.1111/nph.13313

Cesnik, A. J., Shortreed, M. R., Sheynkman, G. M., Frey, B. L., and Smith, L. M. (2016). Human proteomic variation revealed by combining RNA-Seq proteogenomics and global post-translational modification (G-PTM) search strategy. *J. Proteome Res.* 15, 800–808. doi: 10.1021/acs.jproteome.5b00817

Chang, S., Mahon, E. L., Mackay, H. A., Rottmann, W. H., Strauss, S. H., Pijut, P. M., et al. (2018). Genetic engineering of trees: progress and new horizons. *In Vitro Cell Dev. Biol. Plant* 54, 341–376. doi: 10.1007/s11627-018-9914-1

Chen, H. C., Song, J., Wang, J. P., Lin, Y. C., Ducoste, J., Shuford, C. M., et al. (2014). Systems biology of lignin biosynthesis in *Populus trichocarpa*: heteromeric 4-coumaric acid:coenzyme A ligase protein complex formation, regulation, and numerical modeling. *Plant Cell* 26, 876–893. doi: 10.1105/tpc.113.119685

Cheung, W. A., Shao, X., Morin, A., Siroux, V., Kwan, T., Ge, B., et al. (2017). Functional variation in allelic methylomes underscores a strong genetic contribution and reveals novel epigenetic alterations in the human epigenome. *Genome Biol.* 18:50. doi: 10.1186/s13059-017-1173-7

Civelek, M., and Lusis, A. J. (2014). Systems genetics approaches to understand complex traits. *Nat. Rev. Genet.* 15, 34–48. doi: 10.1038/nrg3575

Consoli, L., Lefevre, A., Zivy, M., De Vienne, D., and Damerval, C. (2002). QTL analysis of proteome and transcriptome variations for dissecting the genetic architecture of complex traits in maize. *Plant Mol. Biol.* 48, 575–581. doi: 10.1023/A:1014840810203

de Lange, O., Klavins, E., and Nemhauser, J. (2018). Synthetic genetic circuits in crop plants. *Curr. Opin. Biotechnol.* 49, 16–22. doi: 10.1016/j.copbio.2017.07.003

Drost, D. R., Benedict, C. I., Berg, A., Novaes, E., Novaes, C. R. D. B., Yu, Q., et al. (2010). Diversification in the genetic architecture of gene expression and transcriptional networks in organ differentiation of *Populus*. *Proc. Natl. Acad. Sci. USA* 107, 8492–8497. doi: 10.1073/pnas.0914709107

Drost, D. R., Puranik, S., Novaes, E., Novaes, C. R., Dervinis, C., Gailing, O., et al. (2015). Genetical genomics of *Populus* leaf shape variation. *BMC Plant Biol.* 15:166. doi: 10.1186/s12870-015-0557-7

Engler, C., Youles, M., Gruetzner, R., Ehnert, T. M., Werner, S., Jones, J. D., et al. (2014). A golden gate modular cloning toolbox for plants. *ACS Synth. Biol.* 3, 839–843. doi: 10.1021/sb4001504

Feltus, F. A. (2014). Systems genetics: a paradigm to improve discovery of candidate genes and mechanisms underlying complex traits. *Plant Sci.* 223, 45–48. doi: 10.1016/j.plantsci.2014.03.003

Fukuda, H., and Komamine, A. (1980). Establishment of an experimental system for the tracheary element differentiation from single cells isolated from the mesophyll of *Zinnia elegans*. *Plant Physiol.* 65, 57–60. doi: 10.1104/pp.65.1.57

Goué, N., Mortimer, J. C., Nakano, Y., Zhang, Z., Josserand, M., Ohtani, M., et al. (2013). Secondary cell wall characterization in a BY-2 inductive system. *Plant Cell Tissue Organ Cult.* 115, 223–232. doi: 10.1007/s11240-013-0354-7

Goujon, T., Sibout, R., Eudes, A., Mackay, J., and Joulanin, L. (2003). Genes involved in the biosynthesis of lignin precursors in *Arabidopsis thaliana*. *Plant Physiol. Biochem.* 41, 677–687. doi: 10.1016/S0981-9428(03)00095-0

Hanson, A. D., and Jez, J. M. (2018). Synthetic biology meets plant metabolism. *Plant Sci.* 273, 1–2. doi: 10.1016/j.plantsci.2018.04.004

Heinemann, M., and Panke, S. (2006). Synthetic biology--putting engineering into biology. *Bioinformatics* 22, 2790–2799. doi: 10.1093/bioinformatics/btl469

Jokipii-Lukkari, S., Sundell, D., Nilsson, O., Hvidsten, T. R., Street, N. R., and Tuominen, H. (2017). NorWood: a gene expression resource for evo-devo studies of conifer wood development. *New Phytol.* 216, 482–494. doi: 10.1111/nph.14458

Joseph, B., Atwell, S., Corwin, J. A., Li, B., and Kliebenstein, D. J. (2014). Meta-analysis of metabolome QTLs in *Arabidopsis*: trying to estimate the network size controlling genetic variation of the metabolome. *Front. Plant Sci.* 5:461. doi: 10.3389/fpls.2014.00461

Kirst, M., Basten, C. J., Myburg, A. A., Zeng, Z. B., and Sederoff, R. R. (2005). Genetic architecture of transcript-level variation in differentiating xylem of a *Eucalyptus* hybrid. *Genetics* 169, 2295–2303. doi: 10.1534/genetics.104.039198

Kirst, M., Myburg, A. A., De Leon, J. P., Kirst, M. E., Scott, J., and Sederoff, R. (2004). Coordinated genetic regulation of growth and lignin revealed by quantitative trait locus analysis of cDNA microarray data in an interspecific backcross of *Eucalyptus*. *Plant Physiol.* 135, 2368–2378. doi: 10.1104/pp.103.037960

Kubo, M., Udagawa, M., Nishikubo, N., Horiguchi, G., Yamaguchi, M., Ito, J., et al. (2005). Transcription switches for protoxylem and metaxylem vessel formation. *Genes Dev.* 19, 1855–1860. doi: 10.1101/gad.1331305

Kunaver, M., Anzlovar, A., and Zagar, E. (2016). The fast and effective isolation of nanocellulose from selected cellulosic feedstocks. *Carbohydr. Polym.* 148, 251–258. doi: 10.1016/j.carbpol.2016.04.076

Leplé, J. C., Dauwe, R., Morreel, K., Storme, V., Lapierre, C., Pollet, B., et al. (2007). Downregulation of cinnamoyl-coenzyme A reductase in poplar: multiple-level phenotyping reveals effects on cell wall polymer metabolism and structure. *Plant Cell* 19, 3669–3691. doi: 10.1105/tpc.107.054148

Liu, W., and Stewart, C. N. Jr. (2016). Plant synthetic promoters and transcription factors. *Curr. Opin. Biotechnol.* 37, 36–44. doi: 10.1016/j.copbio.2015.10.001

Mähler, N., Wang, J., Terebieniec, B. K., Ingvarsson, P. K., Street, N. R., and Hvidsten, T. R. (2017). Gene co-expression network connectivity is an important determinant of selective constraint. *PLoS Genet.* 13:e1006402. doi: 10.1371/journal.pgen.1006402

Mansfield, S. D., Kang, K. Y., and Chapple, C. (2012). Designed for deconstruction– poplar trees altered in cell wall lignification improve the efficacy of bioethanol production. *New Phytol.* 194, 91–101. doi: 10.1111/j.1469-8137.2011.04031.x

Marbach, D., Costello, J. C., Küffner, R., Vega, N. M., Prill, R. J., Camacho, D. M., et al. (2012). Wisdom of crowds for robust gene network inference. *Nat. Methods* 9, 796–804. doi: 10.1038/nmeth.2016

Matsuda, F., Nakabayashi, R., Yang, Z., Okazaki, Y., Yonemaru, J., Ebana, K., et al. (2015). Metabolome-genome-wide association study dissects genetic architecture for generating natural variation in rice secondary metabolism. *Plant J.* 81, 13–23. doi: 10.1111/tpj.12681

Mckown, A. D., Klapste, J., Guy, R. D., Geraldes, A., Porth, I., Hannemann, J., et al. (2014). Genome-wide association implicates numerous genes underlying ecological trait variation in natural populations of *Populus trichocarpa*. *New Phytol.* 203, 535–553. doi: 10.1111/nph.12815

Mizrachi, E., and Myburg, A. A. (2016). Systems genetics of wood formation. *Curr. Opin. Plant Biol.* 30, 94–100. doi: 10.1016/j.pbi.2016.02.007

Mizrachi, E., Verbeke, L., Christie, N., Fierro, A. C., Mansfield, S. D., Davis, M. F., et al. (2017). Network-based integration of systems genetics data reveals pathways associated with lignocellulosic biomass accumulation and processing. *Proc. Natl. Acad. Sci. USA* 114, 1195–1200. doi: 10.1073/pnas.1620119114

Morreel, K., Goeminne, G., Storme, V., Sterck, L., Ralph, J., Coppieters, W., et al. (2006). Genetical metabolomics of flavonoid biosynthesis in *Populus*: a case study. *Plant J.* 47, 224–237. doi: 10.1111/j.1365-313X.2006.02786.x

Mottiar, Y., Vanholme, R., Boerjan, W., Ralph, J., and Mansfield, S. D. (2016). Designer lignins: harnessing the plasticity of lignification. *Curr. Opin. Biotechnol.* 37, 190–200. doi: 10.1016/j.copbio.2015.10.009

Myburg, A. A., Grattapaglia, D., Tuskan, G. A., Hellsten, U., Hayes, R. D., Grimwood, J., et al. (2014). The genome of *Eucalyptus grandis*. *Nature* 510, 356–362. doi: 10.1038/nature13308

Obudulu, O., Bygdell, J., Sundberg, B., Moritz, T., Hvidsten, T. R., Trygg, J., et al. (2016). Quantitative proteomics reveals protein profiles underlying major transitions in aspen wood development. *BMC Genomics* 17:119. doi: 10.1186/s12864-016-2458-z

Patron, N. J., Orzaez, D., Marillonnet, S., Warzecha, H., Matthewman, C., Youles, M., et al. (2015). Standards for plant synthetic biology: a common syntax for exchange of DNA parts. *New Phytol.* 208, 13–19. doi: 10.1111/nph.13532

Pilate, G., Guiney, E., Holt, K., Petit-Conil, M., Lapierre, C., Leple, J. C., et al. (2002). Field and pulping performances of transgenic trees with altered lignification. *Nat. Biotechnol.* 20, 607–612. doi: 10.1038/nbt0602-607

Porth, I., Klapste, J., Skyba, O., Friedmann, M. C., Hannemann, J., Ehlting, J., et al. (2013). Network analysis reveals the relationship among wood properties, gene expression levels and genotypes of natural *Populus trichocarpa* accessions. *New Phytol.* 200, 727–742. doi: 10.1111/nph.12419

Pouvreau, B., Vanhercke, T., and Singh, S. (2018). From plant metabolic engineering to plant synthetic biology: the evolution of the design/build/test/learn cycle. *Plant Sci.* 273, 3–12. doi: 10.1016/j.plantsci.2018.03.035

Raes, J., Rohde, A., Christensen, J., Van De Peer, Y., and Boerjan, W. (2003). Genome-wide characterization of the lignification toolbox in *Arabidopsis*. *Plant Physiol.* 133, 1051–1071. doi: 10.1104/pp.103.026484

Saito, M., Nurani, A. M., Kondo, Y., and Fukuda, H. (2017). Tissue culture for xylem differentiation with *Arabidopsis* leaves. *Methods Mol. Biol.* 1544, 59–65. doi: 10.1007/978-1-4939-6722-3_5

Saleme, M. L. S., Cesarino, I., Vargas, L., Kim, H., Vanholme, R., Goeminne, G., et al. (2017). Silencing CAFFEOYL SHIKIMATE ESTERASE affects lignification and improves saccharification in poplar. *Plant Physiol.* 175, 1040–1057. doi: 10.1104/pp.17.00920

Schaumberg, K. A., Antunes, M. S., Kassaw, T. K., Xu, W., Zalewski, C. S., Medford, J. I., et al. (2016). Quantitative characterization of genetic parts and circuits for plant synthetic biology. *Nat. Methods* 13, 94–100. doi: 10.1038/nmeth.3659

Schiffthaler, B., Serrano, A., Delhomme, N., and Street, N.R. (2018). Seidr: a toolkit for calculation of crowd networks. *bioRxiv* [Preprint]. doi: 10.1101/250696

Shi, R., Sun, Y. H., Li, Q., Heber, S., Sederoff, R., and Chiang, V. L. (2010). Towards a systems approach for lignin biosynthesis in *Populus trichocarpa*: transcript abundance and specificity of the monolignol biosynthetic genes. *Plant Cell Physiol.* 51, 144–163. doi: 10.1093/pcp/pcp175

Street, N. R., Skogstrom, O., Sjodin, A., Tucker, J., Rodriguez-Acosta, M., Nilsson, P., et al. (2006). The genetics and genomics of the drought response in *Populus*. *Plant J.* 48, 321–341. doi: 10.1111/j.1365-313X.2006.02864.x

Sundell, D., Street, N. R., Kumar, M., Mellerowicz, E. J., Kucukoglu, M., Johnsson, C., et al. (2017). AspWood: high-spatial-resolution transcriptome

profiles reveal uncharacterized modularity of wood formation in *Populus tremula*. *Plant Cell* 29, 1585–1604. doi: 10.1105/tpc.17.00153

Thomas, B., Raj, M. C., B, A. K., H, R. M., Joy, J., Moores, A., et al. (2018). Nanocellulose, a versatile green platform: from biosources to materials and their applications. *Chem. Rev.* 118, 11575–11625. doi: 10.1021/acs.chemrev.7b00627

Tuskan, G. A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313, 1596–1604. doi: 10.1126/science.1128691

Van de Wouwer, D., Boerjan, W., and Vanholme, B. (2018). Plant cell wall sugars: sweeteners for a bio-based economy. *Physiol. Plant.* 164, 27–44. doi: 10.1111/ppl.12705

Vanholme, R., Cesarino, I., Rataj, K., Xiao, Y., Sundin, L., Goeminne, G., et al. (2013b). Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in *Arabidopsis*. *Science* 341, 1103–1106. doi: 10.1126/science.1241602

Vanholme, R., De Meester, B., Ralph, J., and Boerjan, W. (2019). Lignin biosynthesis and its integration into metabolism. *Curr. Opin. Biotechnol.* 56, 230–239. doi: 10.1016/j.copbio.2019.02.018

Vanholme, R., Demedts, B., Morreel, K., Ralph, J., and Boerjan, W. (2010). Lignin biosynthesis and structure. *Plant Physiol.* 153, 895–905. doi: 10.1104/pp.110.155119

Vanholme, B., Desmet, T., Ronsse, F., Rabaey, K., Van Breusegem, F., De Mey, M., et al. (2013a). Towards a carbon-negative sustainable bio-based economy. *Front. Plant Sci.* 4:174. doi: 10.3389/fpls.2013.00174

Vanholme, R., Storme, V., Vanholme, B., Sundin, L., Christensen, J. H., Goeminne, G., et al. (2012). A systems biology view of responses to lignin biosynthesis perturbations in *Arabidopsis*. *Plant Cell* 24, 3506–3529. doi: 10.1105/tpc.112.102574

Voelker, S. L., Lachenbruch, B., Meinzer, F. C., Jourdes, M., Ki, C., Patten, A. M., et al. (2010). Antisense down-regulation of 4CL expression alters lignification, tree growth, and saccharification potential of field-grown poplar. *Plant Physiol.* 154, 874–886. doi: 10.1104/pp.110.159269

Wang, J. P., Matthews, M. L., Williams, C. M., Shi, R., Yang, C., Tunlaya-Anukit, S., et al. (2018). Improving wood properties for wood utilization through multi-omics integration in lignin biosynthesis. *Nat. Commun.* 9:1579. doi: 10.1038/s41467-018-03863-z

Wang, J. P., Naik, P. P., Chen, H. C., Shi, R., Lin, C. Y., Liu, J., et al. (2014). Complete proteomic-based enzyme reaction and inhibition kinetics reveal how monolignol biosynthetic enzyme families affect metabolic flux and lignin in *Populus trichocarpa*. *Plant Cell* 26, 894–914. doi: 10.1105/tpc.113.120841

Wilkerson, C. G., Mansfield, S. D., Lu, F., Withers, S., Park, J. Y., Karlen, S. D., et al. (2014). Monolignol ferulate transferase introduces chemically labile linkages into the lignin backbone. *Science* 344, 90–93. doi: 10.1126/science.1250161

Xie, M., Muchero, W., Bryan, A. C., Yee, K., Guo, H. B., Zhang, J., et al. (2018). A 5-enolpyruvylshikimate 3-phosphate synthase functions as a transcriptional repressor in *Populus*. *Plant Cell* 30, 1645–1660. doi: 10.1105/tpc.18.00168

Xu, P., Kong, Y. M., Song, D. L., Huang, C., Li, X., and Li, L. G. (2014). Conservation and functional influence of alternative splicing in wood formation of *Populus* and *Eucalyptus*. *BMC Genomics* 15:780. doi: 10.1186/1471-2164-15-780

Yamaguchi, M., Goue, N., Igarashi, H., Ohtani, M., Nakano, Y., Mortimer, J. C., et al. (2010). VASCULAR-RELATED NAC-DOMAIN6 and VASCULAR-RELATED NAC-DOMAIN7 effectively induce transdifferentiation into xylem vessel elements under control of an induction system. *Plant Physiol.* 153, 906–914. doi: 10.1104/pp.110.154013

Zhang, J., Yang, Y., Zheng, K., Xie, M., Feng, K., Jawdy, S. S., et al. (2018). Genome-wide association studies and expression-based quantitative trait loci analyses reveal roles of HCT2 in caffeoylquinic acid biosynthesis and its regulation by defense-responsive transcription factors in *Populus*. *New Phytol.* 220, 502–516. doi: 10.1111/nph.15297

Zhao, Y., Sun, J., Xu, P., Zhang, R., and Li, L. (2014). Intron-mediated alternative splicing of WOOD-ASSOCIATED NAC TRANSCRIPTION FACTOR1B regulates cell wall thickening during fiber development in *Populus* species. *Plant Physiol.* 164, 765–776. doi: 10.1104/pp.113.231134

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor declared a past co-authorship with one of the authors NS.

Opportunities for Innovation in Genetic Transformation of Forest Trees

Michael Nagle[1†], Annabelle Déjardin[2†], Gilles Pilate[2]* and Steven H. Strauss[1]*

[1] Forest Ecosystems and Society, Molecular and Cellular Biology, Oregon State University, Corvallis, OR, United States,
[2] BioForA, INRA, ONF, Orléans, France

The incorporation of DNA into plant genomes followed by regeneration of non-chimeric stable plants (transformation) remains a major challenge for most plant species. Forest trees are particularly difficult as a result of their biochemistry, aging, desire for clonal fidelity, delayed reproduction, and high diversity. We review two complementary approaches to transformation that appear to hold promise for forest trees.

Keywords: transformation, regeneration, *WUSCHEL*, *BABY BOOM*, *Populus*, organogenesis, embryogenesis, Agrobacterium

SUMMARY

Developmental genes (DG) may be useful tools for promoting transformation. DGs, which can act through a wide variety of developmental mechanisms to promote regeneration of transgenic cells, have been widely employed in model plants to promote embryogenesis and in some cases organogenesis. Following initial experimental demonstration in dicots, the DGs *WUSCHEL* and/or *BABY BOOM* have formed the basis of a high efficiency method for a variety of monocot genotypes and species. However, in dicots the utility of these genes as the basis of a robust transformation system has not been demonstrated. Many additional DGs appear capable of promoting regeneration that have not been systematically explored as transformation tools.

Because *in vitro* plant transformation systems are costly and must be customized for each new genotype and species, *in vivo* approaches to transformation hold much appeal. It is possible to produce stable transgenic plants by agro-inoculation of seeds or vegetative/floral buds, but as yet these approaches have not been used routinely in any plant species except for the Arabidopsis floral dip. We will discuss how the Arabidopsis system, and other *in planta* techniques, may be tailored for forest trees, taking into account variations in biology of different taxa.

DEVELOPMENTAL GENES AS TOOLS FOR TRANSFORMATION IMPROVEMENT

When overexpressed, transcriptional or epigenetic regulators of embryo and meristem development (referred to as developmental genes, DGs) have been shown to confer improvements to *in vitro* regenerability. Recent molecular evidence places these genes within a genetic regulatory network, connected by cascades and feedback loops of transactivation. Knowledge of these interactions, detailed in **Supplementary Table 1** and **Figure 1**, can inform selection of individual genes and combinations of genes that may be most effective for improving regeneration. In this mini-review, however, we do not consider the many genes which may enhance regeneration via

FIGURE 1 | Portion of genetic regulatory network with demonstrated ability to enhance regeneration via overexpression. Green nodes represent phytohormones and green arrows indicate transactivation that results from upstream hormone-triggered cascades (not shown). Auxin leads to PLT upregulation through a relatively well-characterized cascade (Horstman et al., 2014), as do cytokinins for B-type ARRs (Xie et al., 2018). Blue nodes represent genes which have been shown to enhance regeneration when overexpressed, and blue arrows indicate direct targets of their transactivation activity (detailed in **Supplementary Table 1**).

epigenetic mechanisms, or by affecting the rate of gene transfer or incorporation of DNA into the genome.

Candidate Genes and Their Modes of Action

The first gain-of-function mutation for a transcriptional regulator of shoot induction was observed in 1941. Bryan and Sass discovered a heritable trait in maize that causes leaves to develop ectopic shoot meristems, or "knots" (Bryan and Sass, 1941). Transposon tagging later revealed the responsible locus to be *Knotted-1*, the first homeodomain transcription factor identified in plants (Hake et al., 1989). *Knotted-1* was found to be overexpressed in mutants as a result of a transposon insertion (Smith et al., 1992), and overexpression of maize *Knotted-1* in tobacco and Arabidopsis was also reported to trigger development of ectopic shoot meristems (Sinha et al., 1993; Lincoln et al., 1994).

A loss-of-function mutation of *SHOOT MERISTEMLESS* (*STM*), an Arabidopsis homolog of *Knotted-1*, was reported to lead to premature termination of shoot meristems and a phenotype of twig-like plants lacking lateral shoots (Long et al., 1996). Transcription of *STM* is regulated in part by a positive feedback loop between *STM* and the transactivator *CUP-SHAPED COTYLEDON 1* (*CUC1*) (**Supplementary Table 1** and **Figure 1**; Takada et al., 2001; Spinelli et al., 2011). Similarly to STM, overexpression of CUC1 enhances regeneration through embryogenic and organogenic pathways (Takada et al., 2001; Hibara et al., 2003). Overexpression of *CUC1/2* in Arabidopsis led to large increases in numbers of shoot from calli, while knockout

had the opposite effect (Daimon et al., 2003). Directly upstream of *CUC1* is *LEAFY COTYLEDON1 (LEC1)*, which was first identified as a regulator of embryogenesis when overexpression induced embryos to develop on leaves of Arabidopsis (Lotan et al., 1998).

A mutant screen for shoot meristem defects led to the discovery of *WUSCHEL (WUS)* loss-of-function mutants, which had similar phenotypes to *stm* mutants (Laux et al., 1996). *WUS* is expressed in the organizing center adjacent to meristematic stem cells (Mayer et al., 1998), then trafficked into the central zone of the stem cell niche (Yadav et al., 2010; Daum et al., 2014), where it is required for maintenance of stem cell identity. Contrarily, *STM* is expressed and active in the peripheral zone (PZ) (Williams et al., 2005; Gordon et al., 2007). In Arabidopsis, overexpression of *WUS* was reported to initiate ectopic organogenesis *in vivo*, although differentiation into organs was incomplete unless *STM* was also overexpressed (Gallois et al., 2002). *WUS* was rediscovered in a T-DNA activation mutagenesis screen for gain-of-function mutations conferring cytokinin-independent potential for induced embryogenesis *in vitro* (Zuo et al., 2002).

Downstream of *WUS* and *STM* are overlapping sets of diverse genes which balance differentiation and dedifferentiation to promote progressive development of meristems. These genes include members of the SE-associated receptor kinase (SERK) family, as well as enzymes for hormone biosynthesis, cell cycle regulators, and numerous transcriptional regulators, some of which function not only downstream, but upstream of *WUS/STM* (Spinelli et al., 2011; Balkunde et al., 2017; Ikeuchi et al., 2018; Scofield et al., 2018).

WUS transcription is activated by a complex of factors including B-type *ARABIDOPSIS RESPONSE REGULATOR (ARR)* and HDIII-ZIP family proteins (**Supplementary Table 1** and **Figure 1**). In separate experiments, overexpression of *ARR2* and *ARR12* led to roughly twofold increases in number of shoots from callus *in vitro* (Dai et al., 2017; Zhang et al., 2017). Contrary to the roles of B-type *ARR* genes, overexpression of the A-type led to strong suppression of *in vitro* organogenic capacity (Osakabe et al., 2002; Buechel et al., 2010).

The combined knockout of several HDIII-ZIP factors in a *wus* background rescued SAM development, indicating that these factors may simultaneously promote and inhibit stem cell differentiation via pathways both dependent and independent of *WUSCHEL* (Lee and Clark, 2015). Orthologs of these TFs found in *Populus trichocarpa* are expressed in the SAM, and overexpression promotes stem cell proliferation and inhibits development of shoot primordia into mature organs (Du et al., 2011; Robischon et al., 2011; Zhu et al., 2013).

ENHANCER OF SHOOT REGENERATION 1/2 acts directly upstream of *WUS* and indirectly upstream of *STM* (**Supplementary Table 1** and **Figure 1**) and numerous other poorly known genes (Chandler et al., 2007; Ikeuchi et al., 2018). *ESR1* overexpression conferred cytokinin-independent competence for regeneration, although constitutive overexpression inhibited differentiation of SAMs. Recovery of transgenic plants was enabled via deactivation of chemoinducible *ESR1* after shoot primordial development (Banno et al., 2001). *ESR2* is transactivated by *ESR1* and shares many downstream targets (Ikeuchi et al., 2018); overexpression of either leads to remarkable improvement in shoot regeneration capacity in Arabidopsis (Ikeda et al., 2006; Mase et al., 2007). Unlike *ESR1* (Banno et al., 2001), overexpression of the upstream gene *WOUND-INDUCED DEDIFFERENTIATION 1 (WIND1)* enhances formation of callus in addition to shoots (Iwase et al., 2016); however, callus develops into shoots upon transfer of chemoinducible *wind1* mutants to media without inducer (Iwase et al., 2011). Several WIND homologs are known to transcriptionally activate ESR1 (**Supplementary Table 1**).

PLETHORA (PLT) 3/5/7 are responsible for direct transactivation of WUS and indirect transactivation of both WUS and STM via *CUP-SHAPED COTYLEDON (CUC) 1/2* (**Supplementary Table 1** and **Figure 1**). In Arabidopsis, overexpression of *PLT5/7* enables cytokinin-independent shoot regeneration, although at a very low rate (Kareem et al., 2015).

BABY BOOM (BBM) overexpression confers the ability for cytokinin-independent *in vitro* somatic embryogenesis in Arabidopsis (Boutilier et al., 2002). In contrast, expression of a *BBM* homolog in tobacco enhanced regeneration via shoot organogenesis without affecting embryogenesis (Srinivasan et al., 2007, 2011).

Effectiveness of Developmental Genes in Non-model Species

WUS and related genes have been found to be effective at promoting regeneration in crop and forest species. Overexpression of the rice homolog of *WUS* in seedlings

was found to cause *de novo* organogenesis of shoots in rice (Kamiya et al., 2003). *Populus tomentosa* transformed to overexpress any of four *WUSCHEL* or *WUSCHEL-ASSOCIATED HOMEOBOX* orthologs showed increased adventitious rooting (Liu et al., 2014; Li et al., 2018). Overexpression of *AtWUS* led to increased embryo and callus formation *in vitro* in coffee (Arroyo-Herrera et al., 2008), increased embryo formation in cotton (Bouchabké-Coussa et al., 2013), and enables *in vitro* ectopic embryogenesis for the otherwise completely recalcitrant *Capsicum chinense* (Solís-Ramos et al., 2009). Overexpression of homologs of *WUS*, or *WUS* in combination with *BABY BOOM (BBM)*, enhances *in vitro* transformation and shoot regeneration in a variety of monocots, including rice, sorghum, and maize. Several completely recalcitrant maize inbred lines became responsive to transformation and regeneration when overexpressing either *WUS* or *BBM* homologs (Lowe et al., 2016; Mookkan et al., 2017).

STM and related genes are also active in crop and forest species. Expression of *STM/Knotted-1* orthologs from apple or maize enhanced shoot regeneration from leaf explants in the absence of exogenous cytokinins in tobacco, though were not effective in plum under the conditions studied (Srinivasan et al., 2011). In citrus, expression of maize *Knotted-1* enhanced *in vitro* regeneration after transformation, with rates varying widely among varieties (Hu et al., 2016). In the gymnosperm forest tree *Picea abies* (Norway spruce), overexpression of a *KNOTTED-1* ortholog similarly promoted *in vitro* somatic embryogenesis (Belmonte et al., 2007).

Other DGs can also stimulate regeneration in non-model species. Overexpression of maize *LEC* in wheat and maize enabled efficient transformation without the use of selectable markers (Lowe et al., 2002). *AtLEC1* overexpression in white spruce, however, had no effect on somatic embryogenesis (Klimaszewska et al., 2010). Overexpression of an *ESR1* ortholog led to a doubling of shoot regeneration during transformation in hybrid poplar (Yordanov et al., 2014). Overexpression of a *BBM* ortholog in *Capsicum annuum* (sweet pepper) enabled *in vitro* somatic embryogenesis and efficient transformation of a genotype which was previously recalcitrant (Heidmann et al., 2011). In *P. tomentosa*, expression of a *Brassica* homolog of *BBM* led calli to develop somatic embryos, which is otherwise rarely seen with poplar regeneration systems (Deng et al., 2009).

Strategies for Using Developmental Genes for Transformation Improvement

When overexpressed during vegetative development, *BBM*, *WUS*, *LEC1* and other genes can promote various regeneration pathways, but then unsurprisingly lead to defects in further development, such as disorganized shoot and floral meristems (Gallois et al., 2002), infertility, and shoot necrosis (Lowe et al., 2016). Clearly, to be useful their expression must be carefully controlled in strength and timing. Three main options exist for targeted expression of transgenes: induced expression using external stimuli such as chemical or physical inducers (e.g., heat or drought); controlled excision of the genes from the genome using similar inducer options; and use of promoters

whose innate expression pattern closely resembles that of native meristematic genes, thus will have much attenuated expression after meristems or embryos are initiated. More complex options include gene editing to engineer the promoters of native DGs to be inducible, or to add miRNA-resistance mutations into the genes' transcribed regions, thus achieving some level of derepression without affecting genomic context and other local regulatory cues (Zhang et al., 2017). Heat-shock, drought-stress, and meristematic promoters have been shown to be effective for driving Cre excision of *BBM* and *WUS* orthologs in monocots (Lowe et al., 2016, 2018; Mookkan et al., 2017), and a heat-shock promoter was used to drive Flp-*FRT* excision of *BBM* after *BBM* induced somatic embryogenesis in *P. tomentosa* (Deng et al., 2009). Clearly many options exist, but the choice of the most effective DGs or combinations thereof, as well as gene expression/removal options, need to be explored in parallel for specific crop and forest tree species.

In vitro culture and selection conditions will strongly affect the value of DGs for promoting regeneration. For example, plant hormones added to culture media can mask or help to amplify DG effects, depending on the specific taxa, tissues, and hormone species and concentrations, and duration of treatments (Hill and Schaller, 2013; Irshad et al., 2018; Kumar et al., 2018). With highly effective regeneration, use of selective agents such as antibiotics can be reduced or eliminated entirely, greatly reducing physiological stresses that retard regeneration (Mookkan et al., 2017; Lowe et al., 2018). The optimal combinations, and their impacts on transformation rate as well as chimerism in the resulting plants, can only be determined by trial and error with specific taxa, and may need to be customized to specific genotypes in highly genetically variable and recalcitrant species such as most forest trees.

Future improvements to the system are likely to include growing use of multiple DGs, in part to complement and balance their differing activities, and to make DG "reagents" effective across a wide range of taxa. For example, strong overexpression of *WUS* or upstream *ESR1* without sufficient balancing activity of the *STM* pathway in the PZ has reportedly led to failure of PZ cells to differentiate (Banno et al., 2001; Ikeda et al., 2006), and to necrosis of shoot primordia (Lowe et al., 2016). Studies of gene combinations have to date received little attention, but may lead to major improvements in transformation efficiency.

TOWARD THE DEVELOPMENT OF *IN PLANTA* TRANSFORMATION METHODS IN FOREST TREES

A major limitation to genetic transformation is the need to develop *in vitro* propagation and regeneration systems, which for many plant species are very time consuming and require a high level of technical expertise. Moreover, it must be customized for each new genotype and species, with many remaining recalcitrant to regeneration and/or transformation. As a result, methods that bypass the need for *in vitro* systems are highly desirable.

In planta transformation techniques take advantage of natural biological processes to produce and regenerate transgenic plants

(**Figure 2**), and are thus in theory applicable to a large panel of genotypes and species. The target tissues are diverse, and can include secondary meristems. The induced somatic sector analysis (ISSA) approach is an example of what can be achieved in various tree species (*Pinus*, *Eucalyptus*, *Populus*, Spokevicius et al., 2005; Van Beveren et al., 2006). A "cambial window" is cut with a sharp razor blade through the bark to get access to the cambial/young xylem tissues, which are then inoculated with an *Agrobacterium tumefaciens* solution. After wound closure and cambium reestablishment, the transformed cells divide and differentiate, producing somatic sectors of transformed cells. Within a few months, and without any *in vitro* steps or complex manipulations, it is possible to analyze transgenes and promoters directly in the woody stem tissues of trees by comparing transformed sectors with adjacent non-transformed ones (Hussey et al., 2011; Creux et al., 2013; Baldacci-Cresp et al., 2015). ISSA has a great potential to study cell fate and pattern formation during secondary growth and xylogenesis, thanks to the development of microscope-derived techniques like Raman or ATR-FTIR, which can give spectroscopic information at the cellular or cell wall level. Although very useful for research, ISSA can, however, not be used to regenerate transgenic plants.

To this end, other tissues can be targeted for transformation, including vegetative meristems protected in axillary or apical buds, as was investigated in sugarcane (Mayavan et al., 2015), *Populus* (Yang et al., 2010), and grapevine (Fujita et al., 2009). *A. tumefaciens* was the DNA vector and reached the meristems after mechanical wounding, possibly complemented by sonication or vacuum infiltration (Mayavan et al., 2015). This approach required an efficient adventitious rooting system to regenerate plants from transgenic shoots excised from mother plants. However, repeated rounds of selection generally failed to avoid chimeric transgenic shoots.

To avoid chimeras, some protocols have been developed on germinating seeds or seedlings, with the goal of reaching apical meristems as early as possible in plant development. Indeed, these meristems will ultimately give rise to reproductive meristems that may produce transformed germ cells. Several attempts were at least partly successful in producing transgenic T1 plants from T0 transformed chimeric embryos or seedlings using *A. tumefaciens* (Lin et al., 2009; Shah et al., 2015; Ahmed et al., 2018) or particle bombardment (Hamada et al., 2017). However, this method would be difficult to apply in forest trees as the elimination of chimeras usually requires sexual reproduction to the T1 generation, requiring a long wait for flowering.

In order to speed up the process, while ensuring the generation of non-chimeric plants, some authors targeted reproductive meristems before flowering and fertilization. The rationale was to transform future germ cells. Arabidopsis flowers were successfully transformed by *A. tumefaciens* vacuum infiltration (Bechtold et al., 1993); Clough and Bent (1998) further improved this method, and found that dipping of flowers was efficient enough as long as a surfactant was also used. Even though the overall efficiency was less than 1%, the method remained viable thanks to the high fertility and small size of Arabidopsis, enabling hundreds to thousands of germinating seeds to be efficiently screened using selectable markers. This method was developed

FIGURE 2 | Virtual plant showing the diverse organs/meristems targeted by *in planta* transformation. **(A)** Primary meristems in vegetative buds leading to the production of transgenic shoots after agro-inoculation. **(B)** Primary meristems in germinating seeds or seedlings leading to T0 chimeric plants and T1 transgenic plants. **(C)** Reproductive cells in floral buds or flowers leading to T0 transgenic plants. **(D)** Cambium secondary meristems leading to transformed wood sectors.

with limited success in other species (*Raphanus sativus*, Curtis and Nam, 2001; *Brassica napus*, Qing et al., 2000; *Medicago truncatula*, Trieu et al., 2000). An alternative method is to target pollen grains using methods such as sonication to penetrate pollen apertures in a DNA-containing sucrose solution. Southern analyses demonstrated the successful transfer of transgenes to progenies by pollination in sorghum (Wang et al., 2007), *Brassica juncea* (Wang et al., 2008), and maize (Yang et al., 2017). Pre-treatment of pollen grains by aeration (e.g., 20 min at 4°C) increased pollen viability, mitigating a common adverse effect of sonication (Yang et al., 2017). Zhao et al. (2017) have recently investigated an innovative method by which DNA was delivered to pollen via magnetic nanoparticles; stable transformants in cotton, pepper, and pumpkin were generated. Pollen-mediated transformation could be easily tested for forest trees as pollination for controlled crosses or seed production is very common as part of conventional breeding. However, to be feasible and conform to most regulatory requirements for containment of transgenic pollen, such crosses would need to be carried out on detached or grafted floral branches in greenhouses, which is possible only for a limited number of forest tree species.

Unfortunately, most of the procedures described above required a high level of technical expertise, as they were hardly transferred to other laboratories. The only exception is Arabidopsis floral dip, which has been used in numerous laboratories worldwide. Could *in planta* techniques such as this be tailored for forest trees? Several convergent studies have shown that the ovules, not the pollen, are the direct targets for transformation through floral dip. It has been shown that manual outcrossing experiments produced transgenic progenies only when *A. tumefaciens* was applied on pollen recipient plants, not on pollen donor plants (Ye et al., 1999; Desfeux et al., 2000). Bechtold et al. (2000) elegantly reached the same conclusion using a genetic approach.

Floral dip has been shown to give rise to transgenic seedlings with genetically independent insertions that are typically hemizygous (carrying the T-DNA at only one allele of a given locus) (Bechtold et al., 1993). The transformation rate is very dependent on the flower developmental stage (Desfeux et al., 2000), with the optimal stage being when the gynoecium is still open (Irepan Reyes-Olalde et al., 2013), thus allowing agrobacteria to penetrate and transform the ovule primordia.

The CRABS-CLAW mutant, which maintains an open gynoecium, gives a sixfold enhanced rate of transformation (Desfeux et al., 2000). Therefore, access by Agrobacterium to the locule of the ovule appears to be critical for the transformation. Trees are perennial species, where flower initiation takes place the year before flowering, thus injection of the Agrobacterium solution into female floral buds before flowering at a stage where the gynoecium is still open needs to occur weeks to months before seed release. However, trees also produce very large number of seeds, thus the *in planta* approach, if it can be optimized and applied to many buds at the right times, together with an efficient selection system for germinating seeds, may be realistic option for some tree species.

CONCLUSION

In summary, both DG and *in planta* approaches to transformation hold promise to solve major problems in plant and tree transformation. DGs appear to hold most promise where a basic *in vitro* regeneration system is in place, thus might benefit from a large elevation of transformation efficiency using an established transformation pathway. It has to date been most effective in species with embryogenic rather than organogenic regeneration systems. *In planta* systems hold most promise where *in vitro* approaches are extremely difficult or impossible, thus alternate pathways are required. It will also be most easily pursued as part of a large-scale breeding program, enabling large numbers of floral buds to be treated and seedling populations screened for transformation and chimerism. *In planta* transformation of cambium and axillary/apical buds has been successful, but is prone to chimerism; research on

germline transformation (i.e., transformation of mother cells within floral buds and embryos within seeds) may help to reduce this problem. *In planta* and DG overexpression approaches to efficient transformation might not be mutually exclusive; research is warranted to elucidate the potential for DGs to enhance *in planta* systems. Given the importance of regeneration as a bottleneck to transformation and gene editing of forest trees in research and application (Chang et al., 2018), acceleration of research using both approaches is warranted.

AUTHOR CONTRIBUTIONS

MN and AD drafted the sections "Developmental Genes as Tools for Transformation Improvement" and "Toward the Development of *in planta* Transformation Methods in Forest Trees," respectively, after extensive consultation with SS and GP. SS and GP revised the drafted text and prepared initial drafts of the sections "Introduction" and "Conclusion."

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2018.01443/full#supplementary-material

REFERENCES

Ahmed, T., Biswas, S., Elias, S. M., Rahman, M. S., Tuteja, N., and Seraj, Z. I. (2018). *In Planta* transformation for conferring salt tolerance to a tissue-culture unresponsive indica rice (*Oryza sativa* L.) cultivar. *In Vitro Cell. Dev. Biol. Plant* 54, 154–165. doi: 10.1007/s11627-017-9870-1

Arroyo-Herrera, A., Gonzalez, A. K., Moo, R. C., Quiroz-Figueroa, F. R., Loyola-Vargas, V. M., Rodriguez-Zapata, L. C., et al. (2008). Expression of WUSCHEL in *Coffea canephora* causes ectopic morphogenesis and increases somatic embryogenesis. *Plant Cell Tiss Organ Cult*. 94, 171–180. doi: 10.1007/s11240-008-9401-1

Baldacci-Cresp, F., Moussawi, J., Leplé, J.-C., Van Acker, R., Kohler, A., Candiracci, J., et al. (2015). PtaRHE1, a *Populus tremula* x *Populus alba* RING-H2 protein of the ATL family, has a regulatory role in secondary phloem fibre development. *Plant J*. 82, 978–990. doi: 10.1111/tpj.12867

Balkunde, R., Kitagawa, M., Xu, X. M., Wang, J., and Jackson, D. (2017). SHOOT MERISTEMLESS trafficking controls axillary meristem formation, meristem size and organ boundaries in Arabidopsis. *Plant J*. 90, 435–446. doi: 10.1111/tpj.13504

Banno, H., Ikeda, Y., Niu, Q.-W., and Chua, N.-H. (2001). Overexpression of Arabidopsis *ESR1* induces initiation of shoot regeneration. *Plant Cell* 13, 2609–2618. doi: 10.1105/tpc.010234

Bechtold, N., Ellis, J., and Pelletier, G. (1993). *In Planta Agrobacterium*-mediated gene-transfer by infiltration of adult *Arabidopsis thaliana* plants. *C. R. Acad. Sci. Ser. III Sci*. 316, 1194–1199. doi: 10.1007/978-3-642-79247-2_3

Bechtold, N., Jaudeau, B., Jolivet, S., Maba, B., Vezon, D., Voisin, R., et al. (2000). The maternal chromosome set is the target of the T-DNA in the *in planta* transformation of *Arabidopsis thaliana*. *Genetics* 155, 1875–1887.

Belmonte, M. F., Tahir, M., Schroeder, D., and Stasolla, C. (2007). Overexpression of HBK3, a class I KNOX homeobox gene, improves the development of Norway spruce (*Picea abies*) somatic embryos. *J. Exp. Bot*. 58, 2851–2861. doi: 10.1093/jxb/erm099

Bouchabké-Coussa, O., Obellianne, M., Linderme, D., Montes, E., Maia-Grondard, A., Vilaine, F., et al. (2013). Wuschel overexpression promotes somatic embryogenesis and induces organogenesis in cotton *Gossypium hirsutum* tissues cultured *in vitro*. *Plant Cell Rep*. 32, 675–686. doi: 10.1007/s00299-013-1402-9

Boutilier, K., Offringa, R., Sharma, V. K., Kieft, H., Ouellet, T., Zhang, L., et al. (2002). Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. *Plant Cell* 14, 1737–1749. doi: 10.1105/tpc.001941

Bryan, A. A., and Sass, J. E. (1941). Knotted" maize plants. *J. Hered*. 32, 343–346.

Buechel, S., Leibfried, A., To, J. P. C., Zhao, Z., Andersen, S. U., Kieber, J. J., et al. (2010). Role of A-type *ARABIDOPSIS RESPONSE REGULATORS* in meristem maintenance and regeneration. *Eur. J. Cell Biol*. 89, 279–284. doi: 10.1016/j.ejcb.2009.11.016

Chandler, J. W., Cole, M., Flier, A., Grewe, B., and Werr, W. (2007). The AP2 transcription factors DORNRÖSCHEN and DORNRÖSCHEN-LIKE redundantly control *Arabidopsis* embryo patterning via interaction with PHAVOLUTA. *Plant Mol. Biol*. 1662, 1653–1662. doi: 10.1007/s11103-010-9721-5

Chang, S., Mahon, E. L., MacKay, H. A., Rottmann, W. H., Strauss, S. H., Pijut, P. M., et al. (2018). Genetic engineering of trees: progress and new horizons. *In Vitro Cell. Dev. Biol*. 54, 341–376.

Clough, S. J., and Bent, A. F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J*. 16, 735–743. doi: 10.1046/j.1365-313x.1998.00343.x

Creux, N. M., Bossinger, G., Myburg, A. A., and Spokevicius, A. V. (2013). Induced somatic sector analysis of cellulose synthase (CesA) promoter regions in woody stem tissues. *Planta* 237, 799–812. doi: 10.1007/s00425-012-1792-x

Curtis, I. S., and Nam, H. G. (2001). Transgenic radish (*Raphanus sativus* L. *longipinnatus* Bailey) by floral-dip method - plant development and surfactant are important in optimizing transformation efficiency. *Transgenic Res.* 10, 363–371. doi: 10.1023/A:1016600517293

Dai, X., Liu, Z., Qiao, M., Li, J., Li, S., and Xiang, F. (2017). ARR12 promotes *de novo* shoot regeneration in *Arabidopsis thaliana* via activation of *WUSCHEL* expression. *J. Integr. Plant Biol.* 59, 747–758. doi: 10.1111/jipb.12567

Daimon, Y., Takabe, K., and Tasaka, M. (2003). The CUP-SHAPED COTYLEDON genes promote adventitious shoot formation on calli. *Plant Cell Physiol.* 44, 113–121. doi: 10.1093/pcp/pcg038

Daum, G., Medzihradszky, A., Suzaki, T., and Lohmann, J. U. (2014). A mechanistic framework for noncell autonomous stem cell induction in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 111, 14619–14624. doi: 10.1073/pnas. 1406446111

Deng, W., Luo, K., Li, Z., and Yang, Y. (2009). A novel method for induction of plant regeneration via somatic embryogenesis. *Plant Sci.* 177, 43–48. doi: 10.1016/j.plantsci.2009.03.009

Desfeux, C., Clough, S. J., and Bent, A. F. (2000). Female reproductive tissues are the primary target of *Agrobacterium*-mediated transformation by the Arabidopsis floral-dip method. *Plant Physiol.* 123, 895–904. doi: 10.1104/pp. 123.3.895

Du, J., Miura, E., Robischon, M., Martinez, C., and Groover, A. (2011). The *Populus* class III HD ZIP transcription factor *POPCORONA* affects cell differentiation during secondary growth of woody stems. *PLoS One* 6:e17458. doi: 10.1371/ journal.pone.0017458

Fujita, K., Matsuoka, T., Suzuki, S., and Takayanagi, T. (2009). *In planta* transformation technique for grapevines (*Vitis vinifera* L) using dormant buds. *J. Plant Biochem. Biotechnol.* 18, 161–167. doi: 10.1007/BF03263314

Gallois, J.-L., Woodward, C., Reddy, G. V., and Sablowski, R. (2002). Combined SHOOT MERISTEMLESS and WUSCHEL trigger ectopic organogenesis in *Arabidopsis*. *Development* 129, 3207–3217. doi: 10.1007/s00425-009-0988-1

Gordon, S. P., Heisler, M. G., Reddy, G. V., Ohno, C., Das, P., and Meyerowitz, E. M. (2007). Pattern formation during de novo assembly of the *Arabidopsis* shoot meristem. *Development* 134, 3539–3548. doi: 10.1242/dev.01 0298

Hake, S., Vollbrecht, E., and Freeling, M. (1989). Cloning Knotted, the dominant morphological mutant in maize using Ds2 as a transposon tag. *EMBO J.* 8, 15–22. doi: 10.1002/j.1460-2075.1989.tb03343.x

Hamada, H., Linghu, Q., Nagira, Y., Miki, R., Taoka, N., and Imai, R. (2017). An *in planta* biolistic method for stable wheat transformation. *Sci. Rep.* 7:11443. doi: 10.1038/s41598-017-11936-0

Heidmann, I., de Lange, B., Lambalk, J., Angenent, G. C., and Boutilier, K. (2011). Efficient sweet pepper transformation mediated by the BABY BOOM transcription factor. *Plant Cell Rep.* 30, 1107–1115. doi: 10.1007/s00299-011-1018-x

Hibara, K., Takada, S., and Tasaka, M. (2003). CUC1 gene activates the expression of SAM-related genes to induce adventitious shoot formation. *Plant J.* 36, 687–696. doi: 10.1046/j.1365-313X.2003.01911.x

Hill, K., and Schaller, G. E. (2013). Enhancing plant regeneration in tissue culture. *Plant Signal. Behav.* 8:e25709. doi: 10.4161/psb.25709

Horstman, A., Willemsen, V., Boutilier, K., and Heidstra, R. (2014). AINTEGUMENTA-LIKE proteins: hubs in a plethora of networks. *Trends Plant Sci.* 19, 146–157. doi: 10.1104/pp.17.00232

Hu, W., Li, W., Xie, S., Fagundez, S., McAvoy, R., Deng, Z., et al. (2016). Kn1 gene overexpression drastically improves genetic transformation efficiencies of citrus cultivars. *Plant Cell Tissue Organ Cult.* 125, 81–91. doi: 10.1007/s11240-015-0931-z

Hussey, S. G., Mizrachi, E., Spokevicius, A. V., Bossinger, G., Berger, D. K., and Myburg, A. A. (2011). *SND2*, a NAC transcription factor gene, regulates genes involved in secondary cell wall development in *Arabidopsis* fibres and increases fibre cell area in *Eucalyptus*. *BMC Plant Biol.* 11:173. doi: 10.1186/1471-2229-11-173

Ikeda, Y., Banno, H., Niu, Q. W., Howell, S. H., and Chua, N. H. (2006). The *ENHANCER* of SHOOT REGENERATION 2 gene in *Arabidopsis* regulates

CUP-SHAPED COTYLEDON 1 at the transcriptional level and controls cotyledon development. *Plant Cell Physiol.* 47, 1443–1456. doi: 10.1093/pcp/ pcl023

Ikeuchi, M., Shibata, M., Rymen, B., Iwase, A., Baagman, A. M., Watt, L., et al. (2018). A gene regulatory network for cellular reprogramming in plant regeneration. *Plant Cell Physiol.* 59, 770–782. doi: 10.1093/pcp/pcy013

Irepan Reyes-Olalde, J., Zuniga-Mayo, V. M., Chavez Montes, R. A., Marsch-Martinez, N., and de Folter, S. (2013). Inside the gynoecium: at the carpel margin. *Trends Plant Sci.* 18, 644–655. doi: 10.1016/j.tplants.2013. 08.002

Irshad, M., Rizwan, H. M., Debnath, B., Anwar, M., Li, M., Liu, S., et al. (2018). Ascorbic acid controls lethal browning and Pluronic F-68 promotes high-frequency multiple shoot regeneration from cotyldonary node explant of Okra (*Abelmoschus esculentus* L.). *HortScience* 53, 183–190. doi: 10.21273/ HORTSCI12315-17

Iwase, A., Harashima, H., Ikeuchi, M., Rymen, B., Ohnuma, M., Komaki, S., et al. (2016). WIND1 promotes shoot regeneration through transcriptional activation of *ENHANCER OF SHOOT REGENERATION1* in Arabidopsis. *Plant Cell* 29, 54–69. doi: 10.1105/tpc.16.00623

Iwase, A., Mitsuda, N., Koyama, T., Hiratsu, K., Kojima, M., Arai, T., et al. (2011). The AP2/ERF transcription factor WIND1 controls cell dedifferentiation in *Arabidopsis*. *Curr. Biol.* 21, 508–514. doi: 10.1016/j.cub.2011.02.020

Kamiya, N., Nagasaki, H., Morikami, A., Sato, Y., and Matsuoka, M. (2003). Isolation and characterization of a rice WUSCHEL-type homeobox gene that is specifically expressed in the central cells of a quiescent center in the root apical meristem. *Plant J.* 35, 429–441. doi: 10.1046/j.1365-313X.2003.01 816.x

Kareem, A., Durgaprasad, K., Sugimoto, K., Du, Y., Pulianmackal, A. J., Trivedi, Z. B., et al. (2015). PLETHORA genes control regeneration by a two-step mechanism. *Curr. Biol.* 25, 1017–1030. doi: 10.1016/j.cub.2015.02.022

Klimaszewska, K., Pelletier, G., Overton, C., Stewart, D., and Rutledge, R. G. (2010). Hormonally regulated overexpression of *Arabidopsis* WUS and conifer LEC1 (CHAP3A) in transgenic white spruce: implications for somatic embryo development and somatic seedling growth. *Plant Cell Rep.* 29, 723–734. doi: 10.1007/s00299-010-0859-z

Kumar, N., Bhatt, V. D., Mastan, S. G., and Reddy, M. P. (2018). "TDZ-induced plant regeneration in *Jatropha curcas*: a promising biofuel plant," in *Thidiazuron: from Urea Derivative to Plant Growth Regulator*, eds N. Ahmad and M. Faisal (Singapore: Springer), 419–428. doi: 10.1007/978-981-10-8004-3_23

Laux, T., Mayer, K. F., Berger, J., and Jürgens, G. (1996). The WUSCHEL gene is required for shoot and floral meristem integrity in *Arabidopsis*. *Development* 122, 87–96. doi: 10.1105/tpc.1.1.37

Lee, C., and Clark, S. E. (2015). A WUSCHEL-independent stem cell specification pathway is repressed by PHB, PHV and CNA in Arabidopsis. *PLoS One* 10:e0126006. doi: 10.1371/journal.pone.0126006

Li, J., Zhang, J., Jia, H., Liu, B., Sun, P., Hu, J., et al. (2018). The WUSCHEL-related homeobox 5a (PtoWOX5a) is involved in adventitious root development in poplar. *Tree Physiol.* 38, 139–153. doi: 10.1093/treephys/tpx118

Lin, J., Zhou, B., Yang, Y., Mei, J., Zhao, X., Guo, X., et al. (2009). Piercing and vacuum infiltration of the mature embryo: a simplified method for *Agrobacterium*-mediated transformation of *indica* rice. *Plant Cell Rep.* 28, 1065–1074. doi: 10.1007/s00299-009-0706-2

Lincoln, C., Long, J., Yamaguchi, J., Serikawa, K., and Hake, S. (1994). A knotted1-like homeobox gene in Arabidopsis is expressed in the vegetative meristem and dramatically alters leaf morphology when overexpressed in transgenic plants. *Plant Cell* 6, 1859–1876. doi: 10.1105/tpc.6.12.1859

Liu, B., Wang, L., Zhang, J., Li, J., Zheng, H., Chen, J., et al. (2014). WUSCHEL-related Homeobox genes in *Populus tomentosa*: diversified expression patterns and a functional similarity in adventitious root formation. *BMC Genomics* 15:296. doi: 10.1186/1471-2164-15-296

Long, J. A., Moan, E. I., Medford, J. I., and Barton, M. K. (1996). A member of the KNOTTED class of homeodomain proteins encoded by the STM gene of *Arabidopsis*. *Nature* 379, 66–69. doi: 10.1038/379066a0

Lotan, T., Ohto, M., Yee, K. M., West, M. A., Lo, R., Kwong, R. W., et al. (1998). *Arabidopsis* LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. *Cell* 93, 1195–1205. doi: 10.1016/S0092-8674(00)81463-4

Lowe, K., Hoerster, G., Sun, X., Rasco-Gaunt, S., Lazerri, P., Ellis, S., et al. (2002). "Maize LEC1 improves transformation in both maize and wheat," in *Plant Biotechnology 2002 and Beyond*, ed. I. K. Vasil (Orlando, FL: Springer), 283–284.

Lowe, K., Rota, M. L., Hoerster, G., Hastings, C., Wang, N., Chamberlin, M., et al. (2018). Rapid genotype "independent" *Zea mays* L. (maize) transformation via direct somatic embryogenesis. *In Vitro Cell. Dev. Biol. Plant* 54, 240–252. doi: 10.1007/s11627-018-9905-2

Lowe, K., Wu, E., Wang, N., Hoerster, G., Hastings, C., Cho, M.-J., et al. (2016). Morphogenic regulators *Baby boom* and *Wuschel* improve Monocot transformation. *Plant Cell* 28, 1998–2015. doi: 10.1105/TPC.16.00124

Mase, H., Hashiba, M., Matsuo, N., and Banno, H. (2007). Expression patterns of *Arabidopsis* ERF VIII-b subgroup genes during *in vitro* shoot regeneration and effects of their overexpression on shoot regeneration efficiency. *Plant Biotechnol.* 24, 481–486. doi: 10.5511/plantbiotechnology.24.481

Mayavan, S., Subramanyam, K., Jaganath, B., Sathish, D., Manickavasagam, M., and Ganapathi, A. (2015). *Agrobacterium*-mediated *in planta* genetic transformation of sugarcane setts. *Plant Cell Rep.* 34, 1835–1848. doi: 10.1007/s00299-015-1831-8

Mayer, K. F. X., Schoof, H., Haecker, A., Lenhard, M., Jürgens, G., and Laux, T. (1998). Role of *WUSCHEL* in regulating stem cell fate in the *Arabidopsis* shoot meristem. *Cell* 95, 805–815. doi: 10.1016/S0092-8674(00)81703-1

Mookkan, M., Nelson-Vasilchik, K., Hague, J., Zhang, Z. J., and Kausch, A. P. (2017). Selectable marker independent transformation of recalcitrant maize inbred B73 and sorghum P898012 mediated by morphogenic regulators BABY BOOM and WUSCHEL2. *Plant Cell Rep.* 36, 1477–1491. doi: 10.1007/s00299-017-2169-1

Osakabe, Y., Miyata, S., Urao, T., Seki, M., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2002). Overexpression of *Arabidopsis* response regulators, ARR4/ATRR1/IBC7 and ARR8/ATRR3, alters cytokinin responses differentially in the shoot and in callus formation. *Biochem. Biophys. Res. Commun.* 293, 806–815. doi: 10.1016/S0006-291X(02)00286-3

Qing, C. M., Fan, L., Lei, Y., Bouchez, D., Tourneur, C., Yan, L., et al. (2000). Transformation of Pakchoi (*Brassica rapa* L. ssp *chinensis*) by *Agrobacterium* infiltration. *Mol. Breed.* 6, 67–72. doi: 10.1023/A:1009658128964

Robischon, M., Du, J., Miura, E., and Groover, A. (2011). The *Populus* Class III HD ZIP, *popREVOLUTA*, influences cambium initiation and patterning of woody stems. *Plant Physiol.* 155, 1214–1225. doi: 10.1104/pp.110.167007

Scofield, S., Murison, A., Jones, A., Fozard, J., Aida, M., Band, L. R., et al. (2018). Coordination of meristem and boundary functions by transcription factors in the SHOOT MERISTEMLESS regulatory network. *Development* 145:dev157081. doi: 10.1242/dev.157081

Shah, S. H., Ali, S., Jan, S. A., Jalal-Ud-Din, and Ali, G. M. (2015). Piercing and incubation method of *in planta* transformation producing stable transgenic plants by overexpressing DREB1A gene in tomato (*Solanum lycopersicum* Mill. *Plant Cell Tissue Organ Cult.* 120, 1139–1157. doi: 10.1007/s11240-014-0670-6

Sinha, N. R., Williams, R. E., and Hake, S. (1993). Overexpression of the maize homeo box gene, KNOTTED-1, causes a switch from determinate to indeterminate cell fates. *Genes Dev.* 7, 787–795. doi: 10.1101/gad.7.5.787

Smith, L. G., Greene, B., Veit, B., and Hake, S. (1992). A dominant mutation in the maize homeobox gene, Knotted-1, causes its ectopic expression in leaf cells with altered fates. *Development* 116, 21–30.

Solís-Ramos, L. Y., González-Estrada, T., Nahuath-Dzib, S., Zapata-Rodriguez, L. C., and Castaño, E. (2009). Overexpression of *WUSCHEL* in *C. chinense* causes ectopic morphogenesis. *Plant Cell Tissue Organ Cult.* 96, 279–287. doi: 10.1007/s11240-008-9485-7

Spinelli, S. V., Martin, A. P., Viola, I. L., Gonzalez, D. H., and Palatnik, J. F. (2011). A mechanistic link between *STM* and *CUC1* during Arabidopsis development. *Plant Physiol.* 156, 1894–1904. doi: 10.1104/pp.111.177709

Spokevicius, A. V., Van Beveren, K., Leitch, M. M., and Bossinger, G. (2005). *Agrobacterium*-mediated *in vitro* transformation of wood-producing stem segments in eucalypts. *Plant Cell Rep.* 23, 617–624. doi: 10.1007/s00299-004-0856-1

Srinivasan, C., Liu, Z., Heidmann, I., Supena, E. D. J., Fukuoka, H., Joosen, R., et al. (2007). Heterologous expression of the *BABY BOOM* AP2/ERF transcription factor enhances the regeneration capacity of tobacco *Nicotiana tabacum* L. *Planta* 225:341. doi: 10.1007/s00425-006-0358-1

Srinivasan, C., Liu, Z., and Scorza, R. (2011). Ectopic expression of class 1 KNOX genes induce adventitious shoot regeneration and alter growth and development of tobacco (*Nicotiana tabacum* L) and European plum (*Prunus domestica* L). *Plant Cell Rep.* 30, 655–664. doi: 10.1007/s00425-006-0358-1

Takada, S., Hibara, K., Ishida, T., and Tasaka, M. (2001). The CUP-SHAPED COTYLEDON1 gene of Arabidopsis regulates shoot apical meristem formation. *Development* 128, 1127–1135.

Trieu, A. T., Burleigh, S. H., Kardailsky, I. V., Maldonado-Mendoza, I. E., Versaw, W. K., Blaylock, L. A., et al. (2000). Transformation of *Medicago truncatula* via infiltration of seedlings or flowering plants with *Agrobacterium*. *Plant J.* 22, 531–541. doi: 10.1046/j.1365-313x.2000.00757.x

Van Beveren, K. S., Spokevicius, A. V., Tibbits, J., Wang, Q., and Bossinger, G. (2006). Transformation of cambial tissue *in vivo* provides an efficient means for induced somatic sector analysis and gene testing in stems of woody plant species. *Funct. Plant Biol.* 33, 629–638. doi: 10.1071/FP06057

Wang, J., Li, Y., and Liang, C. (2008). Recovery of transgenic plants by pollen-mediated transformation in *Brassica juncea*. *Transgenic Res.* 17, 417–424. doi: 10.1007/s11248-007-9115-x

Wang, W., Wang, J., Yang, C., Li, Y., Liu, L., and Xu, J. (2007). Pollen-mediated transformation of *Sorghum bicolor* plants. *Biotechnol. Appl. Biochem.* 48, 79–83. doi: 10.1042/BA20060231

Williams, L., Grigg, S. P., Xie, M., Christensen, S., and Fletcher, J. C. (2005). Regulation of *Arabidopsis* shoot apical meristem and lateral organ formation by microRNA *miR166g* and its *AtHD-ZIP* target genes. *Development* 132, 3657–3668. doi: 10.1242/dev.01942

Xie, M., Chen, H., Huang, L., O'Neil, R. C., Shokhirev, M. N., and Ecker, J. R. (2018). A B-ARR-mediated cytokinin transcriptional network directs hormone cross-regulation and shoot development. *Nat. Commun.* 9:1604. doi: 10.1038/s41467-018-03921-6

Yadav, R. K., Tavakkoli, M., and Reddy, G. V. (2010). WUSCHEL mediates stem cell homeostasis by regulating stem cell number and patterns of cell division and differentiation of stem cell progenitors. *Development* 137, 3581–3589. doi: 10.1242/dev.054973

Yang, L., Cui, G., Wang, Y., Hao, Y., Du, J., Zhang, H., et al. (2017). Expression of foreign genes demonstrates the effectiveness of pollen-mediated transformation in *Zea mays*. *Front. Plant Sci.* 8:383. doi: 10.3389/fpls.2017.00383

Yang, L., Sun, Y., Xie, L., and Liang, A. (2010). A novel approach for *in situ* bud transformation of *Populus* by *Agrobacterium*. *Scand. J. For. Res.* 25, 3–9. doi: 10.1080/02827580903228870

Ye, G. N., Stone, D., Pang, S. Z., Creely, W., Gonzalez, K., and Hinchee, M. (1999). Arabidopsis ovule is the target for *Agrobacterium in planta* vacuum infiltration transformation. *Plant J.* 19, 249–257. doi: 10.1046/j.1365-313X.1999.00520.x

Yordanov, Y. S., Ma, C., Strauss, S. H., and Busov, V. B. (2014). EARLY BUD-BREAK 1 EBB1 is a regulator of release from seasonal dormancy in poplar trees. *Proc. Natl. Acad. Sci. U.S.A.* 111, 10001–10006. doi: 10.1073/pnas.1405621111

Zhang, T.-Q., Lian, H., Zhou, C.-M., Xu, L., Jiao, Y., and Wang, J.-W. (2017). A two-step model for *de novo* activation of WUSCHEL during plant shoot regeneration. *Plant Cell* 29, 1073–1087. doi: 10.1105/tpc.16.00863

Zhao, X., Meng, Z., Wang, Y., Chen, W., Sun, C., Cui, B., et al. (2017). Pollen magnetofection for genetic modification with magnetic nanoparticles as gene carriers. *Nat. Plants* 3, 956–964. doi: 10.1038/s41477-017-0063-z

Zhu, Y., Song, D., Sun, J., Wang, X., and Li, L. (2013). *PtrHB7*, a class III HD-Zip gene, plays a critical role in regulation of vascular cambium differentiation in *Populus*. *Mol. Plant* 6, 1331–1343. doi: 10.1093/mp/sss164

Zuo, J., Niu, Q.-W., Frugis, G., and Chua, N.-H. (2002). The *WUSCHEL* gene promotes vegetative-to-embryonic transition in *Arabidopsis*. *Plant J.* 30, 349–359. doi: 10.1016/S0958-1669(02)00301-4

Genome Editing in Trees: From Multiple Repair Pathways to Long-Term Stability

William Patrick Bewg[1], Dong Ci[1,2] and Chung-Jui Tsai[1]*

[1] Warnell School of Forestry and Natural Resources, Department of Genetics, and Department of Plant Biology, University of Georgia, Athens, GA, United States, [2] Department of Bioscience and Biotechnology, Beijing Forestry University, Beijing, China

The CRISPR technology continues to diversify with a broadening array of applications that touch all kingdoms of life. The simplicity, versatility and species-independent nature of the CRISPR system offers researchers a previously unattainable level of precision and control over genomic modifications. Successful applications in forest, fruit and nut trees have demonstrated the efficacy of CRISPR technology at generating null mutations in the first generation. This eliminates the lengthy process of multigenerational crosses to obtain homozygous knockouts (KO). The high degree of genome heterozygosity in outcrossing trees is both a challenge and an opportunity for genome editing: a challenge because sequence polymorphisms at the target site can render CRISPR editing ineffective; yet an opportunity because the power and specificity of CRISPR can be harnessed for allele-specific editing. Examination of CRISPR/Cas9-induced mutational profiles from published tree studies reveals the potential involvement of multiple DNA repair pathways, suggesting that the influence of sequence context at or near the target sites can define mutagenesis outcomes. For commercial production of elite trees that rely on vegetative propagation, available data suggest an excellent outlook for stable CRISPR-induced mutations and associated phenotypes over multiple clonal generations.

Keywords: mutagenesis, genome engineering, allele dose effect, *Populus*, biallelic, monoallelic, knockout

INTRODUCTION

CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)-based genome editing is rapidly becoming the system of choice for targeted mutagenesis in a growing variety of woody species, including forest trees. Forest trees are an invaluable commodity, providing fiber, energy, materials and climate buffering to the global community, and CRISPR has the potential to further enhance these important traits. Previous-generation methods for gene silencing in plants rely on expression of antisense RNAs, small interfering RNAs or microRNAs to base-pair with target mRNAs for degradation, often with unpredictable and unstable outcomes (Alessandra and Shihshieh, 2010). The specificity and efficiency of CRISPR for targeted DNA mutations, and the ease of adoption in virtually any species are behind the current revolution in genomic editing (Jiang and Doudna, 2017). Meanwhile, CRISPR's popularity is driving the discovery and characterization of new CRISPR-associated (Cas) endonucleases with novel properties that make the system even more versatile (Burstein et al., 2016; Murovec et al., 2017). This review will focus on recent applications

of CRISPR in woody species, with a special focus on forest trees, the mutation patterns observed at target sites, and the long-term stability of CRISPR/Cas9-edited outcomes.

CRISPR APPLICATIONS IN WOODY SPECIES

Phytoene desaturase (PDS) has been a popular marker for evaluating CRISPR in new study systems (**Table 1**). Its mutation disrupts chlorophyll biosynthesis, allowing for visual assessment of knockout (KO) efficiency. CRISPR/Cas9-induced albino mutants have been reported in poplar (Fan et al., 2015), citrus (Jia and Wang, 2014; Zhang et al., 2017), apple (Nishitani et al., 2016), grape (Nakajima et al., 2017), cassava (Odipio et al., 2017), coffee (Breitler et al., 2018), and kiwifruit (Wang Z. et al., 2018). Successful implementation of CRISPR has also been demonstrated by targeting potential developmental and biosynthesis pathway genes in grape (Ren et al., 2016) and the tropical tree *Parasponia andersonii* (van Zeijl et al., 2018; **Table 1**). New CRISPR reagents have been developed to expand genome editing capabilities. One such reagent, SaCas9 from *Staphylococcus aureus*, was shown to effectively generate mutations in Duncan grapefruit (Jia et al., 2017a). Compared to the most commonly used SpCas9 from *Streptococcus pyogenes*, SaCas9 is considerably smaller and recognizes a distinct 5′-NNGRRT protospacer adjacent motif (PAM) sequence (versus 5′-NGG of SpCas9). Using alternative CRISPR/Cas systems such as SaCas9 can increase the number of potential guide-RNA (gRNA) target sites, especially in AT-rich regions which may facilitate promoter editing.

Besides proof-of-concept studies, the CRISPR/Cas9 system has been used to develop disease resistant fruit trees with promising results (**Table 1**). The devastating citrus canker disease is caused by *Xanthomonas citri* subsp. *citri* (Xcc) through effector-activation of a canker susceptibility gene *LOB1* of the Lateral Organ Boundaries transcription factor family (Hu et al., 2014). When the *LOB1* promoter was targeted by CRISPR/Cas9 to disrupt the effector-binding element, canker symptoms after Xcc infection were reduced in Duncan grapefruit (Jia et al., 2016) and Wanjincheng orange (Peng et al., 2017). CRISPR-KO of *LOB1* also increases Xcc resistance in Duncan grapefruit (Jia et al., 2017b). KO-mutations in other susceptibility genes for powdery mildew and fire blight disease have also been achieved in grape and apple protoplasts, respectively (Malnoy et al., 2016), potentially allowing for the regeneration of disease-resistant plants. Several WRKY transcription factors involved in defense regulation have also been targeted for mutagenesis. CRISPR-KO of two positive regulators *PtrWRKY18* and *PtrWRKY35* compromised resistance to *Melampsora* rust in *Populus* (Jiang et al., 2017), whereas KO of grape *VvWRKY52* increased resistance to necrotrophic *Botrytis cinerea* (Wang X. et al., 2018).

To date, the greatest progress in woody species has been made with poplar, the first stably transformed tree to be genome-edited by CRISPR with high efficiency (Zhou et al., 2015). Allele-sensitive bioinformatics resources to facilitate genome editing

in heterozygous species quickly followed, again based on the poplar system (Xue and Tsai, 2015; Xue et al., 2015). The majority of CRISPR studies in poplar have targeted phenylpropanoid metabolism or cell wall traits (**Table 1**). Mutations of individual 4-coumarate:CoA ligase (*4CL*) genes decreased the levels of structural (lignin) or non-structural (proanthocyanidin) phenylpropanoid polymers. CRISPR-KO of MYB transcription factors either increased (*PtoMYB156* and *PtrMYB57*) or decreased (*PtoMYB115* and *PtoMYB170*) phenylpropanoid flux, affecting in turn lignin deposition (*PtoMYB156* and *PtoMYB170*) or flavonoid accrual (*PtrMYB57* and *PtoMYB115*), respectively (Wan et al., 2017; Wang et al., 2017; Xu et al., 2017; Yang et al., 2017). Secondary cell wall synthesis was also compromised by CRISPR-KO of a brassinosteroid biosynthetic gene, supporting a role for brassinosteroids in wood formation (Shen et al., 2018). CRISPR-KO of *BRANCHED1-1* (*BRC1-1*) and *BRC1-2* belonging to the TCP family of transcription factors resulted in altered shoot architecture, and revealed an additional role of BRC2 in leaf development not previously reported for its *Arabidopsis* ortholog (Muhr et al., 2018). A recent study reported successful mutation of essential flowering genes in both male and female poplar genotypes (Elorriaga et al., 2018). The study also collated a large mutation dataset from over 500 transgenic events (Elorriaga et al., 2018) which should prove of value to understanding CRISPR/Cas editing patterns (see below). Although phenotypic evaluation of the flowering traits will require follow-on and multiyear studies in the field, the work underscores a powerful social application of CRISPR in containment of transgenic trees.

DIVERSE INDEL PROFILES INDICATIVE OF cNHEJ, MMEJ, AND TMEJ ACTIVITIES

Small frameshift indels are the most common repair outcomes of single gRNA-directed Cas9 cleavage in trees, with 1 bp insertions (+1), especially +T and +A, predominant in many cases, similar to findings from other plants and animals (Bortesi et al., 2016). However, considerable variations and case-dependent repair outcomes are also noted, suggesting potential influences of target site sequences and/or their genomic contexts (Jacobs et al., 2015; Xu et al., 2015). Meta-analysis of mutation patterns across published tree studies is necessary to gain further insight, but that is made difficult by different reporting formats (not all studies report multi-allele data), and by the use of detection methods that differ in their sensitivity, accuracy, and allele discrimination (Sentmanat et al., 2018). We combined amplicon sequencing data from CRISPR-edited *P. tremula* x *alba* IRNA 717-1B4 (717) generated in our lab (Zhou et al., 2015) with the large 717 dataset from Elorriaga et al. (2018), along with manual inspection of other published tree studies for mutation profile analysis (**Figure 1**). In aggregate, +1 insertions constituted the greatest fraction of mutation types, followed by −1, and then −2, although stereotyped repair patterns are evident (**Figure 1A**). Interestingly, insertions were limited to +1 and +2 across all sites, whereas deletions spanned a much broader size range, though with decreasing frequencies for larger deletions.

TABLE 1 | Summary of published CRISPR/Cas9-mediated knockout in woody species.

Species	Genes targeted	Mutation efficiencies	Transformation source tissue(s)	References
Actinidia chinensis (kiwifruit)	*PDS*	65–92%	*In vitro* leaves	Wang Z. et al., 2018*
Citrus sinensis, Citrus paradise, Poncirus trifoliate x *Citrus sinensis* (citrus)	*Cs2g12470, Cs7g03360, LOB1* (promoter and gene) and *PDS*	3–100%	Greenhouse leaves[+], *in vitro* epicotyl segments	Jia and Wang, 2014*[+]; Jia et al., 2016, 2017a,b*[+]; Peng et al., 2017*; Zhang et al., 2017*
Coffea canephora (coffee)	*PDS*	Up to 30%	Embryogenic calli	Breitler et al., 2018
Malus domestica, Malus prunifolia x *Malus pumila* (apple)	*DIPM1, DIPM2, DIPM4,* and *PDS*	Up to 32%	*In vitro* leaves, protoplasts^	Malnoy et al., 2016^; Nishitani et al., 2016
Manihot esculenta (cassava)	*PDS*	97–99%	Embryogenic calli	Odipio et al., 2017*
Parasponia andersonii (tropical tree)	*EIN2, HK4, NSP1,* and *NSP2*	48–89%	Greenhouse tissues	van Zeijl et al., 2018
Populus tomentosa, Populus tremula x alba, Populus tremula x tremuloides (poplar)	*4CL1, 4CL2, 4CL5, AG1, AG2, BRC1-1, BRC2-1, DWF4, LFY, MYB57, MYB115, MYB156, MYB170, PDS,* and *WRKY18*	Up to 100%	*In vitro* leaves, *in vitro* shoots (leaf, stem and petiole pieces)	Fan et al., 2015; Zhou et al., 2015*; Jiang et al., 2017; Wan et al., 2017; Wang et al., 2017; Xu et al., 2017*; Yang et al., 2017; Elorriaga et al., 2018*; Muhr et al., 2018*; Shen et al., 2018*
Theobroma cacao (cacao)	*NPR3*	Up to 27%	*In vitro* somatic embryo cotyledons	Fister et al., 2018*
Vitis vinifera (grape)	*IdnDH, MLO-7, PDS,* and *WRKY52*	0.1–100%	Embryogenic calli, protoplasts^	Malnoy et al., 2016^; Ren et al., 2016*; Nakajima et al., 2017*; Wang X. et al., 2018*

*Indicates off-target mutations were assessed. All transformations performed via Agrobacterium unless otherwise stated. [+]Xcc (Xanthomonas citri subsp. citri)-facilitated agroinfiltration. ^Direct delivery of Cas9-gRNA ribonucleoproteins.

Small mutagenic indels have often been ascribed to the classical non-homologous end-joining (cNHEJ) DNA repair pathway, but recent studies have demonstrated involvement of the alternative end-joining (alt-EJ) pathway as well (Rodgers and McVey, 2016). It now appears that cNHEJ contributes to the most common +1 insertions and other small indels, whereas larger indels are due to alt-EJ. This is based on studies where impaired cNHEJ drastically changed the repair outcomes of CRISPR/Cas9 in yeast, human cells and *Arabidopsis*, such that the typically predominant +1 insertions as well as other small (<3 bp) indels were greatly reduced, while rates of large indels increased, apparently independently of cNHEJ (van Overbeek et al., 2016; Shen et al., 2017; Lemos et al., 2018). In yeast, the vast majority of the +1 insertions from cNHEJ were templated from 1 bp 5′overhangs at the Cas9 cleavage site (fourth base from the PAM), and dependent on POL4, a low-fidelity X-family DNA polymerase with terminal transferase activity (Lemos et al., 2018). POL4-deficient yeast also lost +2 and +3 insertions, many of which are homonucleotides and apparently templated from the Cas9 cleavage site as well (Lemos et al., 2018). Templated insertions could also explain the majority of +1 events in a large Cas9-induced indel dataset from human cells (van Overbeek et al., 2016), suggesting a conserved mechanism underlying +1 insertions in CRISPR/Cas9-edited organisms (Lemos et al., 2018). In the combined 717 dataset, the majority of +1 insertions were

+T as reported in many CRISPR studies. However, evidence in support of templated +1 insertions was weak, and appeared to be target site-dependent (**Figure 1A**). Clearly, much more data with greater target site diversity and coverage are necessary before a conclusion can been drawn, but such data from trees will require significant and perhaps community-wide efforts. Regardless, the small target site collection used in our analysis supports involvement of more than one mechanism for the commonly observed +T insertions, at least in *Populus*.

cNHEJ-independent repair likely involves different alt-EJ pathways, including microhomology-mediated end-joining (MMEJ), single-strand annealing (SSA), or polymerase theta (POLQ)-mediated end-joining (TMEJ) (Rodgers and McVey, 2016). Both MMEJ and SSA require end resection or unwinding to expose short homologous sequences for annealing (up to ~20–30 bp for MMEJ and longer for SSA) and subsequent repair, and always result in deletions (Sfeir and Symington, 2015). The presence of microhomologous sequences at the deletion junctions can therefore serve as evidence of MMEJ/SSA repair. Indeed, microhomologies of 1–5 bp are readily identifiable in most of the deletion (≥5 bp) alleles we examined, but are rarely found for small deletions (3–4 bp) that might have arisen from cNHEJ (**Figure 1B**). MMEJ has also been associated with abnormal chromosomal translocations and inversions (Sfeir and Symington, 2015). Such modifications have been reported in

A

100 80 60 40 20 0 — Insertion (%)   0 20 40 60 80 100 — Deletion (%)

100% +T
(50% +A) 50% +T
100% +T

Legend:
- 1 bp
- 2 bp
- 3 bp
- 4 bp
- 5–10 bp
- 11–20 bp
- >20 bp

100% +T
100% +T
(33% +C) 10% +T
(10% +C) 33% +T

(18) Potri.003G188500   GAGGATGTTGAAGTCT▾GGA**GGG**
(28) Potri.019G049500   GCTGCTGGGTTATCCAA▾TTT**GGG**
(32) Potri.001G036900   GAGGATGATAAAATCTG▾GAG**GGG**

(171) Potri.015G106900[a]   GCCCCCGCCTCAGCAGC▾CAC**CGG**
(75) Potri.015G106900[a]   GATCACAGAGAGAGAGA▾CAA**GGG**
(116) Potri.011G075800[b]   GCCAAGTCACTTTCTGC▾AAA**AGG**
(112) Potri.004G064300[b]   GCCAAGTCACTTTCTGC▾AAA**AGG**

[a] The gene was targeted for editing at two sites; [b] The paralogs were targeted at the consensus site (Elorriaga et al., 2018).

B

Potential DNA repair pathway	Sequence (top, WT allele; bottom, mutation allele)	Source
(1) cNHEJ templated	CTTGGTGGGCAGGCTTATGTTGAAGCCCAAGATGGCTTGAGCGTAAAAA▾GAC**TGG**ATGAGGAAACAGGTACTCTAATATATTT CTTGGTGGGCAGGCTTATGTTGAAGCCCAAGATGGCTTGAGCGTAAAAA**A**GAC**TGG**ATGAGGAAACAGGTACTCTAATATATTT	Nishitani et al., 2016
(2) cNHEJ non-templated	GAAATGGTGCCCCCGCCTCAGCAGC▾CAC**CGG**CTGCGGCGTTTGCTGTAAGG GAAATGGTGCCCCCGCCTCAGCAGC**T**CAC**CGG**CTGCGGCGTTTGCTGTAAGG	Elorriaga et al., 2018
(3) cNHEJ deletion	CTTGTCTTCTTTGAGGATGATAAAATCTG▾GAG**GGG**CTCCATTGGGCAAGGAACTTGAAGATACTGTCAGAGCTAAGTTTCCTC CTTGTCTTCTTTGAGGATGATAAAA**----**.GAG**GGG**CTCCATTGGGCAAGGAACTTGAAGATACTGTCAGAGCCAAGTTTCCTC	Zhou et al., 2015
(4) MMEJ	AGCAGCACAAGGGCTAAGAACTATAGGCGGCGGAGA▾GAG**GGG**AGCTGCAAGATTTGG AGCAGCACAAGGGCTAAGAACTATAGGCGGCGG**---**.**-------------**AGATTTGG	Jia et al., 2017
(5) MMEJ or SSA	AAACCTATTCTTTTGCCGAAACTCACCTAATATCTCGAAAGGTCGCTGCTGGGTTATCCAA▾TTT**GGG**CATCAAGAAAGGCGAT AAACCTATTCT**---**.TTT**GGG**CATCAAGAAAGGCGAT	Zhou et al., 2015
(6) TMEJ templated in trans	GAGAAATGAGTGCATTGAACTTG▾AGC**TGG**CATAGTAAATCATTAGACTCTCAAGTTGCCTTGAGATGTGGCGCTTATCCTACT GAGAAATGAGTGCATTGAACTTG.**----------**CCTTATCATTAGACTCTCAAGTTGCCTTGAGATGTGGCGCTTATCCTACT	Fan et al., 2015
(7) TMEJ templated in cis	ACGTGAAAATTTGGGGTGATCTTTTGTCTACGCTGTAGGTTGTCTGTATGGACTAT**CCA**AGA▾**CCG**GAC▾CTTGATAACACGGT A**CGTGAAAATTTGGGGTG**ATCTTTTGTCTACGCTGTAGGTTGTCTGTATGGACTATCCAAGA.TCACCC.CAAATTTTCACGGT	Fan et al., 2015
(8) TMEJ	3 in trans sites <-- upstream TCTTTTTCAGGGTATT**CCT**GAT▾CGAG........TAACAGATGAGGTTTTCAT --> 2 in trans sites downstream TCTTTTTCAGGGTATT**CCT**GAT.**---GTTTTTTTT**TAACAGATGAGGTTTTCAT	Nakajima et al., 2017

FIGURE 1 | The mutation profiles and potential DNA repair pathways of CRISPR/Cas9-induced DSBs in *Populus* and other trees. **(A)** Distribution of mutation types at distinct genomic sites based on data from Zhou et al. (2015) (top panel) and Elorriaga et al. (2018) (bottom panel). The corresponding target genes and gRNA sequences are shown on the right, with allele number for each noted in parentheses. The percentages shown inside the +1 (1 bp insertion) bar indicate the fraction that were T insertions. The fraction of templated +1 insertions that deviate from T is shown in parentheses. **(B)** Representative examples of different mutation types and the potential DNA repair pathway involved in each case. PAM sequences are bold underlined, triangles denote predicted Cas9-cleavage sites, indels are shown in red, yellow-shaded regions denote microhomologies, and gray sequences in (6) and (7) were appended from *P. tomentosa* cDNA (GenBank accession KC954700) and *P. tremula* x *alba* 717 genomic sequences (Xue et al., 2015), respectively. Note, the region in (7) contains two overlapping target sites. For (8), there are five possible *in trans* template sites within introns of Phytozome (v12) gene model GSVIVG01016650001, the nearest one 640 bp upstream of the target site.

several studies – including two from *Populus* – where two or more gRNAs were designed to target the same gene to produce large deletions (Fan et al., 2015; Elorriaga et al., 2018). Moreover, large deletions are sometimes accompanied by small insertions (Fan et al., 2015; Nakajima et al., 2017), a pattern that is characteristic of the recently discovered TMEJ pathway (Koole et al., 2014). TMEJ depends on POLQ, an error-prone A-family DNA polymerase that can extend microhomologies in a template-dependent (either *in cis* or *in trans*) or independent manner (Kent et al., 2016). TMEJ is the essential repair pathway in animal germ cells, as embryos of zebrafish *polq* mutants are hypersensitive to DSB-inducing treatments, with low levels of repair producing only +1 insertions (Thyme and Schier, 2016). In *Arabidopsis*, TMEJ is required for T-DNA integration following *Agrobacterium* transformation of either flowers or roots (van Kregten et al., 2016). We found evidence of *in cis* or *in trans* templated insertions in the complex indels reported for poplar and grape (**Figure 1B**; Fan et al., 2015; Nakajima et al., 2017), supporting an active TMEJ pathway in somatic cells of plants.

Examination of published mutation profiles of *Populus* and other tree species suggests differential involvement of multiple repair pathways, probably with cNHEJ contributing to +1, +2 and small (1–4 bp) deletions, MMEJ (and SSA) to larger deletions, and TMEJ to complex indels (**Figure 1B**). The varying dependency of these pathways on sequence contexts (microhomologies) likely underpins the non-random nature of CRISPR/Cas9 repair outcomes reported in many studies, including trees (Jacobs et al., 2015; van Overbeek et al., 2016; Vu et al., 2017; Elorriaga et al., 2018). Incorporation of microhomology modeling into the gRNA design workflow (Bae et al., 2014; Segar et al., 2015) should enable prediction of potential DNA repair outcomes for informed selection of target sites.

LONG-TERM STABILITY OF CRISPR-EDITED TREES THROUGH VEGETATIVE PROPAGATION

For many herbaceous species where CRISPR editing efficiencies are low, or where monoallelic/mosaic mutations predominate in the first-generation (T0) transformants, multi-generation progeny screening is necessary to obtain homozygous mutants (Xu et al., 2015). Although initial transmission rates vary depending on the study system and the nature of CRISPR-induced (somatic or germinal) mutations carried by the

founder plant, stable mutation inheritance can be expected once homozygous lines are obtained, as reported for *Arabidopsis*, rice, tomato and potato (Brooks et al., 2014; Feng et al., 2014; Zhou et al., 2014; Butler et al., 2015). For woody perennials, however, the issues are rather different. Cross-generational screening is difficult to implement for transgenic trees owing to their long generation times and strict regulation of flowering transgenic trees (Strauss et al., 2015). The predominantly outcrossing nature of trees, many of which are also dioecious, adds further challenge to rapid-cycle breeding and introgression of CRISPR-derived mutations into elite germplasms. While advances of early-flowering induction in contained environments (Hoenicka et al., 2014; Klocko et al., 2016) promise to accelerate progress, commercial production of many forest, fruit and nut trees relies on clonal propagation of elite genotypes. For woody perennials, therefore, it is pertinent to address long-term stability of CRISPR editing, both on-target and off-target, in vegetatively-propagated T0 transformants.

In theory, CRISPR-induced DNA modifications should lead to permanent mutations in edited cells that can be inherited mitotically during clonal propagation, yet experimental data are rare. One study used tissue culture to clone CRISPR-derived mutations from T0 diploid and tetraploid potato, and reported stable maintenance of targeted mutations across clonal generations, and in three selected cases, through the germline as well (Butler et al., 2015). In that study, however, somatic mutations were prevalent in T0 plants, as fewer than half of the originally-detected mutation types were captured as single mutations in clonally-propagated plants (Butler et al., 2015). The high levels of somatic mutations likely reflect a high proportion of chimeras, a common problem in tissue culture when plants are regenerated from multiple cells, in this case, with heterogenomic modifications. Fortunately for *Populus*, the proven efficiency of CRISPR (Zhou et al., 2015; Elorriaga et al., 2018) means null mutations with biallelic KO can be readily obtained in T0 transformants and stably inherited through clonal propagation. Wang et al. (2017) reported faithful maintenance of *PtoMYB115* mutations in tissue culture-propagated *Populus tomentosa* somaclones, though in one case low frequencies of new mutations not seen in the parent line were detected, indicative of chimeras. Similarly, the CRISPR editing outcomes of *BRC1-1* and *BRC2-1* were also stable over multiple cycles of vegetative propagation in tissue culture (Muhr et al., 2018). We have maintained a subset of the previously-characterized 717 mutants (Zhou et al., 2015) in the greenhouse for over 4 years by repeatedly cutting back the original transformants and/or propagation using rooted cuttings. The reddish-brown wood discoloration of lignin-reduced *4cl1* mutants has been stable in all re-sprouted shoots or clonally-propagated plants. Repeated amplicon-sequencing of randomly selected lines re-confirmed the targeted *4CL1* mutations 4 years later, with no off-target changes to the paralogous *4CL5* (**Supplementary Table S1**). Another group of transgenic plants harbors a non-functional gRNA for *4CL5* due to SNPs (one per allele) between the genome-sequenced *P. trichocarpa* and the transformation host 717 that prevented Cas9 cleavage as confirmed by amplicon sequencing (Zhou et al., 2015). It should be noted that one of the 717 SNPs alters the PAM site from NGG to NGA, the latter is a non-canonical PAM of SpCas9 thought to cause off-target cleavage in human cells (Zhang Y. et al., 2014b). Reanalyzing this group of plants will inform as to whether the imperfectly-matched *4CL5*-gRNA exhibited any off-target activity over the long term. We found no evidence of CRISPR/Cas9 cleavage after 4 years of coppicing and regrowth (**Supplementary Table S1**). These findings echo other tree studies that showed no or very rare off-targeting (Jia et al., 2017a; Nakajima et al., 2017; Elorriaga et al., 2018; see also **Table 1**), as well as reports from *Arabidopsis*, rice and tomato based on whole-genome re-sequencing (Feng et al., 2014; Zhang et al., 2014a; Peterson et al., 2016; Rodríguez-Leal et al., 2017). The data provide support for long-term stability and specificity of CRISPR/Cas9-mediated mutagenesis, with extremely low off-target potential during vegetative propagation in poplar.

BROAD-SPECTRUM MUTAGENESIS BEYOND KO

Nullizygous mutations harboring either identical (homozygous) or distinct (heterozygous) mutations in all alleles of the genome are the ideal repair outcomes for gene KO investigation. However, monoallelic, in-frame and/or mosaic mutations can expand the phenotypic spectrum to enhance the power of functional analysis. For instance, transgenic grapevine with monoallelic mutations of a defense-related *WRKY* gene exhibited intermediate levels of disease resistance between WT and biallelic mutants (Wang X. et al., 2018). Similarly, monoallelic or in-frame mutations of PDS led to partial albino phenotypes in both poplar and apple (Fan et al., 2015; Nishitani et al., 2016). Given the abundance of duplicate genes in plant genomes, and the proven successes of CRISPR in multi-allele as well as allele-specific editing (Jia et al., 2016; Elorriaga et al., 2018), there is exciting potential to exploit CRISPR for development of allelic series mutations to address functional redundancy of duplicate genes or tandem repeats, and to investigate the allele-dose response of agronomic traits. Thus, the ability to generate novel germplasms is invaluable not only for tree improvement but also for basic functional genomics research.

In contrast to CRISPR-mediated KO, site-specific gene targeting or replacement remains a major challenge in plants, due to the inefficient homology-directed repair pathway. Geminivirus replicons have been shown to increase site-specific gene knockin (KI) efficiencies by orders of magnitude in tobacco, tomato and hexaploid wheat (Baltes et al., 2014; Čermák et al., 2015; Gil-Humanes et al., 2017). In animal systems, the MMEJ and SSA pathways, along with a new CRISPR/Cpf1 system have been harnessed for targeted KI with success (Sakuma et al., 2016; Tóth et al., 2016). These and other emerging approaches represent promising options for developing efficient KI systems in trees. Finally, many economically important tree species or genotypes remain recalcitrant to transformation and/or tissue culture regeneration, hindering applications of CRISPR. Recent

breakthroughs in morphogenic regulator-mediated regeneration (Lowe et al., 2016, 2018) have already stimulated similar research in trees. Direct delivery of pre-assembled Cas9-gRNA ribonucleoproteins into protoplasts for genome editing as already deployed in apple and grape offers a transgene-free alternative to *Agrobacterium* transformation (Malnoy et al., 2016). At the present time, however, protoplast regeneration for other tree species remains a challenge. There is strong incentive to overcome this challenge since avoiding the footprint of foreign DNA and the associated negative perceptions will improve the outlook for integration of CRISPR technology with commercial deployment of designer trees.

AUTHOR CONTRIBUTIONS

C-JT conceived the idea. WPB, DC, and C-JT collected background information and analyzed data. WPB and C-JT wrote the manuscript with contributions from DC. All authors approved the manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2018.01732/full#supplementary-material

REFERENCES

Alessandra, F., and Shihshieh, H. (2010). Tapping RNA silencing pathways for plant biotechnology. *Plant Biotechnol. J.* 8, 655–677. doi: 10.1111/j.1467-7652.2010.00505.x

Bae, S., Kweon, J., Kim, H. S., and Kim, J.-S. (2014). Microhomology-based choice of Cas9 nuclease target sites. *Nat. Methods* 11:705. doi: 10.1038/nmeth.3015

Baltes, N. J., Gil-Humanes, J., Cermak, T., Atkins, P. A., and Voytas, D. F. (2014). DNA replicons for plant genome engineering. *Plant Cell* 26, 151–163. doi: 10.1105/tpc.113.119792

Bortesi, L., Zhu, C., Zischewski, J., Perez, L., Bassié, L., Nadi, R., et al. (2016). Patterns of CRISPR/Cas9 activity in plants, animals and microbes. *Plant Biotechnol. J.* 14, 2203–2216. doi: 10.1111/pbi.12634

Breitler, J.-C., Dechamp, E., Campa, C., Zebral Rodrigues, L. A., Guyot, R., Marraccini, P., et al. (2018). CRISPR/Cas9-mediated efficient targeted mutagenesis has the potential to accelerate the domestication of *Coffea canephora*. *Plant Cell Tissue Organ Cult.* 134, 383–394. doi: 10.1007/s11240-018-1429-2

Brooks, C., Nekrasov, V., Lippman, Z. B., and Van Eck, J. (2014). Efficient gene editing in tomato in the first generation using the clustered regularly interspaced short palindromic repeats/CRISPR-associated9 system. *Plant Physiol.* 166, 1292–1297. doi: 10.1104/pp.114.247577

Burstein, D., Harrington, L. B., Strutt, S. C., Probst, A. J., Anantharaman, K., Thomas, B. C., et al. (2016). New CRISPR–Cas systems from uncultivated microbes. *Nature* 542, 237–241. doi: 10.1038/nature21059

Butler, N. M., Atkins, P. A., Voytas, D. F., and Douches, D. S. (2015). Generation and inheritance of targeted mutations in potato (*Solanum tuberosum* L.) using the CRISPR/Cas system. *PLoS One* 10:e0144591. doi: 10.1371/journal.pone.0144591

Čermák, T., Baltes, N. J., Čegan, R., Zhang, Y., and Voytas, D. F. (2015). High-frequency, precise modification of the tomato genome. *Genome Biol.* 16:232. doi: 10.1186/s13059-015-0796-9

Elorriaga, E., Klocko, A. L., Ma, C., and Strauss, S. H. (2018). Variation in mutation spectra among CRISPR/Cas9 mutagenized poplars. *Front. Plant Sci.* 9:594. doi: 10.3389/fpls.2018.00594

Fan, D., Liu, T., Li, C., Jiao, B., Li, S., Hou, Y., et al. (2015). Efficient CRISPR/Cas9-mediated targeted mutagenesis in Populus in the first generation. *Sci. Rep.* 5:12217. doi: 10.1038/srep12217

Feng, Z., Mao, Y., Xu, N., Zhang, B., Wei, P., Yang, D.-L., et al. (2014). Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in Arabidopsis. *Proc. Natl. Acad. Sci. U.S.A.* 111, 4632–4637. doi: 10.1073/pnas.1400822111

Fister, A. S., Landherr, L., Maximova, S. N., and Guiltinan, M. J. (2018). Transient expression of CRISPR/Cas9 machinery targeting TcNPR3 enhances defense response in *Theobroma cacao*. *Front. Plant Sci.* 9:268. doi: 10.3389/fpls.2018.00268

Gil-Humanes, J., Wang, Y., Liang, Z., Shan, Q., Ozuna, C. V., Sánchez-León, S., et al. (2017). High-efficiency gene targeting in hexaploid wheat using DNA replicons and CRISPR/Cas9. *Plant J.* 89, 1251–1262. doi: 10.1111/tpj.13446

Hoenicka, H., Lehnhardt, D., Nilsson, O., Hanelt, D., and Fladung, M. (2014). Successful crossings with early flowering transgenic poplar: interspecific crossings, but not transgenesis, promoted aberrant phenotypes in offspring. *Plant Biotechnol. J.* 12, 1066–1074. doi: 10.1111/pbi.12213

Hu, Y., Zhang, J., Jia, H., Sosso, D., Li, T., Frommer, W. B., et al. (2014). Lateral organ boundaries 1 is a disease susceptibility gene for citrus bacterial canker disease. *Proc. Natl. Acad. Sci. U.S.A.* 111, E521–E529. doi: 10.1073/pnas.1313271111

Jacobs, T. B., Lafayette, P. R., Schmitz, R. J., and Parrott, W. A. (2015). Targeted genome modifications in soybean with CRISPR/Cas9. *BMC Biotechnol.* 15:16. doi: 10.1186/s12896-015-0131-2

Jia, H., Orbovic, V., Jones, J. B., and Wang, N. (2016). Modification of the PthA4 effector binding elements in Type I CsLOB1 promoter using Cas9/sgRNA to produce transgenic Duncan grapefruit alleviating XccΔpthA4:dCsLOB1.3 infection. *Plant Biotechnol. J.* 14, 1291–1301. doi: 10.1111/pbi.12495

Jia, H., and Wang, N. (2014). Targeted genome editing of sweet orange using Cas9/sgRNA. *PLoS One* 9:e93806. doi: 10.1371/journal.pone.0093806

Jia, H., Xu, J., Orbovic, V., Zhang, Y., and Wang, N. (2017a). Editing citrus genome via SaCas9/sgRNA system. *Front. Plant Sci.* 8:2135. doi: 10.3389/fpls.2017.02135

Jia, H., Zhang, Y., Orbović, V., Xu, J., White, F. F., Jones, J. B., et al. (2017b). Genome editing of the disease susceptibility gene CsLOB1 in citrus confers

resistance to citrus canker. *Plant Biotechnol. J.* 15, 817–823. doi: 10.1111/pbi. 12677

Jiang, F., and Doudna, J. A. (2017). CRISPR–Cas9 structures and mechanisms. *Annu. Rev. Biophys.* 46, 505–529. doi: 10.1146/annurev-biophys-062215- 010822

Jiang, Y., Guo, L., Ma, X., Zhao, X., Jiao, B., Li, C., et al. (2017). The WRKY transcription factors PtrWRKY18 and PtrWRKY35 promote Melampsora resistance in Populus. *Tree Physiol.* 37, 665–675. doi: 10.1093/treephys/tpx008

Kent, T., Mateos-Gomez, P. A., Sfeir, A., and Pomerantz, R. T. (2016). Polymerase θ is a robust terminal transferase that oscillates between three different mechanisms during end-joining. *eLife* 5:e13740. doi: 10.7554/eLife.13740

Klocko, A. L., Ma, C., Robertson, S., Esfandiari, E., Nilsson, O., and Strauss, S. H. (2016). FT overexpression induces precocious flowering and normal reproductive development in Eucalyptus. *Plant Biotechnol. J.* 14, 808–819. doi: 10.1111/pbi.12431

Koole, W., Van Schendel, R., Karambelas, A. E., Van Heteren, J. T., Okihara, K. L., and Tijsterman, M. (2014). A Polymerase Theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. *Nat. Commun.* 5:3216. doi: 10.1038/ncomms4216

Lemos, B. R., Kaplan, A. C., Bae, J. E., Ferrazzoli, A. E., Kuo, J., Anand, R. P., et al. (2018). CRISPR/Cas9 cleavages in budding yeast reveal templated insertions and strand-specific insertion/deletion profiles. *Proc. Natl. Acad. Sci. U.S.A.* 115, E2040–E2047. doi: 10.1073/pnas.1716855115

Lowe, K., La Rota, M., Hoerster, G., Hastings, C., Wang, N., Chamberlin, M., et al. (2018). Rapid genotype "independent" *Zea mays* L. (maize) transformation via direct somatic embryogenesis. *In Vitro Cell. Dev. Biol.* 54, 240–252. doi: 10.1007/s11627-018-9905-2

Lowe, K., Wu, E., Wang, N., Hoerster, G., Hastings, C., Cho, M.-J., et al. (2016). Morphogenic regulators Baby boom and Wuschel improve monocot transformation. *Plant Cell* 28, 1998–2015. doi: 10.1105/tpc.16.00124

Malnoy, M., Viola, R., Jung, M. H., Koo, O. J., Kim, S., Kim, J. S., et al. (2016). DNA-free genetically edited grapevine and apple protoplast using CRISPR/Cas9 ribonucleoproteins. *Front. Plant Sci.* 7:1904. doi: 10.3389/fpls.2016.01904

Muhr, M., Paulat, M., Awwanah, M., Brinkkötter, M., and Teichmann, T. (2018). CRISPR/Cas9-mediated knockout of Populus BRANCHED1 and BRANCHED2 orthologs reveals a major function in bud outgrowth control. *Tree Physiol.* 38, 1588–1597. doi: 10.1093/treephys/tpy088

Murovec, J., Pirc, Ž, and Yang, B. (2017). New variants of CRISPR RNA-guided genome editing enzymes. *Plant Biotechnol. J.* 15, 917–926. doi: 10.1111/pbi. 12736

Nakajima, I., Ban, Y., Azuma, A., Onoue, N., Moriguchi, T., Yamamoto, T., et al. (2017). CRISPR/Cas9-mediated targeted mutagenesis in grape. *PLoS One* 12:e0177966. doi: 10.1371/journal.pone.0177966

Nishitani, C., Hirai, N., Komori, S., Wada, M., Okada, K., Osakabe, K., et al. (2016). Efficient genome editing in apple using a CRISPR/Cas9 system. *Sci. Rep.* 6:31481. doi: 10.1038/srep31481

Odipio, J., Alicai, T., Ingelbrecht, I., Nusinow, D. A., Bart, R., and Taylor, N. J. (2017). Efficient CRISPR/Cas9 genome editing of phytoene desaturase in cassava. *Front. Plant Sci.* 8:1780. doi: 10.3389/fpls.2017.01780

Peng, A., Chen, S., Lei, T., Xu, L., He, Y., Wu, L., et al. (2017). Engineering canker- resistant plants through CRISPR/Cas9-targeted editing of the susceptibility gene CsLOB1 promoter in citrus. *Plant Biotechnol. J.* 15, 1509–1519. doi: 10. 1111/pbi.12733

Peterson, B. A., Haak, D. C., Nishimura, M. T., Teixeira, P. J., James, S. R., Dangl, J. L., et al. (2016). Genome-wide assessment of efficiency and specificity in CRISPR/Cas9 mediated multiple site targeting in Arabidopsis. *PLoS One* 11:e0162169. doi: 10.1371/journal.pone.0162169

Ren, C., Liu, X., Zhang, Z., Wang, Y., Duan, W., Li, S., et al. (2016). CRISPR/Cas9- mediated efficient targeted mutagenesis in Chardonnay (*Vitis vinifera* L.). *Sci. Rep.* 6:32289. doi: 10.1038/srep32289

Rodgers, K., and McVey, M. (2016). Error-prone repair of DNA double-strand breaks. *J. Cell. Physiol.* 231, 15–24. doi: 10.1002/jcp.25053

Rodríguez-Leal, D., Lemmon, Z. H., Man, J., Bartlett, M. E., and Lippman, Z. B. (2017). Engineering quantitative trait variation for crop improvement by genome editing. *Cell* 171, 470.e–480.e. doi: 10.1016/j.cell.2017.08.030

Sakuma, T., Nakade, S., Sakane, Y., Suzuki, K.-I. T., and Yamamoto, T. (2016). MMEJ-assisted gene knock-in using TALENs and CRISPR-Cas9 with the PITCh systems. *Nat. Protoc.* 11, 118–133. doi: 10.1038/nprot.2015.140

Segar, M. W., Sakofsky, C. J., Malkova, A., and Liu, Y. (2015). MMBIRFinder: a tool to detect microhomology-mediated break-induced replication. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 12, 799–806. doi: 10.1109/TCBB.2014.2359450

Sentmanat, M. F., Peters, S. T., Florian, C. P., Connelly, J. P., and Pruett-Miller, S. M. (2018). A survey of validation strategies for CRISPR-Cas9 editing. *Sci. Rep.* 8:888. doi: 10.1038/s41598-018-19441-8

Sfeir, A., and Symington, L. S. (2015). Microhomology-mediated end joining: a back-up survival mechanism or dedicated pathway? *Trends Biochem. Sci.* 40, 701–714. doi: 10.1016/j.tibs.2015.08.006

Shen, H., Strunks, G. D., Klemann, B. J. P. M., Hooykaas, P. J. J., and De Pater, S. (2017). CRISPR/Cas9-induced double-strand break repair in Arabidopsis nonhomologous end-joining mutants. *G3* 7, 193–202. doi: 10.1534/g3.116. 035204

Shen, Y., Li, Y., Xu, D., Yang, C., Li, C., and Luo, K. (2018). Molecular cloning and characterization of a brassinosteriod biosynthesis-related gene PtoDWF4 from *Populus tomentosa*. *Tree Physiol.* 38, 1424–1436. doi: 10.1093/treephys/tpy027

Strauss, S. H., Costanza, A., and Séguin, A. (2015). Genetically engineered trees: paralysis from good intentions. *Science* 349, 794–795. doi: 10.1126/science. aab0493

Thyme, S. B., and Schier, A. F. (2016). Polq-mediated end joining is essential for surviving DNA double-strand breaks during early zebrafish development. *Cell Rep.* 15, 707–714. doi: 10.1016/j.celrep.2016.03.072

Tóth, E., Weinhardt, N., Bencsura, P., Huszár, K., Kulcsár, P. I., Tálas, A., et al. (2016). Cpf1 nucleases demonstrate robust activity to induce DNA modification by exploiting homology directed repair pathways in mammalian cells. *Biol. Direct* 11:46. doi: 10.1186/s13062-016-0147-0

van Kregten, M., De Pater, S., Romeijn, R., Van Schendel, R., Hooykaas, P. J. J., and Tijsterman, M. (2016). T-DNA integration in plants results from polymerase- θ-mediated DNA repair. *Nat. Plants* 2:16164. doi: 10.1038/nplants.2016.164

van Overbeek, M., Capurso, D., Carter, M. M., Thompson, M. S., Frias, E., Russ, C., et al. (2016). DNA repair profiling reveals nonrandom outcomes at Cas9- mediated breaks. *Mol. Cell.* 63, 633–646. doi: 10.1016/j.molcel.2016.06.037

van Zeijl, A., Wardhani, T. A. K., Seifi Kalhor, M., Rutten, L., Bu, F., Hartog, M., et al. (2018). CRISPR/Cas9-mediated mutagenesis of four putative symbiosis genes of the tropical tree *Parasponia andersonii* reveals novel phenotypes. *Front. Plant Sci.* 9:284. doi: 10.3389/fpls.2018.00284

Vu, G. T. H., Cao, H. X., Fauser, F., Reiss, B., Puchta, H., and Schubert, I. (2017). Endogenous sequence patterns predispose the repair modes of CRISPR/Cas9- induced DNA double-stranded breaks in *Arabidopsis thaliana*. *Plant J.* 92, 57–67. doi: 10.1111/tpj.13634

Wan, S., Li, C., Ma, X., and Luo, K. (2017). PtrMYB57 contributes to the negative regulation of anthocyanin and proanthocyanidin biosynthesis in poplar. *Plant Cell Rep.* 36, 1263–1276. doi: 10.1007/s00299-017-2151-y

Wang, L., Ran, L., Hou, Y., Tian, Q., Li, C., Liu, R., et al. (2017). The transcription factor MYB115 contributes to the regulation of proanthocyanidin biosynthesis and enhances fungal resistance in poplar. *New Phytol.* 215, 351–367. doi: 10. 1111/nph.14569

Wang, X., Tu, M., Wang, D., Liu, J., Li, Y., Li, Z., et al. (2018). CRISPR/Cas9- mediated efficient targeted mutagenesis in grape in the first generation. *Plant Biotechnol. J.* 16, 844–855. doi: 10.1111/pbi.12832

Wang, Z., Wang, S., Li, D., Zhang, Q., Li, L., Zhong, C., et al. (2018). Optimized paired-sgRNA/Cas9 cloning and expression cassette triggers high-efficiency multiplex genome editing in kiwifruit. *Plant Biotechnol. J.* 16, 1424–1433. doi: 10.1111/pbi.12884

Xu, C., Fu, X., Liu, R., Guo, L., Ran, L., Li, C., et al. (2017). PtoMYB170 positively regulates lignin deposition during wood formation in poplar and confers drought tolerance in transgenic Arabidopsis. *Tree Physiol.* 37, 1713–1726. doi: 10.1093/treephys/tpx093

Xu, R.-F., Li, H., Qin, R.-Y., Li, J., Qiu, C.-H., Yang, Y.-C., et al. (2015). Generation of inheritable and "transgene clean" targeted genome-modified rice in later generations using the CRISPR/Cas9 system. *Sci. Rep.* 5:11491. doi: 10.1038/ srep11491

Xue, L.-J., Alabady, M. S., Mohebbi, M., and Tsai, C.-J. (2015). Exploiting genome variation to improve next-generation sequencing data analysis and genome editing efficiency in *Populus tremula* x alba 717-1B4. *Tree Genet. Genomics* 11:82. doi: 10.1007/s11295-015-0907-5

Xue, L.-J., and Tsai, C.-J. (2015). AGEseq: analysis of genome editing by sequencing. *Mol. Plant* 8, 1428–1430. doi: 10.1016/j.molp.2015.1006.1001

Yang, L., Zhao, X., Ran, L., Li, C., Fan, D., and Luo, K. (2017). PtoMYB156 is involved in negative regulation of phenylpropanoid metabolism and secondary cell wall biosynthesis during wood formation in poplar. *Sci. Rep.* 7:41209. doi: 10.1038/srep41209

Zhang, F., Leblanc, C., Irish, V. F., and Jacob, Y. (2017). Rapid and efficient CRISPR/Cas9 gene editing in Citrus using the YAO promoter. *Plant Cell Rep.* 36, 1883–1887. doi: 10.1007/s00299-017-2202-4

Zhang, H., Zhang, J., Wei, P., Zhang, B., Gou, F., Feng, Z., et al. (2014a). The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. *Plant Biotechnol. J.* 12, 797–807. doi: 10.1111/pbi.12200

Zhang, Y., Ge, X., Yang, F., Zhang, L., Zheng, J., Tan, X., et al. (2014b). Comparison of non-canonical PAMs for CRISPR/Cas9-mediated DNA cleavage in human cells. *Sci. Rep.* 4:5405. doi: 10.1038/srep05405

Zhou, H., Liu, B., Weeks, D. P., Spalding, M. H., and Yang, B. (2014). Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. *Nucleic Acids Res.* 42, 10903–10914. doi: 10.1093/nar/gku806

Zhou, X., Jacobs, T. B., Xue, L.-J., Harding, S. A., and Tsai, C.-J. (2015). Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial Populus reveals 4-coumarate:CoA ligase specificity and redundancy. *New Phytol.* 208, 298–301. doi: 10.1111/nph.13470

Check for
updates

Strategies for Engineering Reproductive Sterility in Plantation Forests

Steffi Fritsche[1], Amy L. Klocko[2], Agnieszka Boron[1], Amy M. Brunner[3] and Glenn Thorlby[1]*

[1] Scion, Rotorua, New Zealand, [2] Department of Biology, University of Colorado Colorado Springs, Colorado Springs, CO, United States, [3] Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, VA, United States

A considerable body of research exists concerning the development of technologies to engineer sterility in forest trees. The primary driver for this work has been to mitigate concerns arising from gene flow from commercial plantings of genetically engineered (GE) trees to non-GE plantations, or to wild or feral relatives. More recently, there has been interest in the use of sterility technologies as a means to mitigate the global environmental and socio-economic damage caused by the escape of non-native invasive tree species from planted forests. The current sophisticated understanding of the molecular processes underpinning sexual reproduction in angiosperms has facilitated the successful demonstration of a number of control strategies in hardwood tree species, particularly in the model hardwood tree Poplar. Despite gymnosperm softwood trees, such as pines, making up the majority of the global planted forest estate, only pollen sterility, via cell ablation, has been demonstrated in softwoods. Progress has been limited by the lack of an endogenous model system, long timescales required for testing, and key differences between softwood reproductive pathways and those of well characterized angiosperm model systems. The availability of comprehensive genome and transcriptome resources has allowed unprecedented insights into the reproductive processes of both hardwood and softwood tree species. This increased fundamental knowledge together with the implementation of new breeding technologies, such as gene editing, which potentially face a less oppressive regulatory regime, is making the implementation of engineered sterility into commercial forestry a realistic possibility.

Keywords: sterility, reproduction, forest trees, gene editing, genetic engineering, containment

DRIVERS FOR ENGINEERING STERILE FOREST TREES

Increasing global population coupled with transition to a sustainable bio-based economy is predicted to lead to growing pressure on forests to deliver wood-based products, energy, food, and ecosystem services whilst maintaining their role as major reservoirs of biodiversity. To accommodate this growing demand, it is estimated that the amount of wood we take from forests and plantations each year may need to triple by 2050 (WWF, 2015). Planted forests, which in 2015 made up 7% of forest lands, provide a means to sustainably increase production of forest products and reduce pressure on natural forests (FAO, 2015). Alongside improved silviculture, land

management and other technological advances, biotech-based technologies offer tools to enhance the sustainability and productivity of planted forests (Al-Ahmad, 2018). Sterile trees have the ability to impact a number of obstacles to increasing productivity from planted forests.

Containment of Genetically Engineered Trees

Genetic engineering (GE) is able to provide solutions for many of the challenges forestry faces to sustainably increase forest production. Improved wood quantity and quality, processability, biotic, and abiotic stress tolerance and herbicide tolerance (Harfouche et al., 2011; Porth and El-Kassaby, 2014; Etchells et al., 2015; Ault et al., 2016; Zhou et al., 2017) are amongst the traits successfully demonstrated. The recent approval by the Brazilian Regulator for GE Eucalyptus that are able to grow 15–20% faster than the best existing clonal lines (Nature Biotechnology News, 2015) seems likely to lead to first large-scale commercial planting of trees.

There remain well documented regulatory and social challenges associated with commercial planting of GE trees (Porth and El-Kassaby, 2014; Strauss et al., 2016). Gene flow from transgenic trees remains a major concern, particularly as forest trees are virtually undomesticated and pollen is able to disseminate over great distances (DiFazio et al., 2004). Seeds also have the potential to spread, either locally or over distances, depending on the species. Transgene containment through the production of trees that are unable to produce fertile reproductive propagules has the ability to mitigate these concerns and prevent, or severely reduce, the flow of genes via sexual reproduction.

Invasive Tree Species

Increasing attention is being paid to the ecological, economic, and cultural damage caused by invasive tree species that have "escaped" by seed dispersal from planted forests (Breton et al., 2008; Nuñez et al., 2017). Globally, Pinus species are recognized as among the most widespread and influential of all invasive plants (Richardson and Rejmánek, 2004). These escapes, or wildings, are particularly a problem in the Southern Hemisphere where a large percentage of tree plantations are composed of exotic species (Franzese and Raffaele, 2017). South Africa, New Zealand, and Australia, who were early adopters of exotic conifer plantations, have been joined more recently by several South American nations in facing wilding challenges (Simberloff et al., 2010). For example, in New Zealand several exotic conifer species have become established and now occupy ∼1.8 million ha, and are expanding by about 6% per annum (Froude, 2011). Economic and ecological damage resulting from these wildings is challenging the license to operate, with commercially advantageous, but wilding-prone species such as Douglas-fir (Pseudotsuga menziesii). The ability to generate trees that are unable to reproduce would allow control programs to focus on the existing

populations and give forest owners freedom to operate for new plantings.

Increased Wood Production and Other Benefits

The ability to either prevent reproduction or limit the development of reproductive propagules is predicted to boost growth and increase wood production in forest trees by redirecting energy and nutrients to increased vegetative growth (Strauss et al., 1995; Luis and José, 2014). Conclusive evidence for such a reproductive cost is lacking but is supported by evidence that in conifers cone production may utilize a significant proportion of the trees energy and assimilates (Cremer, 1992; Sala et al., 2012; Kramer et al., 2014). Unsurprisingly, in conifers, the long-lived female cones are more energy demanding than the generally more transient male cones (Obeso, 2002). These observations suggest that engineered sterility, particularly female sterility is likely to have a positive impact on vegetative growth and wood production. Long-term growth comparisons between sterile and reproductive trees would provide direct evidence for this and allow quantification of growth differences.

Pollen from many trees cause allergenic reactions and symptoms correlate with exposure (Buters et al., 2012). Planted forests can provide a major source of seasonal allergens (D'amato et al., 2007). For example, allergy to sugi (Cryptomeria japonica) pollen is reported to effect 26.5% of the Japanese population (Taniguchi, 2018). The ability to prevent or limit pollen production from planted forests would provide relief to allergy suffers and mitigate potential social license to operate challenges.

CURRENT UNDERSTANDING OF REPRODUCTIVE PROCESSES IN FOREST TREES

Both angiosperm (hardwood) and gymnosperm/conifer (softwood) trees are used as plantation species. Although they share broad similarities in their reproductive processes, there are distinct differences between them.

Angiosperm Trees

Perhaps no other plant development process has been studied more than flowering. For Arabidopsis, the in planta functions of a large number of flowering genes as well as their regulatory network context are known and studies in plants such as rice and petunia have revealed broad functional conservation (Pajoro et al., 2014). These include genes that regulate the transition to flowering, floral organ identity as well as pollen and ovule development. Although advances in sequencing have enabled the identification of flowering gene homologs in diverse angiosperm trees, there are few cases where in planta functions have been characterized in trees (Brunner et al., 2017; Klocko et al., 2018). This is due to the long non-flowering period that can last years to decades and that for most species, genetic transformation is a formidable hurdle. Trees also differ from herbaceous plants in the prolonged period between the floral transition and anthesis.

In tropical species such as Eucalyptus, this occurs in one season, but temperate species exhibit indirect flowering, with flower development initiated in 1 year and completed the following year (Vining et al., 2015; Brunner et al., 2017). Thus, multi-year field trials that require monitoring of large trees and collecting flowers from the upper portion of the tree crown are typically required to demonstrate sterility or delay of flowering.

Selection of candidate genes for genetic containment in trees based on homology to Arabidopsis flowering genes and gene expression might be straightforward, but such conservation does not necessarily translate to the predicted or desired phenotype. Flowering time genes are attractive targets because prevention of flowering is easier to monitor (e.g., no need to demonstrate flowers are sterile) and to prevent resource allocation to reproduction. However, accumulating evidence supports that tree homologs of various flowering time and floral meristem identity genes have roles in both vegetative and reproductive phenology (Bohlenius et al., 2006; Bielenberg et al., 2008; Hoenicka et al., 2008; Mohamed et al., 2010; Hsu et al., 2011; Azeez et al., 2014; Tylewicz et al., 2015; Parmentier-Line and Coleman, 2016). Targeting such genes for manipulation can thus result in undesired vegetative effects, such as delayed bud flush, in addition to predicted effects on flowering or no effect on flowering (Hoenicka et al., 2012). However, promising results have also been achieved, such as the delayed flowering without growth reduction demonstrated by overexpressing the poplar ortholog of the floral repressor *SHORT VEGETATIVE PHASE (SVP)* (Klocko et al., 2018). Manipulation of floral organ identity genes might be less likely to have vegetative effects as these genes may show stronger conservation of reproductive-only function. For example, considerable evidence supports that the *AGAMOUS (AG)* subgroup of MADS-box genes have reproductive functions not only in angiosperms but also in gymnosperms (Dreni and Kater, 2014). However, even in these cases, results can differ from expectation. For example, downregulation of the conserved floral meristem identity gene *LEAFY (LFY)* in a male poplar genotype induced bisexual and female flowers (Klocko et al., 2018). Despite the challenges, the knowledge gained from gene function and sterility studies in trees along with more detailed and extensive genome-wide expression studies in different angiosperm trees will enable more accurate gene selection for manipulation of only-reproductive traits.

Gymnosperm Trees

Unlike angiosperms, where there is extensive knowledge of the molecular factors involved in the reproduction process, relatively little is known regarding gymnosperms. A number of putative genes have been identified through comparative analyses of orthologous angiosperm genes, tissue-specific expression analysis or genome sequencing. However, several key floral genes including *FD*, *SQUAMOSA-* (*SQUA-*) or *SEPALLATA-like* (*SEP-like*) seem to be absent (Becker, 2003; Abe et al., 2005; Zahn et al., 2005; Melzer et al., 2010; Karlgren et al., 2011; Jaeger et al., 2013). Initial research was able to detect orthologs to only the B- and C-genes involved in the control of meristem formation and organ identity in the developing cones (Tandre et al., 1995, 1998; Mouradov et al., 1998; Rutledge et al., 1998; Fukui et al., 2001;

Sundström and Engström, 2002; Gramzow et al., 2014; Katahata et al., 2014; Uddenberg et al., 2015). More recently, ABCE model prototype transcription factors, genes that define the developmental flower organ model (ABC(DE)) in angiosperms, have been confirmed in gymnosperms (Chen et al., 2017). Conifer-specific genes such as the *DEFICIENS-AGAMOUS-LIKE (DAL)* and *NEEDLY* have also been identified but functional knowledge is limited due to the lack of angiosperm orthologs (Carlsbecker et al., 2003, 2004; Rudall et al., 2011). It is not generally possible to predict which of the many vegetative meristems will undergo the reproductive bud transition before changes are initiated making research on reproductive initiation a bold venture (Williams, 2009).

The biggest bottleneck for conifer reproduction research is the inability to carry out functional characterization in an endogenous system which prevents the definitive elucidation of gene function. Testing of gene function in angiosperm model systems has produced inconclusive results. Whilst some such studies have confirmed the function of putative orthologs, others failed to find flowering related differences, found multiple phenotypic alteration or were unable to complement mutants, highlighting the need for a reliable conifer testing system (Rutledge et al., 1998; Tandre et al., 1998; Shindo et al., 2001; Sundström and Engström, 2002; Carlsbecker et al., 2003, 2004; Nilsson et al., 2007; Shiokawa et al., 2008; Klintenas et al., 2012; Katahata et al., 2014; Liu et al., 2018). As discussed above for angiosperm trees, it might be challenging to identify flowering time gene homologs in conifers that do not have roles in vegetative development that make their manipulation for reproductive sterility problematic. For example, the gymnosperm *FLOWERING LOCUS T-like* subfamily has been suggested to have roles in both vegetative and reproductive phenology (Klintenas et al., 2012; Karlgren et al., 2013; Nystedt et al., 2013; Liu et al., 2016). However, the lack of characterization in an endogenous system means that the function of the sub-family members remains unresolved.

ENGINEERED STERILITY IN TREES

The increasingly sophisticated understanding of the molecular processes underpinning sexual reproduction described above has facilitated the successful demonstration of a number of sterility strategies in plants. Chief amongst these are strategies using ablation of reproductive cells or structures and the inactivation or suppression of genes essential for normal reproductive processes. Here, we highlight only a selection of sterility approaches and refer readers to Brunner et al. (2007) and Hoenicka et al. (2016b) for additional examples.

The use of cell or tissue-specific promoters to direct the expression of cytotoxic genes (Palmiter et al., 1987) to reproductive tissues has been widely used to investigate and modify reproductive development in plants (Goldman et al., 1994; Beals and Goldberg, 1997). Numerous examples exist of using this technology in plants to generate male and female sterility (Mariani et al., 1990; Goldman et al., 1994; De Block et al., 1997). Complete (dual male and female) sterility

TABLE 1 | Examples of Engineered Sterility in Forest Trees.

Species	Sterility target	Strategy	Candidate gene(s)	Reference
Angiosperm trees				
Betula pendula	Male	Cell ablation inflorescences	*BpFULL1::barnase*	Lännenpää et al., 2005
Eucalyptus occidentalis	Male	Cell ablation male flowers	*PrMC2::barnaseH102E*	Zhang et al., 2012
Populus tremula x tremuloides	Male	Cell ablation tapetum	*TA29:: barnase*	Elorriaga et al., 2014
Populus alba	Female	Suppression via RNAi	*PtLFY*	Klocko et al., 2016a
Populus alba	Female	Suppression via RNAi	*PtAG*	Lu and Strauss, 2018
Populus tremula x tremuloides	Male	Suppression via RNAi	*LFY/AG*	Klocko et al., 2018
Gymnosperm trees				
Pinus rigida x. P. taeda	Male	Cell ablation male cones	*PrMC2-barnaseH102E*	Zhang et al., 2012

has also been achieved using either independent male- and female-specific promoters or a single promoter targeting both tissues simultaneously (Liu and Liu, 2008; Huang et al., 2016). Male sterility using this cell ablation strategy has been demonstrated in both hardwood and softwood trees via expression of the *BARNASE* gene from *Bacillus amyloliquefaciens* under the control of reproductive tissue-specific promoters (**Table 1**). A key requirement for a cell ablation strategy is a promoter that tightly directs expression of the cytotoxin to the desired reproductive tissue to prevent pleiotropic effects on non-reproductive tissues. The conservation of expression of some floral genes has facilitated the use of a number of well characterized promoters across species (Strauss et al., 1995). Indeed, an anther-specific promoter derived from *Pinus radiata* has been used to express the *BARNASE* gene in both a softwood (pine) and hardwood (Eucalyptus) tree to deliver male sterility (Zhang et al., 2012). We are unaware of dual male/female *BARNASE*-mediated sterility being demonstrated in trees without negative pleiotropic effects (Lemmetyinen et al., 2004) but this should be possible if a suitable promoter is used.

RNA interference (RNAi) is a well proven homology-dependent gene silencing technology that involves double-stranded RNA directed against a target gene or its promoter region (Mansoor et al., 2006). Numerous demonstrations of engineered sterility through the suppression of genes essential for normal reproduction are available in angiosperm species (Wang et al., 2012). RNAi silencing has been used in angiosperm trees to engineer sterility with constructs targeting *LFY* and *AG* successfully producing sterile trees (**Table 1**). The production of male and female sterile plants via the use of chimeric repressors targeting transcription factors involved in flower development has also been demonstrated (Mitsuda et al., 2006; Katahata et al., 2014). In conifers the use of gene suppression methods to prevent reproduction has not been demonstrated even though such methods have been widely used to investigate wood quality traits (Wagner et al., 2005, 2009; Souza et al., 2007; Trontin et al., 2007). Attempts have been reported of expressing conifer flowering-associated genes in an endogenous system (Karlgren et al., 2013) but these studies have not directly sought to address sterility.

This lack of success in conifers reflects both a lack of fundamental knowledge regarding conifer reproduction and the inherent difficulties in working with conifers including the long timescale required for testing. For example, attempts to investigate the effects of over-expressing the Arabidopsis *LFY* gene in *P. radiata* were not informative as neither modified or control plants initiated reproduction during the 8 years that the trees were grown (NZ-EPA, 2008; Lottmann et al., 2010).

FUTURE OUTLOOK AND CHALLENGES

The social, legal and ecological impacts of sterile trees is still controversial (Williams, 2005; Kazana et al., 2015; Strauss et al., 2017). Although sterility provides mitigation for some of the social and ecological objections to the deployment of both GE trees and species with the potential to become invasive, this may be challenged if the sterility technology is itself GE.

The recent development of a number of new breeding technologies, including gene editing, that are already seeing widespread application in crop species (Nekrasov et al., 2017; Waltz, 2018) have great potential in forest trees. Site-directed mutagenesis would allow the inactivation of genes that are essential for normal reproductive processes and the generation of sterile trees. Gene editing-mediated mutagenesis would be particularly advantageous in forest trees where, to date, mutagenesis breeding has played an extremely limited role. The permanent inactivation of a gene would provide assurance of enduring containment and reduce concerns associated with the stability of long-term transgene expression associated with silencing of over-expression technologies (Li et al., 2008). Site directed mutagenesis via CRISPR-cas9 has been demonstrated in a number of tree species including Poplar (Fan et al., 2015) where mutagenesis of genes involved in flowering (Elorriaga et al., 2018) has also been shown. To date, gene editing has not been published in conifer species. However, the existence of a small number of natural spontaneous sterile conifer mutants (Orr-Ewing, 1977; Wilson and Owens, 2003; Rudall et al., 2011) suggest that a targeted-mutagenesis strategy would be successful if suitable targets can be identified.

Although the regulatory landscape regarding gene editing technologies remains complex, it is likely that in many jurisdictions versions of the technology that do not include foreign DNA in the final organisms will not be regulated as GMOs (Waltz, 2016; Davison and Ammann, 2017;

Ishii and Araki, 2017). This would provide a more straightforward and less costly route to commercial release than is currently the case for products of GE technology (Waltz, 2018). This regulatory approach would hold particular promise in applications where sterility is a standalone trait, such as for the control of invasive tree species, rather than providing a means of containment for other (GE) traits. This strategy would require DNA-free editing technologies as outcrossing of transgenes would not be possible with sterile trees.

The second major challenge has been the inability to carry out timely prototyping of sterility constructs in commercially important species. To facilitate testing in conifers it is desirable to develop a system analogous to the Poplar model system (Jansson and Douglas, 2007; Douglas, 2017) which has allowed relatively rapid prototyping of sterility constructs (Klocko et al., 2018). Although effective transformation systems exist for a number of commercially important conifers including *P. radiata*, *P. taeda,* and *Picea abies*, these species have long pre-reproductive juvenile growth periods that limits their use as sterility-testing platforms (Tang and Newton, 2003; Uddenberg et al., 2015). Some conifer species are able to reproduce at a much younger age (Righter, 1939; Pharis et al., 1987; Uddenberg et al., 2013) or can be induced to undergo early reproduction. Such precocious reproduction has been demonstrated in both hardwoods and softwoods by grafting onto older rootstock (Simak, 1978; Zhang et al., 2012), and by the application of external stimuli such as hormone treatments (Pharis et al., 1965; Ross and Pharis, 1985; Meilan, 1997). Stable introduction of *FT* transgenes induced precocious fertile flowers in Eucalyptus (Klocko et al., 2016b) and in Populus when combined with a low temperature treatment (Hoenicka et al., 2016a). In fruit trees, viral vectors that express floral promoters or silence repressors induced early flowering (Velázquez et al., 2016; Yamagishi et al., 2016). Although these offer potential routes to earlier testing of sterility strategies, developing the required tissue culture and transformation capabilities for a new tree species remains a significant barrier.

The increasing availability of genome and transcriptome resources for forest trees is providing new insights into reproductive processes. This is reducing the reliance on non-tree model systems and providing novel species-specific knowledge of reproductive processes and candidate genes for modification. The development of gene-editing-based targeted mutagenesis is likely to be the most attractive route to engineered sterility as it offers precise and predictable modifications combined with assurance of phenotypic stability. The lack of global consensus on the regulation of gene editing technology remains a barrier to research investment and commercialization and complicates the public debate that must go hand-in-hand with progress toward implementation.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

REFERENCES

Abe, M., Kobayashi, Y., Yamamoto, S., Daimon, Y., Yamaguchi, A., Ikeda, Y., et al. (2005). FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. *Science* 309, 1052–1056. doi: 10.1126/science.1115983

Al-Ahmad, H. (2018). Biotechnology for bioenergy dedicated trees: meeting future energy demands. *Z. Naturforsch.* 73, 15–32. doi: 10.1515/znc-2016-0185

Ault, K., Viswanath, V., Jayawickrama, J., Ma, C., Eaton, J., Meilan, R., et al. (2016). Improved growth and weed control of glyphosate-tolerant poplars. *New For.* 47, 653–667. doi: 10.1007/s11056-016-9536-6

Azeez, A., Miskolczi, P., Tylewicz, S., and Bhalerao, R. P. (2014). A tree ortholog of APETALA1 mediates photoperiodic control of seasonal growth. *Curr. Biol.* 24, 717–724. doi: 10.1016/j.cub.2014.02.037

Beals, T. P., and Goldberg, R. B. (1997). A novel cell ablation strategy blocks tobacco anther dehiscence. *Plant Cell* 9, 1527–1545. doi: 10.1105/tpc.9.9.1527

Becker, A. (2003). The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Mol. Phylogenet. Evol.* 29, 464–489. doi: 10.1016/s1055-7903(03)00207-0

Bielenberg, D. G., Wang, Y., Li, Z. G., Zhebentyayeva, T., Fan, S. H., Reighard, G. L., et al. (2008). Sequencing and annotation of the evergrowing locus in peach [*Prunus persica* (L.) Batsch] reveals a cluster of six MADS-box transcription factors as candidate genes for regulation of terminal bud formation. *Tree Genet. Genomes* 4, 495–507. doi: 10.1007/s11295-007-0126-9

Bohlenius, H., Huang, T., Charbonnel-Campaa, L., Brunner, A. M., Jansson, S., Strauss, S. H., et al. (2006). CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. *Science* 312, 1040–1043. doi: 10.1126/science.1126038

Breton, C., Guerin, J., Ducatillion, C., Medail, F., Kull, C. A., and Berville, A. (2008). Taming the wild and 'wilding'the tame: tree breeding and dispersal in Australia and the Mediterranean. *Plant Sci.* 175, 197–205. doi: 10.1016/j.plantsci.2008.05.013

Brunner, A. M., Li, J., DiFazio, S. P., Shevchenko, O., Montgomery, B. E., Mohamed, R., et al. (2007). Genetic containment of forest plantations. *Tree Genet. Genomes* 3, 75–100. doi: 10.1007/s11295-006-0067-8

Brunner, A. M., Varkonyi-Gasic, E., and Jones, R. C. (2017). "Phase change and phenology in trees," in *Comparative and Evolutionary Genomics of Angiosperm Trees*, eds A. Groover and Q. Cronk (Cham: Springer), 227–274. doi: 10.1007/7397_2016_30

Buters, J. T. M., Thibaudon, M., Smith, M., Kennedy, R., Rantio-Lehtimäki, A., Albertini, R., et al. (2012). Release of Bet v 1 from birch pollen from 5 European countries. Results from the HIALINE study. *Atmos. Environ.* 55, 496–505. doi: 10.1016/j.atmosenv.2012.01.054

Carlsbecker, A., Sundström, J., Tandre, K., Englund, M., Kvarnheden, A., Johanson, U., et al. (2003). The DAL10 gene from norway spruce (*Picea abies*) belongs to a potentially gymnosperm-specific subclass of MADS-box genes and is specifically active in seed cones and pollen cones. *Evol. Dev.* 5, 551–561. doi: 10.1046/j.1525-142X.2003.03060.x

Carlsbecker, A., Tandre, K., Johanson, U., Englund, M., and Engstrom, P. (2004). The MADS-box gene DAL1 is a potential mediator of the juvenile-to-adult transition in Norway spruce (*Picea abies*). *Plant J.* 40, 546–557. doi: 10.1111/j.1365-313X.2004.02226.x

Chen, F., Zhang, X., Liu, X., and Zhang, L. (2017). Evolutionary analysis of MIKC(c)-type MADS-box genes in gymnosperms and angiosperms. *Front. Plant Sci.* 8:895. doi: 10.3389/fpls.2017.00895

Cremer, K. (1992). Relations between reproductive growth and vegetative growth of Pinus radiata. *For. Ecol. Manage.* 52, 179–199. doi: 10.1016/0378-1127(92)90501-Y

D'amato, G., Cecchi, L., Bonini, S., Nunes, C., Annesi-Maesano, I., Behrendt, H., et al. (2007). Allergenic pollen and pollen allergy in Europe. *Allergy* 62, 976–990. doi: 10.1111/j.1398-9995.2007.01393.x

Davison, J., and Ammann, K. (2017). New GMO regulations for old: determining a new future for EU crop biotechnology. *GM Crops Food* 8, 13–34. doi: 10.1080/21645698.2017.1289305

De Block, M., Debrouwer, D., and Moens, T. (1997). The development of a nuclear male sterility system in wheat. Expression of the barnase gene under the control of tapetum specific promoters. *Theor. Appl. Genet.* 95, 125–131. doi: 10.1007/s001220050540

DiFazio, S. P., Slavov, G. T., Burczyk, J., Leonardi, S., and Strauss, S. H. (2004). "Gene flow from tree plantations and implications for transgenic risk assessment," in *Plantation Forest Biotechnology for the 21st Century*, eds C. Walter and M. Carson (Trivandrum: Research Signpost), 405–422.

Douglas, C. J. (2017). "*Populus* as a model tree," in *Comparative and Evolutionary Genomics of Angiosperm Trees*, eds A. Groover and Q. Cronk (Berlin: Springer International publishing), 61–84. doi: 10.1007/7397_2016_12

Dreni, L., and Kater, M. M. (2014). MADS reloaded: evolution of the AGAMOUS subfamily genes. *New Phytol.* 201, 717–732. doi: 10.1111/nph.12555

Elorriaga, E., Klocko, A. L., Ma, C., and Strauss, S. H. (2018). Variation in mutation spectra among CRISPR/Cas9 mutagenized poplars. *Front. Plant Sci.* 9:594. doi: 10.3389/fpls.2018.00594

Elorriaga, E., Meilan, R., Ma, C., Skinner, J. S., Etherington, E., Brunner, A., et al. (2014). A tapetal ablation transgene induces stable male sterility and slows field growth in Populus. *Tree Genet. Genomes* 10, 1583–1593. doi: 10.1007/s11295-014-0781-6

Etchells, J. P., Mishra, L. S., Kumar, M., Campbell, L., and Turner, S. R. (2015). Wood formation in trees is increased by manipulating PXY-regulated cell division. *Curr. Biol.* 25, 1050–1055. doi: 10.1016/j.cub.2015.02.023

Fan, D., Liu, T., Li, C., Jiao, B., Li, S., Hou, Y., et al. (2015). Efficient CRISPR/Cas9-mediated targeted mutagenesis in Populus in the first generation. *Sci. Rep.* 5:12217. doi: 10.1038/srep12217

FAO (2015). *Global Forest Resources Assessement 2015: How are the Worlds Forests Changing?* Rome: Food and Agricultural Organization of the United Nations. Available at: http://www.fao.org/3/a-i4793e.pdf [accessed May 2018 2015].

Franzese, J., and Raffaele, E. (2017). Fire as a driver of pine invasions in the Southern Hemisphere: a review. *Biol. Invasions* 19, 2237–2246. doi: 10.1007/s10530-017-1435-z

Froude, V. A. (2011). *Wilding Conifers in New Zealand: Beyond the Status Report.* Bay of Islands: Report prepared for the Ministry of Agriculture and Forestry, Pacific Eco-Logic, 44.

Fukui, M., Futamura, N., Mukai, Y., Wang, Y., Nagao, A., and Shinohara, K. (2001). Ancestral MADS box genes in sugi, *Cryptomeria japonica* D. Don (Taxodiaceae), homologous to the B function genes in angiosperms. *Plant Cell Physiol.* 42, 566–575. doi: 10.1093/pcp/pce069

Goldman, M., Goldberg, R., and Mariani, C. (1994). Female sterile tobacco plants are produced by stigma-specific cell ablation. *EMBO J.* 13, 2976–2984. doi: 10.1002/j.1460-2075.1994.tb06596.x

Gramzow, L., Weilandt, L., and Theissen, G. (2014). MADS goes genomic in conifers: towards determining the ancestral set of MADS-box genes in seed plants. *Ann. Bot.* 114, 1407–1429. doi: 10.1093/aob/mcu066

Harfouche, A., Meilan, R., and Altman, A. (2011). Tree genetic engineering and applications to sustainable forestry and biomass production. *Trends Biotechnol.* 29, 9–17. doi: 10.1016/j.tibtech.2010.09.003

Hoenicka, H., Lautner, S., Klingberg, A., Koch, G., El-Sherif, F., Lehnhardt, D., et al. (2012). Influence of over-expression of the FLOWERING PROMOTING FACTOR 1 gene (FPF1) from *Arabidopsis* on wood formation in hybrid poplar (*Populus tremula* L. × *P. tremuloides* Michx.). *Planta* 235, 359–373. doi: 10.1007/s00425-011-1507-8

Hoenicka, H., Lehnhardt, D., Briones, V., Nilsson, O., and Fladung, M. (2016a). Low temperatures are required to induce the development of fertile flowers in transgenic male and female early flowering poplar (*Populus tremula* L.). *Tree Physiol.* 36, 667–677. doi: 10.1093/treephys/tpw015

Hoenicka, H., Lehnhardt, D., Nunna, S., Reinhardt, R., Jeltsch, A., Briones, V., et al. (2016b). Level of tissue differentiation influences the activation of a heat-inducible flower-specific system for genetic containment in poplar (*Populus tremula* L.). *Plant Cell Rep.* 35, 369–384. doi: 10.1007/s00299-015-1890-x

Hoenicka, H., Nowitzki, O., Hanelt, D., and Fladung, M. (2008). Heterologous overexpression of the birch FRUITFULL-like MADS-box gene BpMADS4 prevents normal senescence and winter dormancy in *Populus tremula* L. *Planta* 227, 1001–1011. doi: 10.1007/s00425-007-0674-0

Hsu, C.-Y., Adams, J. P., Kim, H., No, K., Ma, C., Strauss, S. H., et al. (2011). FLOWERING LOCUS T duplication coordinates reproductive and vegetative growth in perennial poplar. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10756–10761. doi: 10.1073/pnas.1104713108

Huang, J., Smith, A. R., Zhang, T., and Zhao, D. (2016). Creating completely both male and female sterile plants by specifically ablating microspore and megaspore mother cells. *Front. Plant Sci.* 7:30. doi: 10.3389/fpls.2016.00030

Ishii, T., and Araki, M. (2017). A future scenario of the global regulatory landscape regarding genome-edited crops. *GM Crops Food* 8, 44–56. doi: 10.1080/21645698.2016.1261787

Jaeger, K. E., Pullen, N., Lamzin, S., Morris, R. J., and Wigge, P. A. (2013). Interlocking feedback loops govern the dynamic behavior of the floral transition in *Arabidopsis*. *Plant Cell* 25, 820–833. doi: 10.1105/tpc.113.109355

Jansson, S., and Douglas, C. J. (2007). Populus: a model system for plant biology. *Annu. Rev. Plant Biol.* 58, 435–458. doi: 10.1146/annurev.arplant.58.032806.103956

Karlgren, A., Gyllenstrand, N., Clapham, D., and Lagercrantz, U. (2013). FLOWERING LOCUS T/TERMINAL FLOWER1-like genes affect growth rhythm and bud set in Norway spruce. *Plant Physiol.* 163, 792–803. doi: 10.1104/pp.113.224139

Karlgren, A., Gyllenstrand, N., Källman, T., Sundström, J. F., Moore, D., Lascoux, M., et al. (2011). Evolution of the PEBP gene family in plants: functional diversification in seed plant evolution. *Plant Physiol.* 156, 1967–1977. doi: 10.1104/pp.111.176206

Katahata, S.-I., Futamura, N., Igasaki, T., and Shinohara, K. (2014). Functional analysis of SOC1-like and AGL6-like MADS-box genes of the gymnosperm *Cryptomeria japonica*. *Tree Genet. Genomes* 10, 317–327. doi: 10.1007/s11295-013-0686-9

Kazana, V., Tsourgiannis, L., Iakovoglou, V., Stamatiou, C., Alexandrov, A., Araújo, S., et al. (2015). Public attitudes towards the use of transgenic forest trees: a cross-country pilot survey. *iForest-Biogeosci. For.* 9:344.

Klintenas, M., Pin, P. A., Benlloch, R., Ingvarsson, P. K., and Nilsson, O. (2012). Analysis of conifer FLOWERING LOCUS T/TERMINAL FLOWER1-like genes provides evidence for dramatic biochemical evolution in the angiosperm FT lineage. *New Phytol.* 196, 1260–1273. doi: 10.1111/j.1469-8137.2012.04332.x

Klocko, A. L., Brunner, A. M., Huang, J., Meilan, R., Lu, H., Ma, C., et al. (2016a). Containment of transgenic trees by suppression of LEAFY. *Nat. Biotechnol.* 34:918. doi: 10.1038/nbt.3636

Klocko, A. L., Ma, C., Robertson, S., Esfandiari, E., Nilsson, O., and Strauss, S. H. (2016b). FT overexpression induces precocious flowering and normal reproductive development in Eucalyptus. *Plant Biotechnol. J.* 14, 808–819. doi: 10.1111/pbi.12431

Klocko, A. L., Lu, H., Magnuson, A., Brunner, A. M., Ma, C., and Strauss, S. H. (2018). Phenotypic expression and stability in a large-scale field study of genetically engineered poplars containing sexual containment transgenes. *Front. Bioeng. Biotechnol.* 6:100. doi: 10.3389/fbioe.2018.00100

Kramer, R. D., Sillett, S. C., and Carroll, A. L. (2014). Structural development of redwood branches and its effects on wood growth. *Tree Physiol.* 34, 314–330. doi: 10.1093/treephys/tpu011

Lännenpää, M., Hassinen, M., Ranki, A., Hölttä-Vuori, M., Lemmetyinen, J., Keinonen, K., et al. (2005). Prevention of flower development in birch and other plants using a BpFULL1:: BARNASE construct. *Plant Cell Rep.* 24, 69–78. doi: 10.1007/s00299-004-0903-y

Lemmetyinen, J., Keinonen, K., and Sopanen, T. (2004). Prevention of the flowering of a tree, silver birch. *Mol. Breed.* 13, 243–249. doi: 10.1023/b:Molb.0000022525.96200.53

Li, J., Brunner, A. M., Shevchenko, O., Meilan, R., Ma, C., Skinner, J. S., et al. (2008). Efficient and stable transgene suppression via RNAi in field-grown poplars. *Transgen. Res.* 17, 679–694. doi: 10.1007/s11248-007-9148-1

Liu, Y. Y., Yang, K. Z., Wei, X. X., and Wang, X. Q. (2016). Revisiting the phosphatidylethanolamine-binding protein (PEBP) gene family reveals cryptic FLOWERING LOCUS T gene homologs in gymnosperms and sheds new light on functional evolution. *New Phytol.* 212, 730–744. doi: 10.1111/nph.14066

Liu, Z., and Liu, Z. (2008). The second intron of AGAMOUS drives carpel-and stamen-specific expression sufficient to induce complete sterility in *Arabidopsis*. *Plant Cell Rep.* 27, 855–863. doi: 10.1007/s00299-008-0511-3

Liu, Z. -X., Xiong, H. -Y., Li, L. -Y., and Fei, Y. -J. (2018). Functional conservation of an AGAMOUS orthologous gene controlling reproductive organ development in the gymnosperm species *Taxus chinensis* var. mairei. *J. Plant Biol.* 61, 50–59. doi: 10.1007/s12374-017-0154-4

Lottmann, J., O'Callaghan, M., Baird, D., and Walter, C. (2010). Bacterial and fungal communities in the rhizosphere of field-grown genetically modified pine trees (*Pinus radiata* D.). *Environ. Biosaf. Res.* 9, 25–40. doi: 10.1051/ebr/2010007

Lu, H., and Strauss, S. H. (2018). "Cross-suppression of AG and AG-LIKE 11 genes gives sterility in field grown poplar," in *Proceedings of PAG XXVI*, San Diego.

Luis, S. D. B., and José, C. (2014). Costs of female reproduction in a conifer tree: a whole-tree level assessment. *J. Ecol.* 102, 1310–1317. doi: 10.1111/1365-2745.12283

Mansoor, S., Amin, I., Hussain, M., Zafar, Y., and Briddon, R. W. (2006). Engineering novel traits in plants through RNA interference. *Trends Plant Sci.* 11, 559–565. doi: 10.1016/j.tplants.2006.09.010

Mariani, C., De Beuckeleer, M., Truettner, J., Leemans, J., and Goldberg, R. B. (1990). Induction of male sterility in plants by a chimaeric ribonuclease gene. *Nature* 347:737. doi: 10.1038/347737a0

Meilan, R. (1997). Floral induction in woody angiosperms. *New For.* 14, 179–202. doi: 10.1023/A:1006560603966

Melzer, R., Wang, Y. Q., and Theissen, G. (2010). The naked and the dead: the ABCs of gymnosperm reproduction and the origin of the angiosperm flower. *Semin. Cell Dev. Biol.* 21, 118–128. doi: 10.1016/j.semcdb.2009.11.015

Mitsuda, N., Hiratsu, K., Todaka, D., Nakashima, K., Yamaguchi-Shinozaki, K., and Ohme-Takagi, M. (2006). Efficient production of male and female sterile plants by expression of a chimeric repressor in *Arabidopsis* and rice. *Plant Biotechnol. J.* 4, 325–332. doi: 10.1111/j.1467-7652.2006.00184.x

Mohamed, R., Wang, C. T., Ma, C., Shevchenko, O., Dye, S. J., Puzey, J. R., et al. (2010). Populus CEN/TFL1 regulates first onset of flowering, axillary meristem identity and dormancy release in Populus. *Plant J.* 62, 674–688. doi: 10.1111/j.1365-313X.2010.04185.x

Mouradov, A., Glassick, T. V., Hamdorf, B. A., Murphy, L. C., Marla, S. S., Yang, Y., et al. (1998). Family of MADS-Box genes expressed early in male and female reproductive structures of monterey pine. *Plant Physiol.* 117, 55–62. doi: 10.1104/pp.117.1.55

Nature Biotechnology News (2015). Brazil approves transgenic eucalyptus. *Nat. Biotechnol.* 33:577. doi: 10.1038/nbt0615-577c

Nekrasov, V., Wang, C., Win, J., Lanz, C., Weigel, D., and Kamoun, S. (2017). Rapid generation of a transgene-free powdery mildew resistant tomato by genome deletion. *Sci. Rep.* 7:482. doi: 10.1038/s41598-017-00578-x

Nilsson, L., Carlsbecker, A., Sundas-Larsson, A., and Vahala, T. (2007). APETALA2 like genes from *Picea abies* show functional similarities to their *Arabidopsis* homologues. *Planta* 225, 589–602. doi: 10.1007/s00425-006-03741

Nuñez, M. A., Chiuffo, M. C., Torres, A., Paul, T., Dimarco, R. D., Raal, P., et al. (2017). Ecology and management of invasive Pinaceae around the world: progress and challenges. *Biol. Invasions* 19, 3099–3120. doi: 10.1007/s10530-017-1483-4

Nystedt, B., Street, N. R., Wetterbom, A., Zuccolo, A., Lin, Y. C., Scofield, D. G., et al. (2013). The Norway spruce genome sequence and conifer genome evolution. *Nature* 497, 579–584. doi: 10.1038/nature12211

NZ-EPA (2008). *Annual Report for Field Test GMF99001*. Available at: https://www.epa.govt.nz/assets/FileAPI/hsno-ar/GMF99001/GMF99001-2008-GMF99001-5-annual-report.pdf [accessed 14/08/2018].

Obeso, J. R. (2002). The costs of reproduction in plants. *New Phytol.* 155, 321–348. doi: 10.1046/j.1469-8137.2002.00477.x

Orr-Ewing, A. L. (1977). Female sterility in Douglas fir. *Silvae Genet.* 26, 2–3.

Pajoro, A., Biewers, S., Dougali, E., Leal Valentim, F., Mendes, M. A., Porri, A., et al. (2014). The (r)evolution of gene regulatory networks controlling *Arabidopsis* plant reproduction: a two-decade history. *J. Exp. Bot.* 65, 4731–4745. doi: 10.1093/jxb/eru233

Palmiter, R. D., Behringer, R. R., Quaife, C. J., Maxwell, F., Maxwell, I. H., and Brinster, R. L. (1987). Cell lineage ablation in transgenic mice by cell-specific expression of a toxin gene. *Cell* 50, 435–443. doi: 10.1016/0092-8674(87)90497-1

Parmentier-Line, C. M., and Coleman, G. D. (2016). Constitutive expression of the Poplar FD-like basic leucine zipper transcription factor alters growth and bud development. *Plant Biotechnol. J.* 14, 260–270. doi: 10.1111/pbi.12380

Pharis, R. P., Ruddat, M. D., Phillips, C. C., and Heftmann, E. (1965). Precocious flowering of Arizona cypress with gibberellin. *Can. J. Bot.* 43, 923–927. doi: 10.1139/b65-102

Pharis, R. P., Webber, J. E., and Ross, S. D. (1987). The promotion of flowering in forest trees by gibberellin A47 and cultural treatments: a review of the possible mechanisms. *For. Ecol. Manage.* 19, 65–84. doi: 10.1016/0378-1127(87)90012-0

Porth, I., and El-Kassaby, Y. A. (2014). Current status of the development of genetically modified (GM) forest trees world-wide: a comparison with the development of other GM plants in agriculture. *CAB Rev.* 9:12. doi: 10.1079/PAVSNNR20149008

Richardson, D. M., and Rejmánek, M. (2004). Conifers as invasive aliens: a global survey and predictive framework. *Divers. Distribut.* 10, 321–331. doi: 10.1111/j.1366-9516.2004.00096.x

Righter, F. (1939). Early flower production among the pines. *J. For.* 37, 935–938.

Ross, S., and Pharis, R. (1985). *Promotion of Flowering in Tree Crops: Different Mechanisms and Techniques, with Special Reference to Conifers*. Dorset: Institute of Terrestrial Ecology.

Rudall, P. J., Hilton, J., Vergara-Silva, F., and Bateman, R. M. (2011). Recurrent abnormalities in conifer cones and the evolutionary origins of flower-like structures. *Trends Plant Sci.* 16, 151–159. doi: 10.1016/j.tplants.2010.11.002

Rutledge, R., Regan, S., Nicolas, O., Fobert, P., Côté, C., Bosnich, W., et al. (1998). Characterization of an AGAMOUS homologue from the conifer black spruce (*Picea mariana*) that produces floral homeotic conversions when expressed in *Arabidopsis*. *Plant J.* 15, 625–634. doi: 10.1046/j.1365-313x.1998.00250.x

Sala, A., Hopping, K., McIntire, E. J., Delzon, S., and Crone, E. E. (2012). Masting in whitebark pine (*Pinus albicaulis*) depletes stored nutrients. *New Phytol.* 196, 189–199. doi: 10.1111/j.1469-8137.2012.04257.x

Shindo, S., Sakakibara, K., Sano, R., Ueda, K., and Hasebe, M. (2001). Characterization of a FLORICAULA/LEAFY homologue of *Gnetum parvifolium* and its implications for the evolution of reproductive organs in seed plants. *Int. J. Plant Sci.* 162, 1199–1209. doi: 10.1086/323417

Shiokawa, T., Yamada, S., Futamura, N., Osanai, K., Murasugi, D., Shinohara, K., et al. (2008). Isolation and functional analysis of the CjNdly gene, a homolog in *Cryptomeria japonica* of FLORICAULA/LEAFY genes. *Tree Physiol.* 28, 21–28. doi: 10.1093/treephys/28.1.21

Simak, M. (1978). "Precocius flower induction in Pinus sylvestris by grafting," *Proceedings of the Symposium on Flowering and Seed Development in Trees*, Starkville, 63–76.

Simberloff, D., Nunez, M. A., Ledgard, N. J., Pauchard, A., Richardson, D. M., Sarasola, M., et al. (2010). Spread and impact of introduced conifers in South America: lessons from other southern hemisphere regions. *Austral Ecol.* 35, 489–504. doi: 10.1111/j.1442-9993.2009.02058.x

Souza, A. J. D., Mendes, B. M. J., and Mourão Filho, F. D. A. A. (2007). Gene silencing: concepts, applications, and perspectives in woody plants. *Sci. Agric.* 64, 645–656. doi: 10.1590/S0103-90162007000600014

Strauss, S. H., Jones, K. N., Lu, H., Petit, J. D., Klocko, A. L., Betts, M. G., et al. (2017). Reproductive modification in forest plantations: impacts on biodiversity and society. *New Phytol.* 213, 1000–1021. doi: 10.1111/nph.14374

Strauss, S. H., Ma, C., Ault, K., and Klocko, A. L. (2016). "Lessons from two decades of field trials with genetically modified trees in the USA: biology and regulatory compliance," in *Biosafety of Forest Transgenic Trees*, eds C. Vettori, F. Gallardo, H. Häggman, V. Kazana, F. Migliacci, G. Pilate, et al. (Dordrecht: Springer), 101–124.

Strauss, S. H., Rottmann, W. H., Brunner, A. M., and Sheppard, L. A. (1995). Genetic engineering of reproductive sterility in forest trees. *Mol. Breed.* 1, 5–26. doi: 10.1007/BF01682086

Sundström, J., and Engström, P. (2002). Conifer reproductive development involves B-type MADS-box genes with distinct and different activities in male organ primordia. *Plant J.* 31, 161–169. doi: 10.1046/j.1365-313X.2002.01343.x

Tandre, K., Albert, V. A., Sundås, A., and Engström, P. (1995). Conifer homologues to genes that control floral development in angiosperms. *Plant Mol. Biol.* 27, 69–78. doi: 10.1007/BF00019179

Tandre, K., Svenson, M., Svensson, M. E., and Engström, P. (1998). Conservation of gene structure and activity in the regulation of reproductive organ development of conifers and angiosperms. *Plant J.* 15, 615–623. doi: 10.1046/j.1365-313x.1998.00236.x

Tang, W., and Newton, R. J. (2003). Genetic transformation of conifers and its application in forest biotechnology. *Plant Cell Rep.* 22, 1–15. doi: 10.1007/s00299-003-0670-1

Taniguchi, T. (2018). "Biotechnological research in *Cryptomeria japonica*," in *Biofuels: Greenhouse Gas Mitigation and Global Warming: Next Generation Biofuels and Role of Biotechnology*, eds A. Kumar, S. Ogita and Y.-Y. Yau (New Delhi: Springer), 357–374. doi: 10.1007/978-81-322-3763-1_20

Trontin, J.-F., Walter, C., Klimaszewska, K., Park, Y.-S., and Lelu-Walter, M.-A. (2007). Recent progress in genetic transformation of four Pinus spp. *Trans. Plant J.* 1, 314–329.

Tylewicz, S., Tsuji, H., Miskolczi, P., Petterle, A., Azeez, A., Jonsson, K., et al. (2015). Dual role of tree florigen activation complex component FD in photoperiodic growth control and adaptive response pathways. *Proc. Natl. Acad. Sci. U.S.A.* 112, 3140–3145. doi: 10.1073/pnas.1423440112

Uddenberg, D., Akhter, S., Ramachandran, P., Sundström, J. F., and Carlsbecker, A. (2015). Sequenced genomes and rapidly emerging technologies pave the way for conifer evolutionary developmental biology. *Front. Plant Sci.* 6:970. doi: 10.3389/fpls.2015.00970

Uddenberg, D., Reimegard, J., Clapham, D., Almqvist, C., von Arnold, S., Emanuelsson, O., et al. (2013). Early cone setting in *Picea abies* acrocona is associated with increased transcriptional activity of a MADS box transcription factor. *Plant Physiol.* 161, 813–823. doi: 10.1104/pp.112.207746

Velázquez, K., Agüero, J., Vives, M. C., Aleza, P., Pina, J. A., Moreno, P., et al. (2016). Precocious flowering of juvenile citrus induced by a viral vector based on Citrus leaf blotch virus: a new tool for genetics and breeding. *Plant Biotechnol. J.* 14, 1976–1985. doi: 10.1111/pbi.12555

Vining, K. J., Romanel, E., Jones, R. C., Klocko, A., Alves-Ferreira, M., Hefer, C. A., et al. (2015). The floral transcriptome of *Eucalyptus grandis*. *New Phytol.* 206, 1406–1422. doi: 10.1111/nph.13077

Wagner, A., Donaldson, L., Kim, H., Phillips, L., Flint, H., Steward, D., et al. (2009). Suppression of 4-coumarate-CoA ligase in the coniferous gymnosperm *Pinus radiata*. *Plant Physiol.* 149, 370–383. doi: 10.1104/pp.108.125765

Wagner, A., Phillips, L., Narayan, R. D., Moody, J. M., and Geddes, B. (2005). Gene silencing studies in the gymnosperm species *Pinus radiata*. *Plant Cell Rep.* 24, 95–102. doi: 10.1007/s00299-004-0911-y

Waltz, E. (2016). *CRISPR-Edited Crops Free to Enter Market, Skip Regulation.* London: Nature Publishing Group.

Waltz, E. (2018). With a free pass. CRISPR-edited plants reach market in record time. *Nat. Biotechnol.* 36:6. doi: 10.1038/nbt0118-6b

Wang, X., Singer, S. D., and Liu, Z. (2012). Silencing of meiosis-critical genes for engineering male sterility in plants. *Plant Cell Rep.* 31, 747–756. doi: 10.1007/s00299-011-1193-9

Williams, C. G. (2005). Framing the issues on transgenic forests. *Nat. Biotechnol.* 23:530. doi: 10.1038/nbt0505-530

Williams, C. G. (2009). *Conifer Reproductive Biology.* Berlin: Springer Science & Business Media. doi: 10.1007-978-1-4020-9602-0

Wilson, V. R., and Owens, J. N. (2003). Histology of sterile male and female cones in *Pinus monticola* (western white pine). *Sex. Plant Reproduct.* 15, 301–310.

WWF (2015). *WWF Living Forests Report: Chapter 5: Saving Forests at Risk.* Gland: WWF – World Wide Fund for Nature. Available at: http://awsassets.panda.org/downloads/living_forests_report_chapter_5_1.pdf [accessed May 2018].

Yamagishi, N., Li, C., and Yoshikawa, N. (2016). Promotion of flowering by apple latent spherical virus vector and virus elimination at high temperature allow accelerated breeding of apple and pear. *Front. Plant Sci.* 7:171. doi: 10.3389/fpls.2016.00171

Zahn, L. M., Kong, H., Leebens-Mack, J. H., Kim, S., Soltis, P. S., Landherr, L. L., et al. (2005). The evolution of the SEPALLATA subfamily of MADS-box genes: a preangiosperm origin with multiple duplications throughout angiosperm history. *Genetics* 169, 2209–2223. doi: 10.1534/genetics.104.037770

Zhang, C., Norris-Caneda, K. H., Rottmann, W. H., Gulledge, J. E., Chang, S., Kwan, B. Y.-H., et al. (2012). Control of pollen-mediated gene flow in transgenic trees. *Plant Physiol.* 159, 1319–1334. doi: 10.1104/pp.112.197228

Zhou, S., Runge, T., Karlen, S., Ralph, J., Gonzales-Vigil, E., and Mansfield, S. (2017). Chemical pulping advantages of ziplignin hybrid poplar. *ChemSusChem* 10, 3565–3573. doi: 10.1002/cssc.201701317

Genome-Wide Association Studies to Improve Wood Properties: Challenges and Prospects

Qingzhang Du[1,2,3], Wenjie Lu[1,2,3], Mingyang Quan[1,2,3], Liang Xiao[1,2,3], Fangyuan Song[1,2,3], Peng Li[1,2,3], Daling Zhou[1,2,3], Jianbo Xie[1,2,3], Longxin Wang[1,2,3] and Deqiang Zhang[1,2,3]*

[1] Beijing Advanced Innovation Center for Tree Breeding by Molecular Design, Beijing Forestry University, Beijing, China, [2] National Engineering Laboratory for Tree Breeding, College of Biological Sciences and Technology, Beijing Forestry University, Beijing, China, [3] Key Laboratory of Genetics and Breeding in Forest Trees and Ornamental Plants, Ministry of Education, College of Biological Sciences and Technology, Beijing Forestry University, Beijing, China

Wood formation is an excellent model system for quantitative trait analysis due to the strong associations between the transcriptional and metabolic traits that contribute to this complex process. Investigating the genetic architecture and regulatory mechanisms underlying wood formation will enhance our understanding of the quantitative genetics and genomics of complex phenotypic variation. Genome-wide association studies (GWASs) represent an ideal statistical strategy for dissecting the genetic basis of complex quantitative traits. However, elucidating the molecular mechanisms underlying many favorable loci that contribute to wood formation and optimizing GWAS design remain challenging in this omics era. In this review, we summarize the recent progress in GWAS-based functional genomics of wood property traits in major timber species such as *Eucalyptus*, *Populus*, and various coniferous species. We discuss several appropriate experimental designs for extensive GWAS in a given undomesticated tree population, such as omics-wide association studies and high-throughput phenotyping technologies. We also explain why more attention should be paid to rare allelic and major structural variation. Finally, we explore the potential use of GWAS for the molecular breeding of trees. Such studies will help provide an integrated understanding of complex quantitative traits and should enable the molecular design of new cultivars.

Keywords: GWAS, omics, functional genomics, wood formation, systems biology

INTRODUCTION

Wood, the secondary xylem of long-lived perennial plants, is produced via cell division from the vascular cambium, cell expansion, cell wall thickening, programmed cell death, and heartwood formation (Plomion et al., 2001; Mellerowicz and Sundberg, 2008). In general, the chemical and ultrastructural properties of wood depend on the components of the secondary cell walls, allowing wood to fulfill highly specialized functions that are essential for tree growth and development (Du et al., 2013a; Mizrachi and Myburg, 2016). Wood also represents a major carbon sink that plays a crucial role in carbon cycling in the terrestrial ecosystem, serving as an important renewable resource for the production of lumber, pulp, paper, and biofuels (Mellerowicz and Sundberg, 2008).

Progress has recently been made toward modifying the major wood biopolymers (i.e., lignin and cellulose) in several model plants (Robischon et al., 2011; Chen et al., 2014; Ye and Zhong, 2015), but much remains to be explored about the biosynthetic machinery of the chemistry and ultrastructure of wood cell walls. Most studies to date indicate that secondary wall biosynthesis needs the coordinated activity of transcriptional networks in regulating the diverse metabolic pathways involving polysaccharides and lignin biopolymers (Coleman et al., 2009; Mewalal et al., 2014). This intricate biological process incorporates a diverse set of xylem-forming genes, most with unknown functions (Zinkgraf et al., 2017). The genetic architecture and functional mechanisms that directly affect wood properties have not yet been fully identified and dissected.

Numerous studies have examined the underlying genetic variation in complex polysaccharides (alpha cellulose, hemicelluloses, and holocellulose), lignin (insoluble, soluble, syringyl, and total lignin), cell wall sugars (arabinose, glucose, mannose, and xylose), and ultrastructural traits (average density, crystallinity, fiber length, and microfiber angle) in *Eucalyptus, Populus*, and various coniferous species using forward genetic approaches, such as quantitative trait loci (QTLs) (Sewell et al., 2000; Novaes et al., 2009; Thumma et al., 2010; Yin et al., 2010) and candidate gene-based association mapping (Thumma et al., 2009; Dillon et al., 2010; Wegrzyn et al., 2010; Beaulieu et al., 2011; Du et al., 2013b; Guerra et al., 2013; Porth et al., 2013). However, the QTLs, single nucleotide polymorphism (SNP) loci, and candidate genes identified to date explain only a small proportion of the genetic variation of wood components. Genes cannot work in isolation; instead, multiple genes within complex biological pathways are often jointly involved in phenotypic variation (Du et al., 2015; Zinkgraf et al., 2017). Therefore, a more holistic research approach encompassing whole genome variation must be taken to understand and improve wood property traits.

Given the high genetic diversity, almost undomesticated status, rapid decay of linkage disequilibrium (LD), and minimal genetic structure of forest tree populations, such populations should represent ideal systems for conducting association studies and breeding using molecular marker-assisted selection (MAS). LD is fundamentally important for any genome-wide association study (GWAS) when genotyping does not cover all sequence variants in a genome. Indeed, the advantage of a low extent of LD is that once an association is detected, it is likely that this marker is physically close to the causal variant (likely within the gene itself) or is even the causal variant (Ingvarsson et al., 2016). Dissection of genome–phenotype associations through GWAS in a variety of systems is thus expected to efficiently bridge the gap between QTLs and causal genes in most forest trees (Porth et al., 2013), thanks to the currently available large populations and high-throughput sequencing technology. Here, we (1) summarize the recent progress in functional genomics of wood property traits via GWAS; (2) discuss the statistical methods and experimental designs needed to improve the use of GWAS in trees; and (3) explore opportunities for the use of GWAS in the molecular breeding of trees.

RECENT PROGRESS IN GWAS FOR EXPLORATION OF WOOD PROPERTY TRAITS

Trees have a wide geographical distribution and large wild population sizes and thus exhibit diverse responses to environmental changes. Association mapping principally exploits evolutionary recombination at the natural population level. Thus, a collection of cultivars/natural individuals with unknown ancestry and newly designed nested association mapping (NAM) populations are often used for association analysis in trees. The availability of fast, accurate estimation methods for variance components is a prerequisite for performing GWAS. Yu et al. (2006) proposed a mixed linear model (MLM) method for better controlling population structure and the imbalanced kinships among various individuals (Pritchard et al., 2000). Genome-wide rapid association analysis by mixed model and regression (GRAMMER) was subsequently developed to roughly estimate random effects. Unlike approximate estimation models, the efficient mixed-model association (EMMA) matrix is a method for more accurately estimating the genetic and residual variance of a population, and therefore, may speed up the calculation process (Kang et al., 2008). EMMA eXpedited (EMMAX) and Population Parameters Previously Determined (P3D) are two other estimation methods that reduce the need for computational processing (Kang et al., 2010; Zhang et al., 2010). Factored spectrally transformed linear mixed models (FaST-LMM) (Lippert et al., 2011) and genome-wide efficient mixed-model association (GEMMA) (Zhou and Stephens, 2012) methods were subsequently developed. The advantage of these approaches is that they allow variance components to be directly estimated. Meta-GWAS and Joint-GWAS are used for obtaining higher statistical power in analyses of multiple tree populations (Wu et al., 2016; Müller et al., 2018).

Next-generation sequencing (NGS) and SNP arrays have opened up new possibilities for obtaining almost all SSRs and SNPs variants within a gene space or even within an genome-wide scan (Evans et al., 2014; Ingvarsson et al., 2016). Moreover, NGS has enabled genome-wide discovery of structural variation, insertion/deletions (InDels), and copy number variants (CNV) (Marroni et al., 2014), in a growing number of whole-genome resequencing studies in several model tree species (Evans et al., 2014; Silva-Junior and Grattapaglia, 2015; Wang J. et al., 2015; Wang et al., 2016; Du et al., 2016). These GWAS have yielded important insights into the genetic basis of complex quantitative traits in woody species, primarily focusing on wood compositions and wood property traits (**Table 1**) (Cappa et al., 2013; Porth et al., 2013; McKown et al., 2014; Allwright et al., 2016; Lamara et al., 2016; Gong et al., 2017; Resende et al., 2017; Zinkgraf et al., 2017).

TABLE 1 | Summary of genome-wide association studies (GWAS) of wood property traits in main timber species.

Phenotype	Species	Population	Sample size	No. of markers	Method	Reference
Growth and wood properties	*Eucalyptus globulus*	Families and bulk collections	303	7,680 [Diversity Array Technology markers (DArT)]	General linear model (GLM) and unified mixed model (UMM)	Cappa et al., 2013
Wood density, stiffness, microfibril angle, and ring width	*Picea glauca*	Open-pollinated families	1694	7434 (SNPs)	Mixed linear model (MLM)	Lamara et al., 2016
16 wood chemistry/ ultrastructure traits	*Populus trichocarpa*	Unrelated individuals	334	29,233 (SNPs)	GLM	Porth et al., 2013
Lignin percentage, Lignin S:G ratio, 5-carbon sugars, and 6-carbon sugars	*Populus deltoides*	Unrelated individuals	391	334,679 (consensus SNPs), 185,526 (Common SNPs), 76,804 (functional SNPs)	Single-variant and multiple-variant associations on GLM	Fahrenkrog et al., 2017
Basic wood density (BWD), bleached pulp, pulp yield (SPY), and pulp bleaching content	*Eucalyptus grandis* × *Eucalyptus urophylla*	Hybrid breeding population	768	24 806 (SNPs)	GWAS and regional heritability mapping	Resende et al., 2017
17 wood-quality traits	Norway spruce	Mother trees	517	178101 (SNPs)	Multilocus LASSO penalized regression	Baison et al., 2018
Seven wood properties	*Populus tomentosa*	Unrelated individuals	435	5,482 (InDels)	MLM and Kempthorne model	Gong et al., 2017

Diversity Array Technology (DArT) markers.

Porth et al. (2013) performed the first GWAS for key wood chemistry and ultrastructure traits in a population of 334 unrelated black cottonwood (*Populus trichocarpa*) individuals and found 141 significant SNPs associated with cell wall traits. Only 40% of these associations involve genes previously known to function in wood formation (Wegrzyn et al., 2010). For example, a synonymous SNP within the *FRA8* ortholog can explain 21.0% of the total genetic variance of fiber length, suggesting the GWAS could provide insight into the genetic basis of wood traits. Further integration of datasets from transcription factor binding, transcriptome profiling, and GWAS experiments in *Populus* provided an effective means of annotating conserved gene co-expression modules under diverse environmental conditions (Zinkgraf et al., 2017). The combinatorial utility of QTL and chromosome-wide association mapping enabled the authors to identify six genes (encoding the transcription factors ANGUSTIFOLIA C-terminus Binding Protein [CtBP] and KANADI, a Ca^{2+}-transporting ATPase, an amino acid transporter, the copper transporter ATOX1-related, and a protein kinase) of functional relevance to cell wall recalcitrance in *Populus* (Muchero et al., 2015). These putative causal regulators may be utilized for selective breeding of trees.

These GWAS have largely focused on the common genetic variants using MLM-based methods. Identifying all functional variants that contribute to certain phenotypes has been infeasible. However, novel statistical models that combine the joint genetic effects of all variable loci at the whole-genome level have recently become the most popular types of models,

allowing breeders to use genomic information to advance breeding programs (Yang N. et al., 2014; Resende et al., 2017; Xiao et al., 2017). For example, regional heritability mapping (RHM), has been proposed that provides heritability estimates for sequence segments with common or rare genetic effects (Nagamine et al., 2012). RHM was quite powerful for the detection of true QTLs in 768 hybrid *Eucalyptus* trees (Resende et al., 2017), suggesting that complex traits in *Eucalyptus* are controlled by multiple allele variants with rare effects.

Regardless of the statistical method utilized, the availability of numerous samples is especially important for detecting associated small-effect loci. Fahrenkrog et al. (2017) conducted GWAS for wood composition traits and detected a combination of common and rare SNPs by targeted resequencing of 18,153 genes in 391 unrelated *Populus deltoides* individuals. The results indicated that low-frequency SNPs associated with several bioenergy traits, suggesting that both common and rare variants must be considered in order to show a comprehensive picture of the genetic dissection of complex traits (Resende et al., 2017). Non-SNP allelic variation is another frequently used explanation for "missing heritability" of complex traits. Gong et al. (2017) utilized InDel-based GWAS to detect the causal variants underlying growth and wood properties in 435 unrelated *Populus tomentosa* accessions and identified regulatory InDels with an average of 14.7% phenotypic variance explained. The higher contribution of InDels to phenotypic variance compared to SNP loci, with a median of *c.* 5% explained, supports the notion

that the InDels represent an effective marker system for MAS.

Genome-wide association study might also be feasible for coniferous species, even though their genomes are generally very large (often as much as 20 Gb). Uchiyama et al. (2013) examined the potential of performing GWAS in conifers using 367 unrelated plus trees of *Cryptomeria japonica* D. Don and identified six novel QTLs that were significantly associated with variation in wood property traits and the quantity of male strobili. All six SNPs were identified in sequences sharing similarity with known genes, such as genes encoding microtubule-associated protein RP/EB family members and a CLIP-associated protein. A xylem co-expression network was reconstructed in white spruce (*Picea glauca*) based on 180 wood-associated-genes, and the network hubs of several known NAC and MYB regulators, which were found using integrated GWAS and co-expression networks (Lamara et al., 2016). The genome sequence for Norway spruce (*Picea abies*) (Nystedt et al., 2013) provides a basis for functional multi-locus GWAS of wood properties in this species. A recent study used a multi-locus LASSO penalized regression method to identify 39 candidate genes involved in the formation of both early and late wood, as well as dynamic processes of juvenility, which could be useful for explaining the temporal regulation on secondary growth (Baison et al., 2018).

Over the past 5 years, new whole-genome resequencing and emerging GWAS approaches that have been used to dissect ecologically relevant traits and wood traits in forest trees (Evans et al., 2014; Ingvarsson et al., 2016). Zhang et al. (2018) recently performed GWAS and expression QTL (eQTL) analyses of 917 *P. trichocarpa* accessions and found that *hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl transferase 2* (*PtHCT2*) controlled chlorogenic acid (CGA) and partially characterized metabolite levels, providing novel insights into omics-based inference of gene function in trees. Such practical applications of GWAS require the development of novel statistical methods, databases, and experimental designs for forest trees.

FUTURE APPLICATIONS OF GWAS TO ANALYZE COMPLEX QUANTITATIVE TRAITS IN TREES

Complex phenotypic variations, such as wood formation in long-lived tree populations, involve a series of dynamic biological processes orchestrated in a precise, quantitative manner, including the transcriptional and translational regulation and the flux of metabolic intermediates of diverse biochemical pathways (Mizrachi and Myburg, 2016). We take advantage of GWAS to explore the single-marker effects, pleiotropic relationships, genetic interactions among these loci, and their interactions with different environmental factors (Ingvarsson et al., 2016; Plomion et al., 2016). Thus, based on recent advances in GWAS for identification of loci affecting wood property traits, we will discuss several statistical methods and experimental designs that

could facilitate the extensive application of GWAS in a given undomesticated tree population.

Drive Omics-Wide Association Study (OWAS) in Trees

Knowledge of the genetic basis and hierarchical interactions of higher-order variation at an intermediate level, such as transcriptomic, proteomic, and metabolomic variation ranging from genotypic to phenotypic, has proven to be a great resources as "molecular phenotypes" will be crucial for dissecting the functional pathways underlying complex quantitative traits in trees (**Figure 1**). The joint use of "-omics"-wide data will help disclose candidate genes and functional pathways underlying target traits beyond GWAS (Feltus, 2014; Mizrachi and Myburg, 2016). The comprehensive reconstruction of global biochemical networks across multiple omic layers would entail the use of omics-wide association study (OWAS) techniques proposed based on both multi-omic measurements and computational data integration. A systematic OWAS approach must therefore be adopted to map interactive eQTLs, protein QTLs (pQTLs), and metabolite QTLs (mQTLs) underlying these different profiles for quantitative traits (**Figure 1**). The opening of this new OWAS era in trees will allow us to identify minor QTLs that are masked by major loci (Luo, 2015; Xiao et al., 2017).

As a prerequisite, precise sequencing and rigorous annotation of holistic reference genomes are needed for large heterozygous tree genomes, which remains challenging. Work is also needed to refine new sequencing technologies and to develop relevant bioinformatics approaches, which would lay the foundation for identifying genome-wide allele variation and performing massively parallel RNA sequencing (RNA-seq) of 1000s of genotypes (Alonso-Blanco and Méndez-Vigo, 2014). In particular, non-coding RNAs constitute a specific component of the transcriptome worthy of attention, which could provide new insights into the regulatory mechanisms underlying eQTLs (Zhou et al., 2015; Plomion et al., 2016; Quan et al., 2018). In this new omics era for trees, many challenges remain. For example, if differences between tissues and developmental stages were not considered, conventional mapping methods would not be suitable for identifying new higher-level omics variation.

Development of High-Throughput Phenotyping Technologies

In addition to obtaining complete genome sequence information, the improvement of phenotyping methods and the availability of high-quality trait data are also priority areas for successful GWAS in forest trees (Li Z. et al., 2014; Du et al., 2016). Trees are usually large and have a long lifespan, making it difficult to perform phenotyping in outdoor test plantations on the biological, structural, and temporal levels (for example, tree biomass, wood properties, and biotic and abiotic responses to stress; **Figure 1**). The conventional procedures used for phenotyping of tree populations, which represent a "phenotyping bottleneck" (Furbank and Tester, 2011). The highly dynamic nature of genes, proteins, and metabolite levels along a temporal or developmental trajectory, point to the need for careful

FIGURE 1 | Multiple omics-wide association mapping of complex phenotypes aimed at constructing systems maps in trees. Systematic OWAS was used to map interactive expression QTLs (eQTLs), protein QTLs (pQTLs), and metabolite QTLs (mQTLs) underlying these different profiles for complex quantitative traits. Traditional genotype–phenotype associations to detect the underling QTLs were calculated via GWAS. Systems OWAS mapping treats phenotypic formation as a dynamic system composed of many interactive, functional elements at different levels. In this systems model, mRNA and non-coding RNA expression, protein and polypeptide levels, and metabolite levels are measured using different methods. These data could be integrated with epigenomic and environment interaction data to construct a complete biological working model.

experimental design and sampling conditions to ensure that the samples (Alonso-Blanco and Méndez-Vigo, 2014) will provide relevant insights about the phenotype of interest.

More attention should focus on how to develop high-throughput phenotyping (HTP) technologies to increase the precision of time-series data for functional plant traits. The emerging plant HTP platform uses a variety of imaging methods including visible imaging, imaging spectroscopy, thermal infrared imaging, fluorescence imaging, 3D imaging, and tomographic imaging to collect data for quantitative studies of complex traits related to plant growth, yield, adaptation, morphological and physiological traits (Li L. et al., 2014; Gómez-Candón et al., 2016; Yendrek et al., 2017).

Such a computational ecosystem provides tremendous opportunities for solving foundational problems in predictive phenomics and accelerating breeding efforts (Singh et al., 2016; Shakoor et al., 2017). For example, high-throughput 3D imaging can be used to map QTLs underlying variation in root architecture and seed traits in herbaceous plants (Moore et al., 2013; Topp et al., 2013). Gómez-Candón et al. (2016) used thermal IR data to characterize individual apple tree responses to drought, and it can identify genotypic variation with differential phenotypic responses to water limitation. Plant phenomics facilities have been opening worldwide (Brown et al., 2014). Careful consideration of the appropriate sensors and HTP time points for field traits is important for data collection in long-lived woody species. For example, RGB color camera

networks can be deployed to collect images continuously over the course of periods to assess biomass, growth rate, and disease progression. In this emerging era of phenomics, a tree HTP platform utilizing high-performance computing software and measurement hardwire is needed to collect multi-dimensional phenotyping data for different environments over long periods of time. Therefore, the resolution and precision of phenotyping systems should be higher for long-lived trees than the HTP systems developed for annual herbaceous plants (Meijón et al., 2014; Slovak et al., 2014; Yang W. et al., 2014).

Increase the Focus on Rare Variation and Major Structural Variation

Over the past few years, population-scale resequencing of the human genome has enabled more comprehensive analysis of allelic effects on certain complex diseases. Some rare variants (i.e., minor allele frequency [MAF] < 1% or 5%) are likely to be more extreme than common variants for susceptibility to diseases with high and low heritability (Genomes Project et al., 2010; Nelson et al., 2012; Tennessen et al., 2012). Most association studies in plant species have typically removed rare variants due to the limited power of statistical approaches for detecting their contribution to the phenotypic variation of complex traits (Porth et al., 2013; Du et al., 2015).

Progress has been made in several crops and model plants. For example, Wang S. et al. (2015) exploited rare variation mapping to discover a TCP transcription factor gene vital for

tendril development in cucumber (*Cucumis sativus* L.). Xing et al. (2015) found that a low-frequency SNP within *Brachytic2* can moderately increase yield and reduce stem height in maize. The genomes of woody species have not experienced reductions in genetic diversity due to domestication, and rare variants are abundant among tree genomes but less tractable in GWAS. Such abundant rare variants are actually important for explaining the missing heritability of complex traits (Resende et al., 2017).

Rare beneficial alleles are usually employed in current tree breeding, which once proven useful, undergo selective sweep and become fixed in all major cultivars. Balancing samples across population subdivisions and increasing the sample size can homogenize allele frequencies and elevate globally rare variants to common markers in some subpopulations. In addition, the use of multiple bi-parental crosses can break up population structure, which could increase the power of detecting multiple rare alleles that underlie natural variation. The promising approach of joint mapping with association panels and multiple bi-parental crosses in trees is also valuable for identifying low-frequency or small-effect alleles (Du et al., 2016). Decreases in sequencing costs and improvements in genotyping technologies have promoted the use of exceptionally large diverse populations to identify and analyze rare variants. Other statistical models and approaches, such as RHM (Nagamine et al., 2012) and the SKAT package for estimating the joint effect (Fahrenkrog et al., 2017), should be used to detect and dissect rare variation in tree populations.

Advances in sequencing technology and bioinformatics algorithms offer the potential to test previously undetected structural variation within re-sequencing populations of a species (Xiao et al., 2017). In contrast to the currently known SNPs and SSRs, CNVs can be categorized into deletions, duplications, and insertions that have not been given much attention in tree species (Feltus, 2014; Ingvarsson et al., 2016). Hence, investigating the roles, structures, and functions of CNVs as both biomarkers for mapping efforts and potential functional variants in woody plants should provide vital findings about phenotypic variation (Marroni et al., 2014). The presence of high levels of heterozygosity in perennial forest trees reduces the power of CNV detection. To address this issue, the precision of genome sequencing should be enhanced and the resequencing coverage depth should be increased, which would contribute to fine read-depth-based CNV identification.

Dissect the Often-Ignored Epistasis Effects in GWAS

Genetic variation of quantitative traits is classified into additive, dominant, and epistatic effects, which are conferred by numerous genes/alleles in the multiple biological networks (Du et al., 2015). It is likely that non-additive interactions between separate mutations (epistasis) often reflect the missing heritability and lack of inter-population validation of causal variants (Manolio et al., 2009). However, detecting locus-locus epistasis is challenging experimentally, statistically, and computationally due to the large number of complex interactions to be dissected in GWAS (Mackay, 2014). Several association algorithms, such as BiForce (Gyenesei et al., 2012), genome-wide interaction studies (GWIS,

Goudey et al., 2013), the Kempthorne model (Mao et al., 2007), and FastEpistasis (Schüpbach et al., 2010) make it possible to test for two-way epistasis between all locus pairs. Indeed, testing for pairwise epistasis would shed light on genetic architecture.

Additionally, dissecting the genome-wide genetic basis of exhaustive gene/allele epistasis using different populations could provide insight into the roles of epistasis in response to local adaptation in trees (Du et al., 2018). Once a gene–gene interaction network is constructed using GWAS (**Figure 2**), it can be validated as communities of genes (gene modules) that interact strongly within the complex network using RNA profiling, molecular biology, and biochemistry (Mackay, 2014; Gong et al., 2017). Quantitative traits controlled by multiple gene networks are often regulated by higher-order gene–gene interactions that are too complex to be analyzed by standard two-way epistatic tests (Du et al., 2015). In the future, more attention should focus on the pleiotropic effects of epistasis on gene expression, metabolites, and development in different geographical populations. Such knowledge would provide insights into the mechanisms and functions of the gene networks underlying plant traits.

PERSPECTIVES FOR POST-GWAS AND MOLECULAR BREEDING IN TREES

A deeper knowledge of the genetic basis of complex traits, such as wood properties, can be achieved through the integrated use of diverse statistical models and experimental tools. Systems genetics could be used to understand the molecular mechanisms of candidate genes underlying traits of interest. Some promising studies demonstrate the advantages of this approach across a tree's lifespan (Feltus, 2014; Ingvarsson et al., 2016; Plomion et al., 2016). However, functional validation and annotation of causal loci remain challenging, especially because most loci detected by GWAS are located in intergenic regions or are not canonical components of previously identified pathways (Civelek and Lusis, 2013; Xiao et al., 2017). Given the growing number of genomic variations mapped by GWAS for a given trait, a systems biology approach may be needed to validate the function of these alleles using holistic data that provide evidence for causality. Family-based accession intercrosses using different alleles could be performed to clearly determine whether an allelic variant of a specific gene is causal for the observed natural trait variation (Du et al., 2016). Promising advances in genetic transformation and genomic technologies as well as statistical and computational methods may help to address these issues.

Tree MAS breeding may depend on the use of a pleiotropic allele or a favorable combination of alleles for multiple traits of interest (Lamara et al., 2016), such as the MYB transcription factor and LncRNA genes identified for lignin and polysaccharides traits (Zhou et al., 2017; Quan et al., 2018). Thus, it is particularly important to link specific alleles to corresponding traits. High-throughput whole-genome sequences should enable the use of alternative breeding approaches, including genomics-assisted selection (GAS) and genome selection (GS) breeding strategies, as well as genome-based phenotypic prediction and
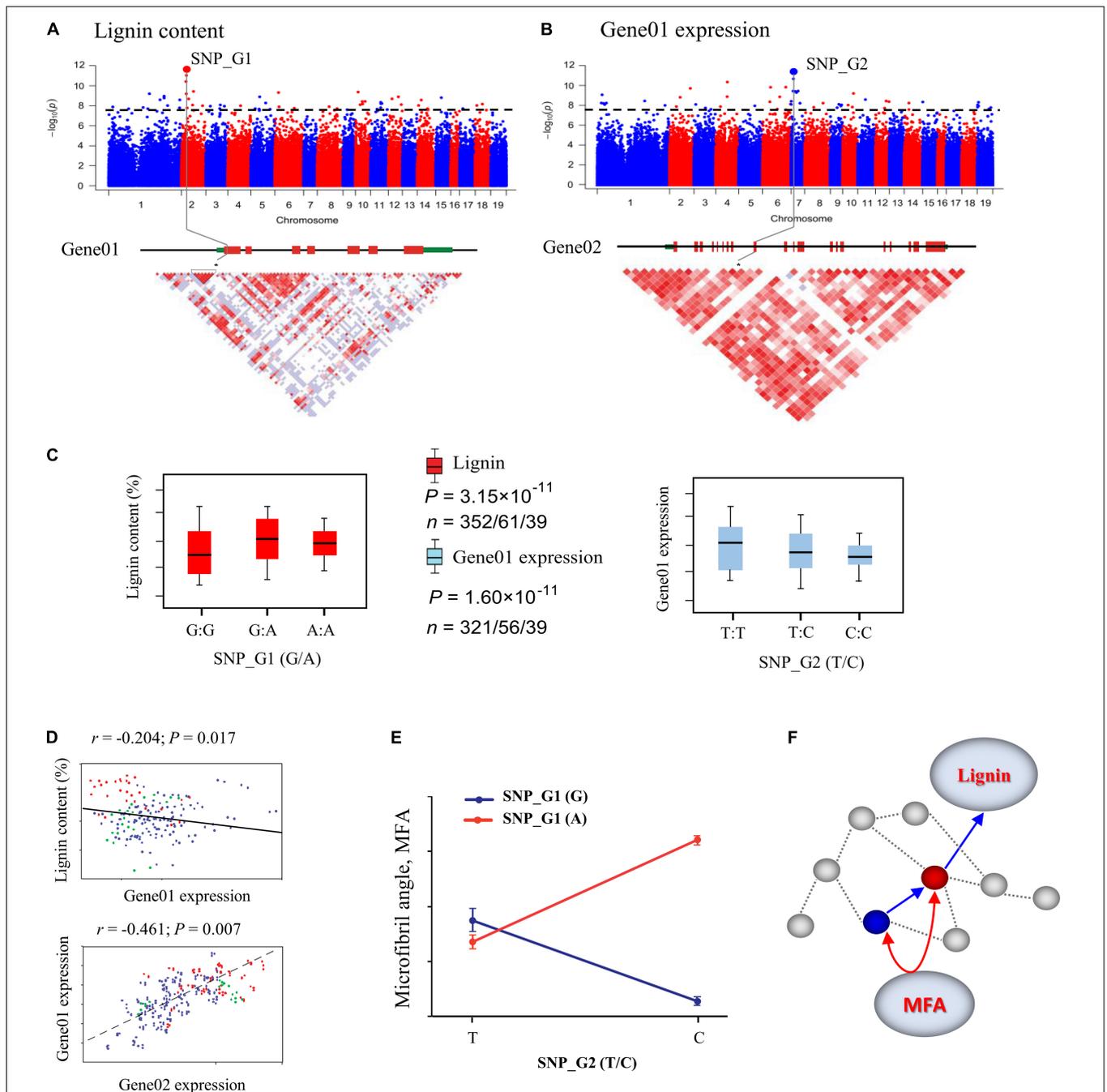
FIGURE 2 | Identification of single marker effects and epistasis for wood traits using the systems genetics integrative framework. **(A)** Manhattan plots displaying genome-wide association study (GWAS) results for lignin content. The most significant SNP (SNP_G1) associated with lignin content was identified, which is located in exon 1 of a gene (Gene01). **(B)** GWAS results for Gene01 expression using express QTL (eQTL) methods. The most significant SNP (SNP_G2) associated with Gene01 expression was identified, which is located in exon 9 of a gene (Gene02). The x-axis shows the chromosome positions and the y-axis shows significance expressed as $-\log_{10}$. The dotted horizontal line depicts the Bonferroni-adjusted significance threshold (0.05/n). Three linked genes are shown at the bottom (red rectangle, coding sequences; black line, introns; green rectangle, 5′ and 3′ untranslated regions). **(C)** Box plot for lignin content trait (red) and expression of Gene01 (sky blue) plotted as an effect of genotypes at the lead SNP. The horizontal line represents the mean and the vertical lines mark the range from the 5th and 95th percentile of the total data. **(D)** Plots of correlation between each pair for lignin content, Gene01 expression, and Gene02 expression level among the genotype classes. The r-value is based on the Pearson correlation coefficient. The P-value was calculated using the t approximation. **(E)** Pairwise interactions between SNP_G1 and SNP_G2 control Microfibril angle (MFA) with different genotypic combinations at the two loci. **(F)** These GWAS and eQTL data were used to construct the expression network shown in the figure. Gene02 (blue dot) is located at the identified eQTL significance peak and is the network hub that affects the downstream Gene01 (red dot) to affect lignin content. Gene02 can also interact with Gene01 in the network, which is associated with MFA.

design breeding using rare, recessive, large-effect mutations (Resende et al., 2012; Evans et al., 2014; Varshney et al., 2016). Specifically, using GWAS tools, it will be possible to identify many more causative genes and their roles in multiple linked traits, highlighting the potential of GWAS for systematically uncovering the balanced regulation between these traits and identifying the key hub genes that link them (Feltus, 2014; Lamara et al., 2016). However, QTL and GWAS information has not yet been successfully used for the molecular breeding of perennial trees.

High-throughput GWAS combined with precision genome editing has huge potential for combining currently available causal alleles in the gene pool (Zhou et al., 2015; Xiao et al., 2017). This GAS strategy could be used to accelerate the improvement of wood traits, as well as plant growth and stress resistance. In addition, whole-genome prediction of hybrid performance will be very important for tree breeding design, as heterosis has been successfully exploited for many major forest species. Selection and crosses based on integrated genomics approaches, as well as genetic modification via transformation, could potentially be used to develop new and superior cultivars.

REFERENCES

1000 Genomes Project Consortium, Abecasis, G. R., Altshuler, D., Auton, A., Brooks, L. D., Durbin, R. M., et al. (2010). A map of human genome variation from population-scale sequencing. *Nature* 467, 1061–1073. doi: 10. 1038/nature09534

Allwright, M. R., Payne, A., Emiliani, G., Milner, S., Viger, M., Rouse, F., et al. (2016). Biomass traits and candidate genes for bioenergy revealed through association genetics in coppiced European *Populus nigra* (L.). *Biotechnol. Biofuels* 9:195. doi: 10.1186/s13068-016-0603-1

Alonso-Blanco, C., and Méndez-Vigo, B. (2014). Genetic architecture of naturally occurring quantitative traits in plants: an updated synthesis. *Curr. Opin. Plant Biol.* 18, 37–43. doi: 10.1016/j.pbi.2014.01.002

Baison, J., Vidalis, A., Zhou, L., Chen, Z., Li, Z., Sillanpää, M. J., et al. (2018). Association mapping identified novel candidate loci affecting wood formation in Norway spruce. *bioRxiv*. [Preprint]. doi: 10.1101/292847

Beaulieu, J., Doerksen, T., Boyle, B., Clément, S., Deslauriers, M., Beauseigle, S., et al. (2011). Association genetics of wood physical traits in the conifer white spruce and relationships with gene expression. *Genetics* 188, 197–214. doi: 10.1534/genetics.110.125781

Brown, T. B., Cheng, R., Sirault, X. R., Rungrat, T., Murray, K. D., Trtilek, M., et al. (2014). TraitCapture: genomic and environment modelling of plant phenomic data. *Curr. Opin. Plant Biol.* 18, 73–79. doi: 10.1016/j.pbi.2014.02.002

Cappa, E. P., El-Kassaby, Y. A., Garcia, M. N., Acuña, C., Borralho, N. M., Grattapaglia, D., et al. (2013). Impacts of population structure and analytical models in genome-wide association studies of complex traits in forest trees: a case study in *Eucalyptus globulus*. *PLoS One* 8:e81267. doi: 10.1371/journal. pone.0081267

Chen, H. C., Song, J., Wang, J. P., Lin, Y. C., Ducoste, J., Shuford, C. M., et al. (2014). Systems biology of lignin biosynthesis in *Populus trichocarpa*: heteromeric 4-coumaric acid:coenzyme A ligase protein complex formation, regulation, and numerical modeling. *Plant Cell* 26, 876–893. doi: 10.1105/tpc.113.119685

Civelek, M., and Lusis, A. J. (2013). Systems genetics approaches to understand complex traits. *Nat. Rev. Genet.* 15, 34–48. doi: 10.1038/nrg3575

Coleman, H. D., Yan, J., and Mansfield, S. (2009). Sucrose synthase affects carbon partitioning to increase cellulose production and altered cell wall ultrastructure. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13118–13123. doi: 10.1073/pnas.090018 8106

Dillon, S. K., Nolan, M., Li, W., Bell, C., Wu, H. X., and Southerton, S. G. (2010). Allelic variation in cell wall candidate genes affecting solid wood properties in natural populations and land races of *Pinus radiata*. *Genetics* 185, 1477–1487. doi: 10.1534/genetics.110.116582

Du, Q., Gong, C., Wang, Q., Zhou, D., Yang, H., Pan, W., et al. (2016). Genetic architecture of growth traits in *Populus* revealed by integrated quantitative trait locus (QTL) analysis and association studies. *New Phytol.* 209, 1067–1082. doi: 10.1111/nph.13695

Du, Q., Pan, W., Tian, J., Li, B., and Zhang, D. (2013a). The UDP-glucuronate decarboxylase gene family in *Populus*: structure, expression, and association genetics. *PLoS One* 8:e60880. doi: 10.1371/journal.pone.0060880

Du, Q., Pan, W., Xu, B., Li, B., and Zhang, D. (2013b). Polymorphic simple sequence repeat (SSR) loci within cellulose synthase (PtoCesA) genes are associated with growth and wood properties in *Populus tomentosa*. *New Phytol.* 197, 763–776. doi: 10.1111/nph.12072

Du, Q., Tian, J., Yang, X., Pan, W., Xu, B., Li, B., et al. (2015). Identification of additive, dominant, and epistatic variation conferred by key genes in cellulose biosynthesis pathway in *Populus tomentosa*. *DNA Res.* 22, 53–67. doi: 10.1093/ dnares/dsu040

Du, Q., Yang, X., Xie, J., Quan, M., Xiao, L., Lu, W., et al. (2018). Time-specific and pleiotropic quantitative trait loci coordinately modulate stem growth in *Populus*. *Plant Biotech. J.* doi: 10.1111/pbi.13002 [Epub ahead of print].

Evans, L., Slavov, G., Rodgers-Melnick, E., Martin, J., Ranjan, P., Muchero, W., et al. (2014). Population genomics of *Populus trichocarpa* identifies signatures of selection and adaptive trait associations. *Nat. Genet.* 46, 1089–1096. doi: 10.1038/ng.3075

Fahrenkrog, A. M., Neves, L. G., Resende, M. F. Jr., Vazquez, A. I., de Los Campos, G., Dervinis, C., et al. (2017). Genome-wide association study reveals putative regulators of bioenergy traits in *Populus deltoides*. *New Phytol.* 213, 799–811. doi: 10.1111/nph.14154

Feltus, F. (2014). Systems genetics: a paradigm to improve discovery of candidate genes and mechanisms underlying complex traits. *Plant Sci.* 223, 45–48. doi: 10.1016/j.plantsci.2014.03.003

Furbank, R., and Tester, M. (2011). Phenomics-technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* 16, 635–644. doi: 10.1016/j.tplants. 2011.09.005

Gómez-Candón, D., Virlet, N., Labbé, S., Jolivot, A., and Regnard, J. L. (2016). Field phenotyping of water stress at tree scale by UAV-sensed imagery: new

insights for thermal acquisition and calibration. *Precis. Agric.* 17, 786–800. doi: 10.1007/s11119-016-9449-6

Gong, C., Du, Q., Xie, J., Quan, M., Chen, B., and Zhang, D. (2017). Dissection of insertion-deletion (InDel) variants within differentially-expressed genes involved in wood formation in Populus. *Front. Plant Sci.* 8:2199. doi: 10.3389/fpls.2017.02199

Goudey, B., Rawlinson, D., Wang, Q., Shi, F., Ferra, H., Campbell, R. M., et al. (2013). GWIS-model-free, fast and exhaustive search for epistatic interactions in case-control GWAS. *BMC Genom.* 14:S10. doi: 10.1186/1471-2164-14-S3-S10

Guerra, F. P., Wegrzyn, J. L., Sykes, R., Davis, M. F., Stanton, B. J., and Neale, D. B. (2013). Association genetics of chemical wood properties in black poplar (*Populus nigra*). *New Phytol.* 197, 162–176. doi: 10.1111/nph.12003

Gyenesei, A., Moody, J., Semple, C., Haley, C., and Wei, W. (2012). High-throughput analysis of epistasis in genome-wide association studies with BiForce. *Bioinformatics* 28, 1957–1964. doi: 10.1093/bioinformatics/bts304

Ingvarsson, P. K., Hvidsten, T. R., and Street, N. R. (2016). Towards integration of population and comparative genomics in forest trees. *New Phytol.* 212, 338–344. doi: 10.1111/nph.14153

Kang, H. M., Sul, J. H., Service, S. K., Zaitlen, N. A., Kong, S. Y., Freimer, N. B., et al. (2010). Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* 42, 348–354. doi: 10.1038/ng.548

Kang, H. M., Zaitlen, N. A., Wade, C. M., Kirby, A., Heckerman, D., Daly, M. J., et al. (2008). Efficient control of population structure in model organism association mapping. *Genetics* 178, 1709–1723. doi: 10.1534/genetics.107.080101

Lamara, M., Raherison, E., Lenz, P., Beaulieu, J., Bousquet, J., and MacKay, J. (2016). Genetic architecture of wood properties based on association analysis and co-expression networks in white spruce. *New Phytol.* 210, 240–255. doi: 10.1111/nph.13762

Li, L., Zhang, Q., and Huang, D. (2014). A review of imaging techniques for plant phenotyping. *Sensors* 14, 20078–20111. doi: 10.3390/s141120078

Li, Z., Hallingbäck, H. R., Abrahamsson, S., Fries, A., and Gull, B. A. (2014). Functional multi-locus QTL mapping of temporal trends in Scots pine wood traits. *G3* 4, 2365–2379. doi: 10.1534/g3.114.014068

Lippert, C., Listgarten, J., Liu, Y., Kadie, C. M., Davidson, R. I., and Heckerman, D. (2011). FaST linear mixed models for genome-wide association studies. *Nat. Methods* 8, 833–835. doi: 10.1038/nmeth.1681

Luo, J. (2015). Metabolite-based genome-wide association studies in plants. *Curr. Opin. Plant Biol.* 24, 31–38. doi: 10.1016/j.pbi.2015.01.006

Mackay, T. (2014). Epistasis and quantitative traits: using model organisms to study gene-gene interactions. *Nat. Rev. Genet.* 15, 22–33. doi: 10.1038/nrg3627

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., et al. (2009). Finding the missing heritability of complex diseases. *Nature* 461, 747–753. doi: 10.1038/nature08494

Mao, Y., London, N., Ma, L., Dvorkin, D., and Da, Y. (2007). Detection of SNP epistasis effects of quantitative traits using an extended Kempthorne model. *Physiol. Genom.* 28, 46–52. doi: 10.1152/physiolgenomics.00096.2006

Marroni, F., Pinosio, S., and Morgante, M. (2014). Structural variation and genome complexity: is dispensable really dispensable? *Curr. Opin. Plant Biol.* 18, 31–36. doi: 10.1016/j.pbi.2014.01.003

McKown, A. D., Klápště, J., Guy, R. D., Geraldes, A., Porth, I., Hannemann, J., et al. (2014). Genome-wide association implicates numerous genes underlying ecological trait variation in natural populations of *Populus trichocarpa*. *New Phytol.* 203, 535–553. doi: 10.1111/nph.12815

Meijón, M., Satbhai, S. B., Tsuchimatsu, T., and Busch, W. (2014). Genome-wide association study using cellular traits identifies a new regulator of root development in *Arabidopsis*. *Nat. Genetics* 46, 77–81. doi: 10.1038/ng.2824

Mellerowicz, E., and Sundberg, B. (2008). Wood cell walls: biosynthesis, developmental dynamics and their implications for wood properties. *Curr. Opin. Plant Biol.* 11, 293–300. doi: 10.1016/j.pbi.2008.03.003

Mewalal, R., Mizrachi, E., Mansfield, S., and Myburg, A. (2014). Cell wall-related proteins of unknown function: missing links in plant cell wall development. *Plant Cell Physiol.* 55, 1031–1043. doi: 10.1093/pcp/pcu050

Mizrachi, E., and Myburg, A. (2016). Systems genetics of wood formation. *Curr. Opin. Plant Biol.* 30, 94–100. doi: 10.1016/j.pbi.2016.02.007

Moore, C. R., Johnson, L. S., Kwak, I. Y., Livny, M., Broman, K. W., and Spalding, E. P. (2013). High-throughput computer vision introduces the time axis to a quantitative trait map of a plant growth response. *Genetics* 195, 1077–1086. doi: 10.1534/genetics.113.153346

Muchero, W., Guo, J., DiFazio, S. P., Chen, J. G., Ranjan, P., Slavov, G. T., et al. (2015). High-resolution genetic mapping of allelic variants associated with cell wall chemistry in *Populus*. *BMC Genom.* 16:24. doi: 10.1186/s12864-015-1215-z

Müller, B., de Almeida Filho, J., Lima, B., Garcia, C., Missiaggia, A., Aguiar, A., et al. (2018). Independent and Joint-GWAS for growth traits in *Eucalyptus* by assembling genome-wide data for 3373 individuals across four breeding populations. *New Phytol.* doi: 10.1111/nph.15449 [Epub ahead of print].

Nagamine, Y., Pong-Wong, R., Navarro, P., Vitart, V., Hayward, C., Rudan, I., et al. (2012). Localising loci underlying complex trait variation using regional genomic relationship mapping. *PLoS One* 7:e46501. doi: 10.1371/journal.pone.0046501

Nelson, M. R., Wegmann, D., Ehm, M. G., Kessner, D., St, Jean P, Verzilli, C., et al. (2012). An abundance of rare functional variants in 202 drug target genes sequenced in 14,002 people. *Science* 337, 100–104. doi: 10.1126/science.1217876

Novaes, E., Osorio, L., Drost, D., Miles, B., Boaventura-Novaes, C., Benedict, C., et al. (2009). Quantitative genetic analysis of biomass and wood chemistry of Populus under different nitrogen concentrations. *New Phytol.* 182, 878–890. doi: 10.1111/j.1469-8137.2009.02785.x

Nystedt, B., Street, N. R., Wetterbom, A., Zuccolo, A., Lin, Y. C., Scofield, D. G., et al. (2013). The Norway spruce genome sequence and conifer genome evolution. *Nature* 497, 579–584. doi: 10.1038/nature12211

Plomion, C., Bastien, C., Bogeat-Triboulot, M. B., Bouffier, L., Déjardin, A., Duplessis, S., et al. (2016). Forest tree genomics: 10 achievements from the past 10 years and future prospects. *Ann. Forest Sci.* 73, 77–103. doi: 10.1007/s13595-015-04883

Plomion, C., Leprovost, G., and Stokes, A. (2001). Wood formation in trees. *Plant Physiol.* 127, 1513–1523. doi: 10.1104/pp.127.4.1513

Porth, I., Klapšte, J., Skyba, O., Hannemann, J., McKown, A. D., Guy, R. D., et al. (2013). Genome-wide association mapping for wood characteristics in Populus identifies an array of candidate single nucleotide polymorphisms. *New Phytol.* 200, 710–726. doi: 10.1111/nph.12422

Pritchard, J., Stephens, M., Rosenberg, N., and Donnelly, P. (2000). Association mapping in structured populations. *Am. J. Hum. Genet.* 67, 170–181. doi: 10.1086/302959

Quan, M., Du, Q., Xiao, L., Lu, W., Wang, L., Xie, J., et al. (2018). Genetic architecture underlying the lignin biosynthesis pathway involves noncoding RNAs and transcription factors for growth and wood properties in Populus. *Plant Biotech. J.* doi: 10.1111/pbi.12978 [Epub ahead of print].

Resende, M. D., Resende, M. F. Jr., Sansaloni, C. P., Petroli, C. D., Missiaggia, A. A., Aguiar, A. M., et al. (2012). Genomic selection for growth and wood quality in *Eucalyptus*: capturing the missing heritability and accelerating breeding for complex traits in forest trees. *New Phytol.* 194, 116–128. doi: 10.1111/j.1469-8137.2011.04038.x

Resende, R. T., Resende, M. D., Silva, F. F., Azevedo, C. F., Takahashi, E. K., Silva-Junior, O. B., et al. (2017). Regional heritability mapping and genome-wide association identify loci for complex growth, wood and disease resistance traits in *Eucalyptus*. *New Phytol.* 213, 1287–1300. doi: 10.1111/nph.14266

Robischon, M., Du, J., Miura, E., and Groover, A. (2011). The *Populus* class III HD ZIP, popREVOLUTA, influences cambium initiation and patterning of woody stems. *Plant Physiol.* 155, 1214–1225. doi: 10.1104/pp.110.167007

Schüpbach, T., Xenarios, I., Bergmann, S., and Kapur, K. (2010). FastEpistasis: a high performance computing solution for quantitative trait epistasis. *Bioinformatics* 26, 1468–1469. doi: 10.1093/bioinformatics/btq147

Sewell, M. M., Bassoni, D. L., Megraw, R. A., Wheeler, N. C., and Neale, D. B. (2000). Identification of QTLs influencing wood property traits in loblolly pine (*Pinus taeda*, L.). I. Physical wood properties. *Theor. Appl. Genet.* 101, 1273–1281. doi: 10.1007/s001220051607

Shakoor, N., Lee, S., and Mockler, T. C. (2017). High throughput phenotyping to accelerate crop breeding and monitoring of diseases in the field. *Curr. Opin. Plant Biol.* 38, 184–192. doi: 10.1016/j.pbi.2017.05.006

Silva-Junior, O. B., and Grattapaglia, D. (2015). Genome-wide patterns of recombination, linkage disequilibrium and nucleotide diversity from pooled

resequencing and single nucleotide polymorphism genotyping unlock the evolutionary history of *Eucalyptus grandis*. *New Phytol.* 208, 830–845. doi: 10.1111/nph.13505

Singh, A., Ganapathysubramanian, B., Singh, A. K., and Sarkar, S. (2016). Machine learning for high-throughput stress phenotyping in plants. *Trends Plant Sci.* 21, 110–124. doi: 10.1016/j.tplants.2015.10.015

Slovak, R., Göschl, C., Su, X., Shimotani, K., Shiina, T., and Busch, W. (2014). A scalable open-source pipeline for large-scale root phenotyping of *Arabidopsis*. *Plant Cell* 26, 2390–2403. doi: 10.1105/tpc.114.124032

Tennessen, J. A., Bigham, A. W., O'Connor, T. D., Fu, W., Kenny, E. E., Gravel, S., et al. (2012). Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science* 337, 64–69. doi: 10.1126/science.1219240

Thumma, B. R., Matheson, B. A., Zhang, D., Meeske, C., Meder, R., Downes, G. M., et al. (2009). Identification of a cis-acting regulatory polymorphism in a Eucalypt COBRA-like gene affecting cellulose content. *Genetics* 183, 1153–1164. doi: 10.1534/genetics.109.106591

Thumma, B. R., Nolan, M. F., Evans, R., and Moran, G. F. (2010). Quantitative trait locus (QTL) analysis of wood quality traits in *Eucalyptus nitens*. *Tree Genet. Genomes* 6, 305–317. doi: 10.1007/s11295-009-0250-9

Topp, C. N., Iyer-Pascuzzi, A. S., Anderson, J. T., Lee, C. R., Zurek, P. R., Symonova, O., et al. (2013). 3D phenotyping and quantitative trait locus mapping identify core regions of the rice genome controlling root architecture. *Proc. Natl. Acad. Sci. U.S.A.* 110, E1695–E1704. doi: 10.1073/pnas.1304354110

Uchiyama, K., Iwata, H., Moriguchi, Y., Ujino-Ihara, T., Ueno, S., Taguchi, Y., et al. (2013). Demonstration of genome-wide association studies for identifying markers for wood property and male strobili traits in *Cryptomeria japonica*. *PLoS One* 8:e79866. doi: 10.1371/journal.pone.0079866

Varshney, R. K., Singh, V. K., Hickey, J. M., Xun, X., Marshall, D. F., Wang, J., et al. (2016). Analytical and decision support tools for genomics-assisted breeding. *Trends Plant Sci.* 21, 354–363. doi: 10.1016/j.tplants.2015.10.018

Wang, J., Street, N. R., Scofield, D. G., and Ingvarsson, P. K. (2015). Natural selection and recombination rate variation shape nucleotide polymorphism across the genomes of three related *Populus* species. *Genetics* 202, 1185–1200. doi: 10.1534/genetics.115.183152

Wang, S., Yang, X., Xu, M., Lin, X., Lin, T., Qi, J., et al. (2015). A rare SNP identified a TCP transcription factor essential for tendril development in cucumber. *Mol. Plant* 8, 1795–1808. doi: 10.1016/j.molp.2015.10.005

Wang, J., Street, N. R., Scofield, D. G., and Ingvarsson, P. K. (2016). Variation in linked selection and recombination drive genomic divergence during allopatric speciation of European and American aspens. *Mol. Biol. Evol.* 33, 1754–1767. doi: 10.1101/029561

Wegrzyn, J. L., Eckert, A. J., Choi, M., Lee, J. M., Stanton, B. J., Sykes, R., et al. (2010). Association genetics of traits controlling lignin and cellulose biosynthesis in black cottonwood (*Populus trichocarpa*, Salicaceae) secondary xylem. *New Phytol.* 188, 515–532. doi: 10.1111/j.1469-8137.2010.03415.x

Wu, X., Li, Y., Shi, Y., Song, Y., Zhang, D., Li, C., et al. (2016). Joint-linkage mapping and GWAS reveal extensive genetic loci that regulate male inflorescence size in maize. *Plant Biotech. J.* 14, 1551–1562. doi: 10.1111/pbi.12519

Xiao, Y., Liu, H., Wu, L., Warburton, M., and Yan, J. (2017). Genome-wide association studies in maize: praise and stargaze. *Mol. Plant* 10, 359–374. doi: 10.1016/j.molp.2016.12.008

Xing, A., Gao, Y., Ye, L., Zhang, W., Cai, L., Ching, A., et al. (2015). A rare SNP mutation in Brachytic2 moderately reduces plant height and increases yield potential in maize. *J. Exp. Bot.* 66, 3791–3802. doi: 10.1093/jxb/erv182

Yang, N., Lu, Y., Yang, X., Huang, J., Zhou, Y., Ali, F., et al. (2014). Genome wide association studies using a new nonparametric model reveal the genetic architecture of 17 agronomic traits in an enlarged maize association panel. *PLoS Genet* 10:e1004573. doi: 10.1371/journal.pgen.1004573

Yang, W., Guo, Z., Huang, C., Duan, L., Chen, G., Jiang, N., et al. (2014). Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. *Nat. Commun.* 5:5087. doi: 10.1038/ncomms6087

Ye, Z., and Zhong, R. (2015). Molecular control of wood formation in trees. *J. Exp. Bot.* 66, 4119–4131. doi: 10.1093/jxb/erv081

Yendrek, C., Tomaz, T., Montes, C. M., Cao, Y., Morse, A. M., Brown, P. J., et al. (2017). High-throughput phenotyping of maize leaf physiology and biochemistry using hyperspectral reflectance. *Plant Physiol.* 173, 614–626. doi: 10.1104/pp.16.01447

Yin, T., Zhang, X., Gunter, L., Priya, R., Sykes, R., Davis, M., et al. (2010). Differential detection of genetic loci underlying stem and root lignin content in Populus. *PLoS One* 5:e14021. doi: 10.1371/journal.pone.0014021

Yu, J., Pressoir, G., Briggs, W. H., Vroh, Bi I, Yamasaki, M., Doebley, J. F., et al. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 38, 203–208. doi: 10.1038/ng1702

Zhang, J., Yang, Y., Zheng, K., Xie, M., Feng, K., Jawdy, S. S., et al. (2018). Genome-wide association studies and expression-based quantitative trait loci analyses reveal roles of HCT2 in caffeoylquinic acid biosynthesis and its regulation by defense-responsive transcription factors in *Populus*. *New Phytol.* 220, 502–516. doi: 10.1111/nph.15297

Zhang, Z., Ersoz, E., Lai, C. Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., et al. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* 42, 355–360. doi: 10.1038/ng.546

Zhou, D., Du, Q., Chen, J., Wang, Q., and Zhang, D. (2017). Identification and allelic dissection uncover roles of lncRNAs in secondary growth of *Populus tomentosa*. *DNA Res.* 24, 473–486. doi: 10.1093/dnares/dsx018

Zhou, X., Jacobs, T., Xue, L., Harding, S., and Tsai, C. (2015). Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial *Populus* reveals 4-coumarate: CoA ligase specificity and redundancy. *New Phytol.* 208, 298–301. doi: 10.1111/nph.13470

Zhou, X., and Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* 44, 821–824. doi: 10.1038/ng.2310

Zinkgraf, M., Liu, L., Groover, A., and Filkov, V. (2017). Identifying gene coexpression networks underlying the dynamic regulation of wood-forming tissues in *Populus* under diverse environmental conditions. *New Phytol.* 214, 1464–1478. doi: 10.1111/nph.14492

Quantitative Genetics and Genomics Converge to Accelerate Forest Tree Breeding

Dario Grattapaglia [1,2,3,4*], Orzenil B. Silva-Junior [1,2], Rafael T. Resende [1], Eduardo P. Cappa [5,6], Bárbara S. F. Müller [1,3], Biyue Tan [7], Fikret Isik [4], Blaise Ratcliffe [8] and Yousry A. El-Kassaby [8]

[1] EMBRAPA Recursos Genéticos e Biotecnologia, Brasília, Brazil, [2] Programa de Ciências Genômicas e Biotecnologia, Universidade Católica de Brasília, Brasília, Brazil, [3] Departamento de Biologia Celular, Universidade de Brasília, Brasília, Brazil, [4] Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC, United States, [5] Centro de Investigación de Recursos Naturales, Instituto de Recursos Biológicos, INTA, Buenos Aires, Argentina, [6] Consejo Nacional de Investigaciones Científicas y Técnicas, Buenos Aires, Argentina, [7] Biomaterials Division, Stora Enso AB, Stockholm, Sweden, [8] Department of Forest and Conservation Sciences, Faculty of Forestry, University of British Columbia, Vancouver, BC, Canada

Forest tree breeding has been successful at delivering genetically improved material for multiple traits based on recurrent cycles of selection, mating, and testing. However, long breeding cycles, late flowering, variable juvenile-mature correlations, emerging pests and diseases, climate, and market changes, all pose formidable challenges. Genetic dissection approaches such as quantitative trait mapping and association genetics have been fruitless to effectively drive operational marker-assisted selection (MAS) in forest trees, largely because of the complex multifactorial inheritance of most, if not all traits of interest. The convergence of high-throughput genomics and quantitative genetics has established two new paradigms that are changing contemporary tree breeding dogmas. Genomic selection (GS) uses large number of genome-wide markers to predict complex phenotypes. It has the potential to accelerate breeding cycles, increase selection intensity and improve the accuracy of breeding values. Realized genomic relationships matrices, on the other hand, provide innovations in genetic parameters' estimation and breeding approaches by tracking the variation arising from random Mendelian segregation in pedigrees. In light of a recent flow of promising experimental results, here we briefly review the main concepts, analytical tools and remaining challenges that currently underlie the application of genomics data to tree breeding. With easy and cost-effective genotyping, we are now at the brink of extensive adoption of GS in tree breeding. Areas for future GS research include optimizing strategies for updating prediction models, adding validated functional genomics data to improve prediction accuracy, and integrating genomic and multi-environment data for forecasting the performance of genetic material in untested sites or under changing climate scenarios. The buildup of phenotypic and genome-wide data across large-scale breeding populations and advances in computational prediction of discrete genomic features should also provide opportunities to enhance the application of genomics to tree breeding.

Keywords: genomic selection (GS), tree breeding, quantitative genetics, whole-genome regression, single nucleotide polymorphisms (SNP), marker assisted selection (MAS), realized genomic relationship

INTRODUCTION

Forest tree breeding encompasses a number of steps to increase the frequency of advantageous alleles for several traits concurrently in a target population. Recurrent cycles of selection ultimately result in genetically improved planting material by maximizing genetic gain per unit time in the most cost-effective way (Namkoong et al., 1988; White et al., 2007). Long breeding cycles, late and poor flowering, weak juvenile-mature correlations and changes in climate, market demands and emerging pest and disease pressures, pose, however, daunting challenges. The advancement and ultimate output of tree breeding programs are therefore highly conditional on the length of a breeding cycle. To maximize genetic gains per unit time, extensive efforts in tree breeding were devoted to the two fundamental means by which the length of a breeding cycle can be decreased, namely, early selection and accelerated breeding. While the former is based on the understanding of juvenile-mature correlations and the practice of selection on juvenile traits (Williams, 1988), the latter involved early flower induction methods with hormone, stress treatments, and top grafting (Greenwood et al., 1991; Hasan and Reid, 1995). In the late 80s, the advent of DNA markers and two seminal papers on the dissection of discrete Mendelian factors underlying quantitative traits (Lander and Botstein, 1989), and marker-assisted selection (MAS) (Lande and Thompson, 1990), were seen as powerful tools to overcome the time challenge of tree breeding (El-Kassaby, 1982; Neale and Williams, 1991; Grattapaglia et al., 1992; Williams and Neale, 1992).

Here we cover the current state of the science on the general theme of optimizing and accelerating tree breeding using genomic technologies. A brief overview of the path across QTL (quantitative trait loci) and association mapping is first presented. It provides a quick historical perspective on how and why we reached the point of convergence between quantitative genetics and genomics. Furthermore, it also serves to substantiate the fact that reductionist "genetic dissection" approaches or attempts to use single candidate genes or diffuse, indirect information from transcriptomics, have not proven useful for breeding practice and are therefore not discussed further. We focus on the factors that affect and the challenges that remain to fully integrate genomic data in tree breeding in light of the recent promising results of whole-genome prediction. Although making predictions is difficult "especially about the future" as Niels Bohr and others once amusingly said, we attempt to look at the near future of tree breeding when genotyping, whole genome sequencing and computational prediction of genomic features for thousands of trees will not be limiting. We anticipate a future where the progressive advances made possible by routine genomic selection (GS) in multiple large populations will provide a more powerful platform to revisit the discovery of discrete genomic elements that may further enhance whole-genome phenotype prediction and eventually allow direct discrete interventions at the DNA sequence level.

THE PATH FROM GENETIC DISSECTION TO GENOMIC SELECTION

The prospects of MAS for forest trees was properly doubted early on, limiting its potential value to specific genetic backgrounds resulting from linkage equilibrium of forest tree populations, QTLs interacting with environments and changes of allele frequencies across generations (Strauss et al., 1992). Notwithstanding those sound advices, a number of QTL mapping experiments in the major conifers and eucalypts advanced, encouraged by the promising results of QTL mapping in inbred crops and model systems. In retrospect it is startling to consider how far removed from real-life tree breeding those bi-parental QTL mapping studies in forest trees were (Grattapaglia, 2017). The motivating hypothesis was that it would be possible to locate and estimate the effects of most individual QTLs underlying complex traits in every population and environment and implement them in tree breeding practice. A substantial number of studies reporting hundreds QTLs in forest trees was reported (reviewed in (Kirst et al., 2004; Grattapaglia et al., 2009; Neale and Kremer, 2011). Although several supposedly "major effect" QTLs were found in those early studies, those proved to be largely overestimated in effect size and underestimated in number. Indeed, subsequent multi-family experiments and larger sample sizes, revealed significantly larger numbers of QTLs with correspondingly smaller effects and inconsistent performance across environments and genetic backgrounds (Ukrainetz et al., 2008; Novaes et al., 2009; Thumma et al., 2010; Gion et al., 2011).

To solve the perceived shortcomings of QTL detection in single mapping families, association genetics was put forward as a way to provide population-wide marker–trait associations applicable to breeding (Neale and Savolainen, 2004). The limitation of methods to interrogate DNA polymorphisms at the time only allowed candidate-gene approaches (Thumma et al., 2005; Gonzalez-Martinez et al., 2007), which were then followed by genome-wide association mapping (GWAS) in several forest tree species (Beaulieu et al., 2011; Cumbie et al., 2011; Cappa et al., 2013; Porth et al., 2013; Mckown et al., 2014). However, irrespective of the marker density used, population size and improved analytical methods to account for low-frequency variants (Fahrenkrog et al., 2017; Müller et al., 2017, 2018; Resende et al., 2017a), only few polymorphisms of very modest effect have been detected, largely still lacking independent validation, the cornerstone for the scientific credibility of GWAS results. In effect, after 25 years of research efforts based on the principle and experimental approaches of genetic dissection of quantitative traits, no translation of such efforts to operational tree breeding was achieved (Grattapaglia et al., 2009; Grattapaglia, 2014; Isik, 2014).

The ineffectiveness of fully dissecting complex traits, and the limitations of MAS has not been exclusive to forest trees, but has also been recognized in crops (Bernardo, 2008) and domestic animals (Dekkers, 2004). This realization has caused a significant shift in the paradigm and technical approach to

plant and animal MAS. These fields have now moved from the *a priori* discovery of discrete marker-trait associations to the capture of the whole-genome effect assisted by DNA marker data, harmonizing with the multifactorial polygenic nature of quantitative genetics, as predicted by Fisher's infinitesimal model (Fisher, 1918). This shift was only possible following the development of improved and accessible genomic technologies that allow interrogating thousands of genome-wide single-nucleotide polymorphisms (SNPs) using cost effective platforms. The concept of using the "total allelic" (Nejati-Javaremi et al., 1997) or "total genomic" (Haley and Visscher, 1998) relationship from marker data to derive estimates of breeding values was later termed "Genomic Selection" (GS) in Meuwissen et al. (2001) seminal paper. It demonstrated that "*the selection on genetic values predicted from markers could substantially increase the rate of genetic gain per unit time in animals and plants, especially if combined with techniques to shorten the generation interval.*"

GS employs a genome-wide panel of markers, typically SNP (single nucleotide polymorphism), whose effects on the phenotype are estimated in a "training" population. In forest trees, such a training set is usually composed by sampling one to a few thousand individuals in progeny trials derived from mating a few dozen parents that constitute the target elite germplasm bred. SNPs are used to build prediction models to be later applied to "selection candidates" for which only genotypes are gathered and phenotypes are predicted by the genomic data. The prediction models are cross-validated against a "validation" population, a set of genetically related individuals to the training set but that did not participate in the estimation of marker effects. A prediction model that delivers a high correlation between the observed and predicted breeding values is subsequently used in the breeding phase to calculate the genomic estimated breeding or total genotypic values of the selection candidates (**Figure 1**). GS fundamentally exploits the genetic relationship between the training population and the prospective selection candidates and to a lesser extent the linkage disequilibrium (LD) between marker data and QTL effects. By precluding prior discrete marker selection derived from rigorous significance tests, and by estimating marker effects in a larger and breeding-representative population of trees, GS captures substantial proportions of the heritability contributed by the large numbers of genomic effects that QTL mapping or GWAS are, on principle, neither able nor intended to capture.

PERSPECTIVES OF GENOMIC SELECTION IN TREE BREEDING

GS can have a substantial impact on the rate of genetic gain. Let's recall Falconer's breeder's equation ($\Delta G = ir\sigma_A/L$) (Falconer, 1989), where i is the selection intensity; r is the accuracy of selection, or heritability in the original Falconer's expression, corresponding to the correlation between the estimated and true breeding values; σ_A is the additive-genetic standard deviation of the trait under inspection; and L is the generation interval. GS can increase the rate of genetic gain of breeding cycle by increasing (*i*) because the phenotypes of a much larger number of seedlings in the nursery can be predicted with marker data compared to the number of trees that can be tested in conventional field progeny trials. Additionally, the use of realized genomic relationships is associated with increased accuracy in estimating σ_A and breeding values (*r*) (Hayes et al., 2009; El-Dien et al., 2018). Yet, in forest trees, the potentially greatest impact of GS on the rate of genetic progress will originate from decreasing (*L*). Phenotypes of the selection candidates can be predicted at very early ages, for example, when the seedlings are a few weeks old. GS not only could preclude or at least enhance the efficiency of progeny testing but would also optimize clonal testing phases by advancing a smaller number of pre-selected trees to be assessed in multi-site expanded clonal trials (Resende et al., 2012a) (**Figure 1**). In conifers, GS coupled to somatic embryogenesis for clonal propagation of elite genotypes could allow selecting elite zygotic embryos based on their genomic value saving a significant amount of time, and avoiding the costs and uncertainties currently involved in cryopreservation rescue (Resende et al., 2012b). Additionally, GS will allow simultaneous and early selection for multiple trait in large numbers of individuals, an impossible task in conventional tree breeding that currently largely adopts tandem selection. The final impact of GS would therefore be a significant improvement in the general efficiency of a tree breeding program, provided, of course, that genotyping is inexpensive and GS models are accurate.

What makes GS distinctive from what tree breeders have done so far is that instead of relying uniquely on the expected pedigree, frequently prone to errors, DNA data allows one to build additive and non-additive genomic relationship matrices that more accurately specify the relationships among individuals and simultaneously account for contemporary as well as historical pedigree. This procedure not only allows rectifying pedigree inaccuracies, but critically it captures the within family variation resulting from random Mendelian segregation term. Accordingly, the realized genetic covariances are now based on the actual fraction of the genome that is identical by descent or by state between individuals (Vanraden, 2008). It has been shown in a number of studies in forest trees that realized genomic relationships can produce more accurate predictions than pedigrees alone (Munoz et al., 2014; El-Dien et al., 2015, 2018; Bouvet et al., 2016; Cappa et al., 2017, 2018; Tan et al., 2018). Additionally, the realized genomic relationships of a small subset of the progeny testing population have been effectively combined with a substantially large proportion of un-genotyped individuals in a single-step analysis (Legarra et al., 2009). This method was dubbed "HBLUP," since the best linear unbiased predictors (BLUPs) of breeding values are derived using a single (**H**) genetic covariance matrix that combines the pedigree-based average numerator relationship matrix (**A**) with the marker-based relationship matrix (**G**). HBLUP increases the precision of the genetic parameters generated from traditional pedigrees as shown in recent studies with forest trees (Cappa et al., 2017, 2018; Ratcliffe et al., 2017).

FIGURE 1 | Genomic selection in forest trees. GS begins with the development of a predictive model for the traits of interest **(Left panel)**, which are then used in the GS cycles **(Right panel)** and progressively updated. GS uses genome-wide markers whose effects on the phenotype are estimated concurrently in a large and representative "training population" of individuals without applying severe significance tests. Markers are retained as forecasters of phenotypes in prediction models to be later applied to "selection candidates" for which only genotypes are collected. The prediction models are cross-validated against a "validation population," a set of individuals of the same reference population that were not used for the estimation of marker effects. Once a prediction model is shown to provide adequate accuracy, it can be used in the GS cycle. An array of selection candidates - full of half-sib families derived from crossing either the original elite parents of the training set, or elite individuals selected in the training set - are genotyped and have their breeding values (GEBV) and/or genotypic values (GEGV; additive + non-additive effects) estimated using the model developed earlier. Top ranked seedlings for GEBV are subject to early flower induction and inter-mated to create the next generation of breeding. Top ranked seedlings for GEGV are clonally propagated and tested in verification clonal trials where elite clones are eventually selected for operational plantation. Additionally, all or subsets of the already genotyped selection candidates are planted in experimental design and phenotyped at the target selection age to provide genotype and trait data for GS model updating as GS generations advance and climate changes.

GS: ADVANCES AND CHALLENGES IN FOREST TREES

A comprehensive and time-lined list of empirical GS reports in forest tree species was recently published (Grattapaglia, 2017) and it is now updated in **Table 1**. Prediction accuracies have been largely very good, matching or surpassing those obtainable by pedigree-based phenotypic selection, in line with former simulations (Grattapaglia and Resende, 2011; Iwata et al., 2011; Denis and Bouvet, 2013). When considering the practicalities of tree breeding, however, a number of factors that affect the prospects of GS have to be considered, including the composition of training populations, analytical methods, genotype x environment interaction (G*E), age-age correlations, the long term models performance and cost and quality of DNA

marker data. All these have been the subject of research and reviewed in detail in the context of tree breeding by Grattapaglia (2017) and are briefly discussed below in light of the experimental results reported to date in forest trees.

GS experiments in forest trees have capitalized on the existing structure and diversity of breeding populations and their designs that account for the expected relationship between training and prospective selection candidates. Training populations of several hundred to a few thousand individuals sampled from existing progeny trials with effective population sizes consistent with those used in operational breeding have provided good predictions in essentially all studies and for all traits. Analytical methods differing with respect to the presumed trait architecture have been used and compared. In all studies, the ridge regression best linear unbiased prediction (RR-BLUP), with

TABLE 1 | Timeline summary of experimental genomic selection studies in forest tree species published to date.

Forest tree	DNA marker data	Traits analyzed	References
Eucalypts (*Eucalyptus grandis x E. urophylla, E. camaldulensis* - hybrids)	3,129 DArT array	Growth and wood quality	Grattapaglia et al., 2011 Resende et al., 2012a
Loblolly pine (*Pinus taeda*)	4,852 SNP chip	Growth, wood and disease	Resende et al., 2012b Resende et al., 2012c Munoz et al., 2014 De Almeida et al., 2016
Loblolly pine (*Pinus taeda*)	3,461 SNP chip	Growth and wood quality	Zapata-Valenzuela et al., 2012 Zapata-Valenzuela et al., 2013
Eucalypts (*Eucalyptus grandis x E. urophylla* - hybrids)	29,090 SNP EuCHIP60K	Growth and wood quality	Lima, 2014
White spruce (*Picea glauca*)	6,932 SNP chip	Growth and wood quality	Beaulieu et al., 2014a Beaulieu et al., 2014b Ratcliffe et al., 2017
Interior spruce (*Picea glauca x P. engelmannii*)	8,868 to 62,198 GbS SNPs	Growth and wood quality	El-Dien et al., 2015 Ratcliffe et al., 2015 El-Dien et al., 2018
Maritime pine (*Pinus pinaster*)	4,332 SNP chip	Growth and form	Isik et al., 2016 Bartholome et al., 2016
Eucalypts (*Eucalyptus grandis x E. urophylla* - hybrids)	3,303 DArTseq SNPs	Growth	Bouvet et al., 2016
Eucalypt (*Eucalyptus pellita* and *E. benthamii*)	19,506 SNP EuCHIP60K	Growth	Müller et al., 2017
Eucalypt (*Eucalyptus grandis x E. urophylla* - hybrids)	24,806 SNP EuCHIP60K	Growth and wood quality	Resende et al., 2017b
Eucalypt (*Eucalyptus globulus*)	12,000 SNP EuCHIP60K	Growth and wood quality	Duran et al., 2017
Black spruce (*Picea mariana*)	4,993 SNP chip	Growth and wood quality	Lenz et al., 2017
Eucalypt (*Eucalyptus grandis*)	2,816 DArT array	Growth	Cappa et al., 2017 Cappa et al., 2018
Eucalypt (*Eucalyptus grandis x E. urophylla* - hybrids)	41,304 SNP EuCHIP60K	Growth and wood quality	Tan et al., 2017 Tan et al., 2018
Douglas fir (*Pseudotsuga menziesii*)	69,551 exome capture SNPs	Growth and wood quality	Thistlethwaite et al., 2017
Norway spruce (*Picea abies*)	116,765 exome capture SNPs	Growth and wood quality	Chen et al., 2018
Eucalypt (*Eucalyptus nitens*)	12,236 SNP EuCHIP60K	Growth and wood quality	Suontama et al., 2018
Eucalypt (*Eucalyptus polybractea*)	Shallow whole genome sequencing; up to 500,000 SNPs	Foliar terpene yield traits	Kainer et al., 2018
Eucalypt (*Eucalyptus grandis x E. urophylla* - hybrids)	40,932 SNP EuCHIP60K and 55,772 capture probe SNP	Growth and wood quality	De Moraes et al., 2018

Studies that investigated different aspects of GS but used partially or totally the same breeding population data (genotypes and/or phenotypes) are listed in the same entry.

marker effects treated as random, normally distributed with common variance, has been very efficient. RR-BLUP has been equivalent to Genomic BLUP (GBLUP), providing the best conciliation between prediction efficiency and fast computation, while also revealing that essentially all major traits in forest trees fit the infinitesimal model (Resende et al., 2012c; Beaulieu et al., 2014b; Ratcliffe et al., 2015; Isik et al., 2016; Müller et al., 2017; Tan et al., 2017; Chen et al., 2018). Still, additional research in this area is warranted especially as prior functional information on genomic regions of slightly larger effect might emerge, for example, for disease resistance traits as shown for prediction of fusiform rust resistance in loblolly pine (Resende et al., 2012c).

Ever since the first experimental GS studies in forest trees (Grattapaglia et al., 2011; Resende et al., 2012a,b), it became clear that prediction accuracies are mainly driven by genetic relationship between training and validation sets and are dependent on G*E and age-age correlations. Predictions will be most effective at the same age and in the same environment where the prediction model was trained. Further studies in conifers (Zapata-Valenzuela et al., 2013; Beaulieu et al., 2014a,b; El-Dien et al., 2015; Ratcliffe et al., 2015; Thistlethwaite et al., 2017; Chen et al., 2018), and eucalypts (Müller et al., 2017; Tan et al., 2017), corroborated the key significance of genetic relationships and the impact of G*E and age-age correlations, consistent with findings in domestic animals and crop plants (Lin et al., 2014; Van Eenennaam et al., 2014). While data from G*E or age-age correlation studies will shed light on what to expect from genomic prediction, assuring that the target environment of future selection candidates will be equivalent to the one where models were originally trained is a challenging

issue for GS (Heslot et al., 2015). Regular retraining of GS models by incorporating phenotypes collected in breeding generations closer to the current (Iwata et al., 2011) are expected to mitigate this problem, and will be especially essential in light of climate fluctuations. Research efforts in this area are highly needed and will come as GS programs advance, coupled to innovations in phenotyping platforms that integrate remote sensing, spatial and geographic information systems (Dungey et al., 2018).

Notwithstanding the encouraging estimates of predictive ability, most studies in forest trees used contemporary training and validation sets and thus have not yet been able to adequately assess the realized performance of GS across generations at a larger scale, but results on this topic are imminent. However, given that the relationship between parents and progeny are accurately captured by DNA marker data, and environments should be relatively stable across close generations, it is expected that the performance will be equivalent to current estimates in contemporary sets. In *Pinus pinaster*, preliminary promising results of inter-generation prediction were reported by training models with parents and progeny in the same set (Isik et al., 2016), and later using parents and grandparents to predict in the subsequent generation, albeit with limited effective population sizes (Bartholome et al., 2016). However, this outcome was not observed in a three-generation study of *Pseudotsuga menziesii* (El-Kassaby, personal communication). Model updating strategies will therefore be crucial to counteract the decay of relatedness and LD between the original training set and selection candidates as generations of breeding advance, as shown by simulations for eucalypt breeding (Denis and Bouvet, 2013).

In the past 2 years, a number of additional experimental GS studies have been reported (Cappa et al., 2017, 2018; Duran et al., 2017; Lenz et al., 2017; Müller et al., 2017; Ratcliffe et al., 2017; Tan et al., 2017, 2018; Thistlethwaite et al., 2017) (**Table 1**; Resende et al., 2017b; Chen et al., 2018; De Moraes et al., 2018; El-Dien et al., 2018; Kainer et al., 2018; Suontama et al., 2018). Many of them in species of *Eucalyptus* for which public high-throughput genotyping platforms of DArT (Sansaloni et al., 2010) and SNPs (Silva-Junior et al., 2015) have been available. Access to such resources for eucalypts also allowed improved precision of genetic parameter estimates, pedigree reconstruction and inbreeding studies (Telfer et al., 2015; Klápště et al., 2017; Müller et al., 2017). This clearly points to the fact that the advancement of research and operational adoption of genomics into breeding is strongly dependent on the availability of public, robust, cost-accessible and portable SNP genotyping platforms. The success of GS or any other genomic-based breeding approach will rely on high data quality, as one has to be able to genotype SNPs across generations with high reproducibility and negligible missing data. Although shallow whole genome sequencing (Kainer et al., 2018), genotyping-by-sequencing (GbS) (El-Dien et al., 2015) and sequence capture (Thistlethwaite et al., 2017; Chen et al., 2018; De Moraes et al., 2018) have also been used for GS in trees, currently fixed SNP arrays provide the gold standard of data reproducibility across samples batches and laboratories. Additionally, SNP array data are breeder friendly, available from multiple service providers, easily manageable and

stored without the cost and logistics of sequence data transfer, storage and analysis. This and a significant recent drop in array costs, making them as cost-effective as sequence-based methods, has motivated a large international effort to develop SNP arrays for all main planted conifers (F. Isik pers. comm.), and a second generation, higher density optimized SNP array for species of *Eucalyptus* and *Corymbia* (O.B. Silva-Junior and D. Grattapaglia pers. comm.). The use of a common SNP genotyping array across breeding programs of different organizations will be a key issue to provide the necessary economy of scale to integrate genomics into breeding.

A LOOK TO THE NEAR FUTURE

With easy access to SNP genotyping and positive results in essentially all major forest trees, we are now at the brink of widespread adoption of genomic prediction data, thus realizing the early promises of MAS in forest tree breeding. In addition to the outstanding research challenges discussed above, a promising area to enhance the value of genomic data will involve the inclusion of environmental co-variables in GS models as already shown in crops (Jarquin et al., 2014; Saint Pierre et al., 2016). The integration of multi-environment trials data will be strategic for predicting performance in unobserved environments, identifying suitable sites for evaluating or deploying genetic material and predicting climate change scenarios. While predicting the performance of untested clones or families can be accurate when there is knowledge of genomic relatedness, correspondingly, the performance in yet unobserved or future environments could be forecasted if there is data about those environments as shown for recommendation of *Eucalyptus* clones (Marcatti et al., 2017). Resources such as ClimateNA (Wang et al., 2016) and the NASA POWER project (Stackhouse, 2014) offer multitudes of historical and predicted future environmental data. Because environmental variables that define the correlation between growing conditions are trait specific, research on those most appropriate for inclusion in genomic prediction models will be essential.

Another area that will demand research comes from the evolution of sequencing technologies in moving from sparse SNP data to sequence data for GS. Apart from the challenge and cost of managing massive next generation sequencing data sets for large numbers of individuals in a breeding program framework, in theory, if sequence data were used instead of dense SNPs, accuracy should increase because rare causal alleles would be better captured in predictive models. Until now, however, simulation and experimental studies in domestic animals have shown that whole-genome sequence data does not increase accuracy when LD has a slow decay pattern (Macleod et al., 2014; Forneris et al., 2017; Vanraden et al., 2017), unless very precise prior estimates on the functionality of particular SNPs exist (Perez-Enciso et al., 2015). Increasing the availability and quality of functional data on specific genomic regions might therefore, be warranted.

The success of whole-genome prediction and the poor outcome of dissection approaches in identifying functional

quantitative trait nucleotides, have contributed nevertheless to an exciting new perspective for the study of complex trait variation. A clear pattern has emerged in annual plants indicating that the association signal of common variants in large sample sizes, although spread across the entire genome, is heavily concentrated in regulatory DNA in open chromatin marked by deoxyribonuclease hypersensitive sites (Sullivan et al., 2014; Rodgers-Melnick et al., 2016; Swinnen et al., 2016). In these plants, cis-regulatory elements (CREs) associated with open chromatin such as promoters and enhancers regulating gene expression may contain close to half of all variants influencing traits. As GS implementation advances and large datasets of several thousand trees across unrelated populations are collected, opportunities will emerge for joint and meta-GWAS, as recently described in *Eucalyptus* (Müller et al., 2018). At the same time, chromatin accessibility and gene network data as reported for *Eucalyptus* (Hussey et al., 2017) and *Populus* (Zinkgraf et al., 2017) will become increasingly available which, combined with data from highly powered SNP-trait association studies, should provide new avenues for computational predictive discovery of key regulatory elements in the genome. The progress of such integrative approaches based on large genotype and phenotype datasets might, thus, result in additional clues toward understanding the complex connections and interactions between discrete genomic elements and continuous phenotypic trait variation, ultimately enhancing tree breeding practice.

AUTHOR CONTRIBUTIONS

DG drafted the first version of the manuscript and all co-authors subsequently contributed to it by editing and formatting the final version.

ACKNOWLEDGMENTS

REFERENCES

Bartholome, J., Van Heerwaarden, J., Isik, F., Boury, C., Vidal, M., Plomion, C., et al. (2016). Performance of genomic prediction within and across generations in maritime pine. *BMC Genomics* 17:604. doi: 10.1186/s12864-016-2879-8

Beaulieu, J., Doerksen, T., Boyle, B., Clement, S., Deslauriers, M., Beauseigle, S., et al. (2011). Association genetics of wood physical traits in the conifer white spruce and relationships with gene expression. *Genetics* 188, 1197–1214. doi: 10.1534/genetics.110.125781

Beaulieu, J., Doerksen, T., Clement, S., Mackay, J., and Bousquet, J. (2014a). Accuracy of genomic selection models in a large population of open-pollinated families in white spruce. *Heredity* 113, 343–352. doi: 10.1038/hdy.2014.36

Beaulieu, J., Doerksen, T. K., Mackay, J., Rainville, A., and Bousquet, J. (2014b). Genomic selection accuracies within and between environments and small breeding groups in white spruce. *BMC Genomics* 15:1048. doi: 10.1186/1471-2164-15-1048

Bernardo, R. (2008). Molecular markers and selection for complex traits in plants: learning from the last 20 years. *Crop Sci.* 48, 1649–1664. doi: 10.2135/cropsci2008.03.0131

Bouvet, J. M., Makouanzi, G., Cros, D., and Vigneron, P. (2016). Modeling additive and non-additive effects in a hybrid population using genome-wide genotyping: prediction accuracy implications. *Heredity* 116, 146–157. doi: 10.1038/hdy.2015.78

Cappa, E. P., El-Kassaby, Y. A., Garcia, M. N., Acuna, C., Borralho, N. M. G., Grattapaglia, D., et al. (2013). Impacts of population structure and analytical models in genome-wide association studies of complex traits in forest trees: a case study in *Eucalyptus globulus*. *PLoS ONE* 8:e81267. doi: 10.1371/journal.pone.0081267

Cappa, E. P., El-Kassaby, Y. A., Muñoz, F., Garcia, M. N., Villalba, P. V., Klápšt,ě, J., et al. (2017). Improving accuracy of breeding values by incorporating genomic information in spatial-competition mixed models. *Mol. Breed.* 37:125. doi: 10.1007/s11032-017-0725-6

Cappa, E. P., El-Kassaby, Y. A., Muñoz, F., Garcia, M. N., Villalba, P. V., Klápšt,ě, J., et al. (2018). Genomic-based multiple-trait evaluation in *Eucalyptus grandis* using dominant DArT markers. *Plant Sci.* 271, 27–33. doi: 10.1016/j.plantsci.2018.03.014

Chen, Z. Q., Baison, J., Pan, J., Karlsson, B., Gull, B. A., Westin, J., et al. (2018). Accuracy of genomic selection for growth and wood quality traits in two control-pollinated progeny trials using exome capture as

genotyping platform in Norway spruce. *bioRxiv [Preprint]*. doi: 10.1101/293696

Cumbie, W. P., Eckert, A., Wegrzyn, J., Whetten, R., Neale, D., and Goldfarb, B. (2011). Association genetics of carbon isotope discrimination, height and foliar nitrogen in a natural population of Pinus taeda L. *Heredity* 107, 105–114. doi: 10.1038/hdy.2010.168

De Almeida, J. E. D., Guimaraes, J. F. R., Silva, F. F. E., De Resende, M. D. V., Munoz, P., Kirst, M., et al. (2016). The contribution of dominance to phenotype prediction in a pine breeding and simulated population. *Heredity* 117, 33–41. doi: 10.1038/hdy.2016.23

De Moraes, B. F. X., Dos Santos, R. F., De Lima, B. M., Aguiar, A. M., Missiaggia, A. A., Da Costa Dias, D., et al. (2018). Genomic selection prediction models comparing sequence capture and SNP array genotyping methods. *Mol. Breed.* 38:115. doi: 10.1007/s11032-018-0865-3

Dekkers, J. C. M. (2004). Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons. *J. Anim. Sci.* 82, E313–E328. doi: 10.2527/2004.8213_supplE313x

Denis, M., and Bouvet, J. M. (2013). Efficiency of genomic selection with models including dominance effect in the context of Eucalyptus breeding. *Tree Genet. Genomes* 9, 37–51. doi: 10.1007/s11295-012-0528-1

Dungey, H. S., Dash, J. P., Pont, D., Clinton, P. W., Watt, M. S., and Telfer, E. J. (2018). Phenotyping whole forests will help to track genetic performance. *Trends Plant Sci.* 23, 854–864. doi: 10.1016/j.tplants.2018.08.005

Duran, R., Isik, F., Zapata-Valenzuela, J., Balocchi, C., and Valenzuela, S. (2017). Genomic predictions of breeding values in a cloned Eucalyptus globulus population in Chile. *Tree Genet. Genomes* 13:74. doi: 10.1007/s11295-017-1158-4

El-Dien, O. G., Ratcliffe, B., Klapste, J., Chen, C., Porth, I., and El-Kassaby, Y. A. (2015). Prediction accuracies for growth and wood attributes of interior spruce in space using genotyping-by-sequencing. *BMC Genomics* 16:370. doi: 10.1186/s12864-015-1597-y

El-Dien, O. G., Ratcliffe, B., Klápště, J., Porth, I., Chen, C., and El-Kassaby, Y. A. (2018). Multienvironment genomic variance decomposition analysis of open-pollinated Interior spruce (*Picea glauca* x engelmannii). *Mol. Breed.* 38:26. doi: 10.1007/s11032-018-0784-3

El-Kassaby, Y. A. (1982). Associations between Allozyme genotypes and quantitative traits in Douglas-Fir [*Pseudotsuga Menziesii* (Mirb.) Franco]. *Genetics* 101, 103–115.

Fahrenkrog, A. M., G., N.L., R., R.M.F., I., V. A., Gustavo, C., Christopher, D., Robert, S., Mark, D., et al. and Matias, K. (2017). Genome-wide association study reveals putative regulators of bioenergy traits in *Populus deltoides*. *N. Phytol.* 213, 799–811. doi: 10.1111/nph.14154

Falconer, D. S. (1989). *Introduction to Quantitative Genetics*. Essex: Longman Scientific and Technical.

Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edinb.* 52, 399–433. doi: 10.1017/S0080456800012163

Forneris, N. S., Vitezica, Z. G., Legarra, A., and Pérez-Enciso, M. (2017). Influence of epistasis on response to genomic selection using complete sequence data. *Genet. Select. Evol.* 49:66. doi: 10.1186/s12711-017-0340-3

Gion, J. M., Carouche, A., Deweer, S., Bedon, F., Pichavant, F., Charpentier, J. P., et al. (2011). Comprehensive genetic dissection of wood properties in a widely-grown tropical tree: eucalyptus. *Bmc Genomics* 12:301. doi: 10.1186/1471-2164-12-301

Gonzalez-Martinez, S. C., Wheeler, N. C., Ersoz, E., Nelson, C. D., and Neale, D. B. (2007). Association genetics in *Pinus taeda* L. I. Wood property traits. *Genetics* 175, 399–409. doi: 10.1534/genetics.106.061127

Grattapaglia, D. (2014). "Breeding forest trees by genomic selection: current progress and the way forward," in *Advances in Genomics of Plant Genetic Resources,* eds R. Tuberosa, A. Graner, and E. Frison (New York, NY: Springer), 652–682. doi: 10.1007/978-94-007-7572-5_26

Grattapaglia, D. (2017). "Status and perspectives of genomic selection in forest tree breeding," in *Genomic Selection for Crop Improvement: New Molecular Breeding Strategies for Crop Improvement*, eds R. K. Varshney, M. Roorkiwal, and M. E. Sorrells (Cham: Springer International Publishing), 199–249. doi: 10.1007/978-3-319-63170-7_9

Grattapaglia, D., Chaparro, J., Wilcox, P., Mccord, S., Werner, D., Amerson, H., et al. (1992). "Mapping in woody plants with RAPD markers: applications to breeding in forestry and horticulture," in *Proceedings of the Symposium Applications of RAPD Technology to Plant Breeding* (Minneapolis, MN: Crop Science Society of America, American Society of Horticultural Science, American Genetic Association), 37–40.

Grattapaglia, D., Plomion, C., Kirst, M., and Sederoff, R. R. (2009). Genomics of growth traits in forest trees. *Curr. Opin. Plant Biol.* 12, 148–156. doi: 10.1016/j.pbi.2008.12.008

Grattapaglia, D., and Resende, M. D. V. (2011). Genomic selection in forest tree breeding. *Tree Genet. Genomes* 7, 241–255. doi: 10.1007/s11295-010-0328-4

Grattapaglia, D., Resende, M. D. V., Resende, M., Sansaloni, C., Petroli, C., Missiaggia, A., et al. (2011). Genomic selection for growth traits in Eucalyptus: accuracy within and across breeding populations. *BMC Proc.* 5:O16. doi: 10.1186/1753-6561-5-S7-O16

Greenwood, M. S., Adams, G. W., and Gillespie, M. (1991). Stimulation of flowering by grafted black spruce and white spruce - a comparative-study of the effects of Gibberellin A4/7, cultural treatments, and environment. *Can. J. Forest Res. Revue Can. Rech. Forest.* 21, 395–400. doi: 10.1139/x91-049

Haley, C. S., and Visscher, P. M. (1998). Strategies to utilize marker-quantitative trait loci associations. *J. Dairy Sci.* 81, 85–97. doi: 10.3168/jds.S0022-0302(98)70157-2

Hasan, O., and Reid, J. B. (1995). Reduction of generation time in *Eucalyptus globulus*. *Plant Growth Regul.* 17, 53–60.

Hayes, B. J., Visscher, P. M., and Goddard, M. E. (2009). Increased accuracy of artificial selection by using the realized relationship matrix. *Genet. Res (Camb).* 91, 47–60. doi: 10.1017/S0016672308009981

Heslot, N., Jannink, J. L., and Sorrells, M. E. (2015). Perspectives for genomic selection applications and research in plants. *Crop Sci.* 55, 1–12. doi: 10.2135/cropsci2014.03.0249

Hussey, S. G., Loots, M. T., Van Der Merwe, K., Mizrachi, E., and Myburg, A. A. (2017). Integrated analysis and transcript abundance modelling of H3K4me3 and H3K27me3 in developing secondary xylem. *Sci. Rep.* 7:3370. doi: 10.1038/s41598-017-03665-1

Isik, F. (2014). Genomic selection in forest tree breeding: the concept and an outlook to the future. *N. Forests* 45, 379–401. doi: 10.1007/s11056-014-9422-z

Isik, F., Bartholome, J., Farjat, A., Chancerel, E., Raffin, A., Sanchez, L., et al. (2016). Genomic selection in maritime pine. *Plant Sci.* 242, 108–119. doi: 10.1016/j.plantsci.2015.08.006

Iwata, H., Hayashi, T., and Tsumura, Y. (2011). Prospects for genomic selection in conifer breeding: a simulation study of *Cryptomeria japonica*. *Tree Genet. Genomes* 7, 747–758. doi: 10.1007/s11295-011-0371-9

Jarquin, D., Crossa, J., Lacaze, X., Du Cheyron, P., Daucourt, J., Lorgeou, J., et al. (2014). A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theor. Appl. Genet.* 127, 595–607. doi: 10.1007/s00122-013-2243-1

Kainer, D., Stone, E. A., Padovan, A., Foley, W. J., and Külheim, C. (2018). accuracy of genomic prediction for foliar terpene traits in *Eucalyptus polybractea* G3 8, 2573–2583. doi: 10.1534/g3.118.200443

Kirst, M., Myburg, A., and Sederoff, R. (2004). Genetic mapping in forest trees: markers, linkage analysis and genomics. *Genet. Eng.* 26, 105–141. doi: 10.1007/978-0-306-48573-2_7

Klápště, J., Suontama, M., Telfer, E., Graham, N., Low, C., Stovold, T., et al. (2017). Exploration of genetic architecture through sib-ship reconstruction in advanced breeding population of Eucalyptus nitens. *PLoS ONE* 12:e0185137. doi: 10.1371/journal.pone.0185137

Lande, R., and Thompson, R. (1990). Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics* 124, 743–756.

Lander, E. S., and Botstein, D. (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121, 185–199.

Legarra, A., Aguilar, I., and Misztal, I. (2009). A relationship matrix including full pedigree and genomic information. *J. Dairy Sci.* 92, 4656–4663. doi: 10.3168/jds.2009-2061

Lenz, P. R. N., Beaulieu, J., Mansfield, S. D., Clement, S., Desponts, M., and Bousquet, J. (2017). Factors affecting the accuracy of genomic selection for growth and wood quality traits in an advanced-breeding population of black spruce (*Picea mariana*). *BMC Genomics* 18:335. doi: 10.1186/s12864-017-3715-5

Lima, B. M. (2014). *Bridging Genomics and Quantitative Genetics of Eucalyptus: Genome-Wide Prediction and Genetic Parameter Estimation for Growth and Wood Properties Using High-Density SNP Data*. Dissertation, Piracicaba. University of São Paulo. Available online at http://www.teses.usp.br/teses/disponiveis/11/11137/tde-25062014-085814/pt-br.php

Lin, Z., Hayes, B. J., and Daetwyler, H. D. (2014). Genomic selection in crops, trees and forages: a review. *Crop Pasture Sci.* 65, 1177–1191. doi: 10.1071/CP13363

Macleod, I. M., Hayes, B. J., and Goddard, M. E. (2014). The effects of demography and long-term selection on the accuracy of genomic prediction with sequence data. *Genetics* 198, 1671–1684. doi: 10.1534/genetics.114.168344

Marcatti, G. E., Resende, R. T., Resende, M. D. V., Ribeiro, C. A. A. S., Dos Santos, A. R., Cruz, J. P., et al. (2017). GIS-based approach applied to optimizing recommendations of Eucalyptus genotypes. *For. Ecol. Manage.* 392, 144–153. doi: 10.1016/j.foreco.2017.03.006

Mckown, A. D., Klapste, J., Guy, R. D., Geraldes, A., Porth, I., Hannemann, J., et al. (2014). Genome-wide association implicates numerous genes underlying ecological trait variation in natural populations of *Populus trichocarpa*. *N. Phytol.* 203, 535–553. doi: 10.1111/nph.12815

Meuwissen, T. H., Hayes, B. J., and Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829.

Müller, B. S. F., De Almeida Filho, J. E., Lima, B. M., Garcia, C. C., Missiaggia, A., Aguiar, A. M., et al. (2018). Independent and Joint-GWAS for growth traits in *Eucalyptus* by assembling genome-wide data for 3373 individuals across four breeding populations. *N. Phytol.* doi: 10.1111/nph.15449. [Epub ahead of print].

Müller, B. S. F., Neves, L. G., Almeida-Filho, J. E., Resende, M. F. R. J., Munoz Del Valle, P., Santos, P. E. T., et al. (2017). Genomic prediction in contrast to a genome-wide association study in explaining heritable variation of complex growth traits in breeding populations of Eucalyptus. *BMC Genomics* 18:524. doi: 10.1186/s12864-017-3920-2

Munoz, P. R., Resende, M. F. R., Gezan, S. A., Resende, M. D. V., De Los Campos, G., Kirst, M., et al. (2014). Unraveling additive from nonadditive effects using genomic relationship matrices. *Genetics* 198, 1759–1768. doi: 10.1534/genetics.114.171322

Namkoong, G., Kang, H. C., and Brouard, J. S. (1988). *Tree Breeding: Principles and Strategies*. New York, NY: Springer Verlag. doi: 10.1007/978-1-4612-3892-8

Neale, D. B., and Kremer, A. (2011). Forest tree genomics: growing resources and applications. *Nat. Revi. Genet.* 12, 111–122. doi: 10.1038/nrg2931

Neale, D. B., and Savolainen, O. (2004). Association genetics of complex traits in conifers. *Trends Plant Sci.* 9, 325–330. doi: 10.1016/j.tplants.2004.05.006

Neale, D. B., and Williams, C. G. (1991). Restriction-Fragment-Length-Polymorphism mapping in conifers and applications to forest genetics and tree improvement. *Can. J. Forest Res. Revue Can. Rech. Forest.* 21, 545–554. doi: 10.1139/x91-076

Nejati-Javaremi, A., Smith, C., and Gibson, J. P. (1997). Effect of total allelic relationship on accuracy of evaluation and response to selection. *J. Anim. Sci.* 75, 1738–1745. doi: 10.2527/1997.7571738x

Novaes, E., Osorio, L., Drost, D. R., Miles, B. L., Boaventura-Novaes, C. R. D., Benedict, C., et al. (2009). Quantitative genetic analysis of biomass and wood chemistry of Populus under different nitrogen levels. *N. Phytol.* 182, 878–890. doi: 10.1111/j.1469-8137.2009.02785.x

Perez-Enciso, M., Rincon, J. C., and Legarra, A. (2015). Sequence- vs. chip-assisted genomic selection: accurate biological information is advised. *Genet. Select. Evol.* 47:43. doi: 10.1186/s12711-015-0117-5

Porth, I., Klapste, J., Skyba, O., Hannemann, J., Mckown, A. D., Guy, R. D., et al. (2013). Genome-wide association mapping for wood characteristics in Populus identifies an array of candidate single nucleotide polymorphisms. *N. Phytol.* 200, 710–726. doi: 10.1111/nph.12422

Ratcliffe, B., El-Dien, O. G., Cappa, E. P., Porth, I., Klápště, J., Chen, C., et al. (2017). Single-step BLUP with varying genotyping effort in open-pollinated *Picea glauca. G3* 7, 935–942. doi: 10.1534/g3.116.037895

Ratcliffe, B., El-Dien, O. G., Klapste, J., Porth, I., Chen, C., Jaquish, B., et al. (2015). A comparison of genomic selection models across time in interior spruce (*Picea engelmannii* x glauca) using unordered SNP imputation methods. *Heredity* 115, 547–555. doi: 10.1038/hdy.2015.57

Resende, M. D. V., Resende, M. F. R., Sansaloni, C. P., Petroli, C. D., Missiaggia, A. A., Aguiar, A. M., et al. (2012a). Genomic selection for growth and wood quality in Eucalyptus: capturing the missing heritability and accelerating breeding for complex traits in forest trees. *N. Phytol.* 194, 116–128. doi: 10.1111/j.1469-8137.2011.04038.x

Resende, M. F. R., Munoz, P., Acosta, J. J., Peter, G. F., Davis, J. M., Grattapaglia, D., et al. (2012b). Accelerating the domestication of trees using genomic selection: accuracy of prediction models across ages and environments. *N. Phytol.* 193, 617–624. doi: 10.1111/j.1469-8137.2011.03895.x

Resende, M. F. R., Munoz, P., Resende, M. D. V., Garrick, D. J., Fernando, R. L., Davis, J. M., et al. (2012c). Accuracy of genomic selection methods in a standard data set of loblolly pine (*Pinus taeda* L.). *Genetics* 190, 1503–1510. doi: 10.1534/genetics.111.137026

Resende, R. T., Resende, M. D. V., Silva, F. F., Azevedo, C. F., Takahashi, E. K., Silva, O. B., et al. (2017b). Assessing the expected response to genomic selection of individuals and families in Eucalyptus breeding with an additive-dominant model. *Heredity* 119, 245–255. doi: 10.1038/hdy.2017.37

Resende, R. T., Resende, M. D. V., Silva, F. F., Azevedo, C. F., Takahashi, E. K., Silva-Junior, O. B., et al. (2017a). Regional heritability mapping and genome-wide association identify loci for complex growth, wood and disease resistance traits in Eucalyptus. *N. Phytol.* 213, 1287–1300. doi: 10.1111/nph.14266

Rodgers-Melnick, E., Vera, D. L., Bass, H. W., and Buckler, E. S. (2016). Open chromatin reveals the functional maize genome. *Proc. Natl. Acad. Sci. U.S.A.* 113, E3177–E3184. doi: 10.1073/pnas.1525244113

Saint Pierre, C., Burgueño, J., Crossa, J., Fuentes Dávila, G., Figueroa López, P., Solís Moya, E., et al. (2016). Genomic prediction models for grain yield of spring bread wheat in diverse agro-ecological zones. *Sci. Rep.* 6:27312. doi: 10.1038/srep27312

Sansaloni, C. P., Petroli, C. D., Carling, J., Hudson, C. J., Steane, D. A., Myburg, A. A., et al. (2010). A high-density Diversity Arrays Technology (DArT) microarray for genome-wide genotyping in *Eucalyptus. Plant Methods* 6:16. doi: 10.1186/1746-4811-6-16

Silva-Junior, O. B., Faria, D. A., and Grattapaglia, D. (2015). A flexible multi-species genome-wide 60K SNP chip developed from pooled resequencing 240 *Eucalyptus* tree genomes across 12 species. *N. Phytol.* 206, 1527–1540. doi: 10.1111/nph.13322

Stackhouse, P. W. (2014). "Prediction of worldwide energy resource," ed N. L. R. Ctr. (Hampton, VA). Available online at: http://power.larc.nasa.gov.

Strauss, S. H., Lande, R., and Namkoong, G. (1992). Limitations of molecular-marker-aided selection in forest tree breeding. *Can. J. Forest Res. Revue Can. Rech. Forest.* 22, 1050–1061. doi: 10.1139/x92-140

Sullivan, A. M., Arsovski, A. A., Lempe, J., Bubb, K. L., Weirauch, M. T., Sabo, P. J., et al. (2014). Mapping and dynamics of regulatory DNA and transcription factor networks in *A. thaliana. Cell Rep.* 8, 2015–2030. doi: 10.1016/j.celrep.2014.08.019

Suontama, M., Klápště, J., Telfer, E. J., Graham, N., Stovold, T., Low, C., et al. (2018). Efficiency of genomic prediction across two Eucalyptus nitens seed orchards with different selection histories. *Heredity.* doi: 10.1038/s41437-018-0119-5. [Epub ahead of print].

Swinnen, G., Goossens, A., and Pauwels, L. (2016). Lessons from domestication: targeting *Cis*-regulatory elements for crop improvement. *Trends Plant Sci.* 21, 506–515. doi: 10.1016/j.tplants.2016.01.014

Tan, B., Grattapaglia, D., Martins, G. S., Ferreira, K. Z., Sundberg, B., and Ingvarsson, P. K. (2017). Evaluating the accuracy of genomic prediction of growth and wood traits in two *Eucalyptus* species and their F1 hybrids. *BMC Genomics* 17:110. doi: 10.1186/s12870-017-1059-6

Tan, B., Grattapaglia, D., Wu, H. X., and Ingvarsson, P. K. (2018). Genomic relationships reveal significant dominance effects for growth in hybrid Eucalyptus. *Plant Sci.* 267, 84–93. doi: 10.1016/j.plantsci.2017.11.011

Telfer, E. J., Stovold, G. T., Li, Y., Silva-Junior, O. B., Grattapaglia, D. G., and Dungey, H. S. (2015). Parentage reconstruction in *Eucalyptus nitens* using SNPs and microsatellite markers: a comparative analysis of marker data power and robustness. *PLoS ONE* 10:e0130601. doi: 10.1371/journal.pone.0130601

Thistlethwaite, F. R., Ratcliffe, B., Klapste, J., Porth, I., Chen, C., Stoehr, M. U., et al. (2017). Genomic prediction accuracies in space and time for height and wood density of Douglas-fir using exome capture as the genotyping platform. *BMC Genomics* 18:930. doi: 10.1186/s12864-017-4258-5

Thumma, B. R., Nolan, M. R., Evans, R., and Moran, G. F. (2005). Polymorphisms in cinnamoyl CoA reductase (CCR) are associated with variation in microfibril angle in *Eucalyptus* spp. *Genetics* 171, 1257–1265. doi: 10.1534/genetics.105.042028

Thumma, B. R., Southerton, S. G., Bell, J. C., Owen, J. V., Henery, M. L., and Moran, G. F. (2010). Quantitative trait locus (QTL) analysis of wood quality traits in *Eucalyptus nitens. Tree Genet. Genomes* 6, 305–317. doi: 10.1007/s11295-009-0250-9

Ukrainetz, N., Ritland, K., and Mansfield, S. (2008). Identification of quantitative trait loci for wood quality and growth across eight full-sib coastal Douglas-fir families. *Tree Genet. Genomes* 4, 159–170. doi: 10.1007/s11295-007-0097-x

Van Eenennaam, A. L., Weigel, K. A., Young, A. E., Cleveland, M. A., and Dekkers, J. C. M. (2014). Applied animal genomics: results from the field. *Ann. Rev. Anim. Biosci.* 2, 105–139. doi: 10.1146/annurev-animal-022513-114119

Vanraden, P. M. (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91, 4414–4423. doi: 10.3168/jds.2007-0980

Vanraden, P. M., Tooker, M. E., O'connell, J. R., Cole, J. B., and Bickhart, D. M. (2017). Selecting sequence variants to improve genomic predictions for dairy cattle. *Genet. Select. Evol.* 49, 32. doi: 10.1186/s12711-017-0307-4

Wang, T., Hamann, A., Spittlehouse, D., and Carroll, C. (2016). Locally downscaled and spatially customizable climate data for historical and future periods for North America. *PLoS ONE* 11:e0156720. doi: 10.1371/journal.pone.0156720

White, T. L., Adams, W. T., and Neale, D. B. (2007). *Forest Genetics.* Cambridge, MA: CABI Publishing. p. 682 doi: 10.1079/9781845932855.0000

Williams, C. G. (1988). Accelerated short-term genetic testing for loblolly-pine families. *Can. J. Forest Res. Revue Can. Rech. Forest.* 18, 1085–1089. doi: 10.1139/x88-165

Williams, C. G., and Neale, D. B. (1992). Conifer wood quality and marker-aided selection - a case-study. *Can. J. Forest Res. Revue Can. Rech. Forest.* 22, 1009–1017. doi: 10.1139/x92-135

Zapata-Valenzuela, J., Isik, F., Maltecca, C., Wegrzyn, J., Neale, D., Mckeand, S., et al. (2012). SNP markers trace familial linkages in a cloned population of *Pinus*

taeda - prospects for genomic selection. *Tree Genet. Genomes* 6, 1307–1318. doi: 10.1007/s11295-012-0516-5

Zapata-Valenzuela, J., Whetten, R. W., Neale, D., Mckeand, S., and Isik, F. (2013). Genomic estimated breeding values using genomic relationship matrices in a cloned population of loblolly pine. *G3* 3, 909–916. doi: 10.1534/g3.113.0 05975

Zinkgraf, M., Liu, L., Groover, A., and Filkov, V. (2017). Identifying gene coexpression networks underlying the dynamic regulation of wood-forming tissues in Populus under diverse environmental conditions. *N. Phytol.* 214, 1464–1478. doi: 10.1111/nph. 14492

Manipulation of Growth and Architectural Characteristics in Trees for Increased Woody Biomass Production

Victor B. Busov*

School of Forest Resources and Environmental Science, Michigan Technological University, Houghton, MI, United States

Growth and architectural traits in trees are economically and environmentally important and thus of considerable importance to the improvement of forest and fruit trees. These traits are complex and result from the operation of a number of molecular mechanisms. This review will focus on the regulation of crown architecture, secondary woody growth and adventitious rooting. These traits and processes have significant impact on deployment, management, and productivity of tree crops. The majority of the described work comes from experiments in model plants, poplar, apple, peach, and plum because these species allow functional analysis of the involved genes and have significant genomics resources. However, these studies convincingly show conserved mechanisms for elaboration of specific growth and architectural traits. The conservation of these mechanisms suggest that they can be used as a blueprint for the improvement of these traits and processes in phylogenetically diverse tree crops. We will specifically consider the involvement of flowering time, transcription factors and hormone-associated genes. The review will also discuss the impact of recent technological advances as well as the challenges to the dissection of these traits in trees.

Keywords: hormones, transcription factors, woody biomass growth, molecular mechanisms, crown architecture, adventitious rooting, tree biotechnology

INTRODUCTION

Intensive forest plantation can alleviate the harvesting pressure on native forests via allowing production of the same or larger amount of wood on a much smaller land base (Paquette and Messier, 2010). Improved genetics through breeding is one, if not the leading factor in this increased productivity (Fenning and Gershenzon, 2002; Ruotsalainen, 2014). However, tree breeding is slow due to long generation times, traits that need a long time to evaluate and complex genetic architecture of these traits (Fenning and Gershenzon, 2002). Understanding the involved genetic mechanism could significantly accelerate the process through both conventional breeding and genetic engineering.

Here we review the current knowledge about the molecular mechanisms that underpin three developmental processes in trees with significant impact on intensive plantation deployment, management and growth. The review will focus on mechanisms and genes that can provide positive effects and thus are of breeding value rather than exhaustively discuss progress in the dissection of each process. Where available, the reader will be pointed to reviews that deal with these processes in a more comprehensive manner.

MANIPULATION OF CROWN ARCHITECTURE

Crown architecture is a compound trait resulting from the position, size, periodicity, angle and density of the branches. Crown characteristics affect plantation density, interception of photosynthetic light and quality of the derived wood. Depending on the plantation purpose, the direction and extent to which these characteristics need to be changed can vary. Branches originate in axillary meristems (AMs) and thus establishment and outgrowth of AMs has a profound effect on branch characteristics and crown architecture. AM initiation is exclusively characterized in herbaceous plants and there is no information about the effect of these genes in trees. We therefore will not cover here these developments. Excellent reviews on AM initiation discuss in detail these genes and mechanisms (Janssen et al., 2014; Yang and Jiao, 2016).

Branch Outgrowth

Once established, the AM outgrowth is typically suppressed, a phenomenon known as apical dominance. Auxin is central to the establishment and maintenance of apical dominance (**Figure 1**). The regulatory roles of auxin in apical dominance are indirect and are explained by the canalization and secondary messengers' models (Domagalska and Leyser, 2011; Teichmann and Muhr, 2015). However, only the latter provides genes and mechanisms manipulated in trees and is thus covered here. According to this model, auxin synthesized in the shoot apex, moves basipetally to the roots to generate a secondary signal that travels acropetally to regulate bud outgrowth. Cytokinin was the first candidate for a second messenger because it has a strong positive effect on axillary bud outgrowth when exogenously applied (**Figure 1**). However, cytokinin acropetal transport was not able to activate bud outgrowth (Faiss et al., 1997). This led to the discovery of the shoot branching hormone strigolactones (SLs). SLs have strong negative effects on bud outgrowth (**Figure 1**), are synthesized in roots, acropetally transported to shoots and biosynthetic genes are positively regulated by auxin. SLs metabolic and signaling genes are strong regulators of bud outgrowth in a number of plant species including several trees (Domagalska and Leyser, 2011; Muhr et al., 2016; Foster et al., 2018). RNAi knockdown of poplar and apple orthologs of SLs biosynthetic genes resulted in increased sylleptic branching (branches developed from lateral buds that have not undergone dormancy) (Muhr et al., 2016; Foster et al., 2018).

Branch Angle

Significant progress has been made in trees in elucidating the mechanism underpinning branch angle characteristics. Using an innovative sequencing approach and a distinct peach mutant with acute branch angle, the causative gene was isolated to be *TILLER ANGLE CONTROL1* (*TAC1*) (Dardick et al., 2013). TAC1 was originally found to control tiller angle in rice (Yu et al., 2007). *TAC1* belongs to a small family of genes. All the genes in the family, characterized to date in several plant species, including trees (poplar and plum) control branch or lateral root angles

(**Figure 1**; Xu et al., 2017; Hollender et al., 2018). Depending on presence of a conserved domain, members of the family can increase (*TAC1*) or decrease (*LAZY1*) branch angles (**Figure 1**; Dardick et al., 2013; Hollender and Dardick, 2015; Xu et al., 2017).

Roles of Gibberellins

Gibberellins control stem elongation, but can also regulate crown characteristics (**Figure 1**). *Gibberellin 2-oxidase* (*GA2ox*) overexpression leads to low levels of bioactive GAs, and proliferation of long sylleptic branches at a wide, almost perpendicular angle to the main stem (Mauriat et al., 2011; Zawaski et al., 2011). After 2 years in the field, *GA2ox* overexpressors produced a wide oval crown (Zawaski et al., 2011). A similar effect was also observed in turf grass and rice (Agharkar et al., 2007; Lo et al., 2008). These effects are possibly mediated via the GAs regulation of PIN auxin efflux carrier abundance (Willige et al., 2011; Lofke et al., 2013; Mauriat et al., 2014). In contrast, modifications of GA signaling via DELLA domain proteins produces a highly compact crown consisting of short branches with narrow acute angle (Zawaski et al., 2011). The effect of DELLA domain proteins on branching may be due their interactions with the transcription factor BRANCHED1 (Daviere et al., 2014).

Flowering and Crown Architecture

The determinacy of the meristem is genetically programmed, heritable and significantly affects plant architecture, including crown characteristics in trees (McGarry and Ayre, 2012). Indeterminate meristems typically produce monopodial growth characterized by a pronounced primary stem. In contrast, plants with determinate meristems show sympodial growth, a process of repeated loss of the shoot apical meristem (SAM) through terminal differentiation and lateral outgrowth from the axillary meristem resulting in a compound shoot architecture. Monopodial and sympodial growth types result from differences in the expression of genes and localization of proteins from the CENTRORDIALIS/TERMINAL FLOWER/SELF PRUNING (CETS) family that control flowering (McGarry and Ayre, 2012). CETS genes form a small gene family in Arabidopsis and other plant species. Very small (few amino acid) changes in the sequence of the proteins can reverse their function (Hanzawa et al., 2005). For example, FT promotes, while a close family member, TFL1 inhibits flowering (Hanano and Goto, 2011). FT is a mobile signal originating in the leaf that moves through the phloem stream to reach the shoot or axillary meristems and initiates terminal flower development (Pin and Nilsson, 2012). TFL1 plays an antagonistic role to FT in the SAM (Kobayashi et al., 1999). Low and high FT/TFL1 ratio in the SAM results in indeterminate and determinate growth respectively, (McGarry and Ayre, 2012; **Figure 1**). This model has been confirmed through transgenic overexpression of *FT* orthologs in several tree species (Hsu et al., 2006, 2011; Srinivasan et al., 2012; Klocko et al., 2016). FT overexpression leads to early flowering and highly branched, sympodial growth. Increase in FT/TFL1 balance via downregulation of TFL1/CEN genes in apple leads to similar effects as with *FT* overexpression (Kotoda et al., 2006;

FIGURE 1 | Regulation of crown architecture in trees. Diagram in the middle shows a WT tree with the different regulators and their putative roles indicated. The tree drawings to the left and right shows the effect of the different regulators on the crown architecture. Arrows show positive while blunted line negative effect. Vertical up red arrow indicates upregulation, while blue down arrow downregulation of the gene.

Flachowsky et al., 2012). RNAi downregulation of two *TFL1/CEN*-like homologs in poplar (*PopCEN1* and *PoCEN2*) produced similar but more moderate flowering and architectural phenotypes (Mohamed et al., 2010). Interacting factors and regulators of FT and TFL1/CEN, can also produce changes in tree architecture. Overexpression of CsRAV1, a chestnut ortholog of *TEMPRANILLO,* a regulator of FT (Castillejo and Pelaz, 2008), led to upregulation of a poplar ortholog of *FT* (*PttFT2*) expression (Triozzi et al., 2018) and consequently to increased branching both under greenhouse and field conditions (Moreno-Cortes et al., 2012, 2017). Under field conditions, increased branching led to increased biomass (Moreno-Cortes et al., 2017). Similarly, overexpression of a poplar ortholog of GIGANTEA, a positive regulator of *FT*, upregulated *PttFT2* and increased sylleptic branching in poplar (Ding et al., 2018). FT interacts with FD to promote flowering and overexpression of poplar *FD* homolog led to precocious flowering and sympodial, highly branched growth (Parmentier-Line and Coleman, 2015).

INCREASE OF SECONDARY WOODY GROWTH

Secondary growth originates in a lateral meristem known as vascular cambium, which in trees shows exaggerated and perennial activity, compared to herbaceous plants, resulting in production of massive amounts of conductive and supportive tissues, referred to as wood (Helariutta and Bhalerao, 2003; Elo et al., 2009; Barra-Jimenez and Ragni, 2017). Bifacial periclinal division of the cambium cells, followed by growth

and differentiation results in production of phloem/bark to the outside and xylem/wood to the inside of the tree trunk (Helariutta and Bhalerao, 2003; Zhang et al., 2014). Excellent reviews comprehensively discuss the process (Groover, 2005; Demura and Fukuda, 2007; Groover et al., 2010; Spicer and Groover, 2010; Mizrachi and Myburg, 2016). Here we focus on genes and mechanisms that have positive effects on secondary woody growth and thus are of potential breeding/improvement value (**Figure 2A**). The only exception would be the genes that affect bark development, Bark is typically considered as waist byproduct and thus decrease of bark production would be favored. However, notable exceptions where bark increase would be the goal would be special plantation for production of cork as well as breeding for resistance to pests, fires and drought.

Gibberellins

The first demonstration of increased secondary growth was via transgenic modifications of gibberellin biosynthesis (Eriksson et al., 2000). Overexpression of the Arabidopsis GA-20 oxidase (GA20ox), a key biosynthetic enzyme, resulted in significant (2-fold) increase in wood production (**Figure 2**). Similarly, overexpression of pine *PdGA20ox1* in poplar resulted in nearly 3-fold increase in woody biomass (Jeon et al., 2016; **Figure 2A**). In addition, overexpression of the poplar orthologs of the GA receptor *PttGIBBERELLIN-INSENSITIVE DWARF1* (*PttGID1*) resulted in similar wood biomass enhancement (Mauriat and Moritz, 2009; **Figure 2A**). Increase in bioactive gibberellins also increased fiber length and cellulose/xylan content (Eriksson et al., 2000; Jeon et al., 2016). Increase in GA signaling, however, did
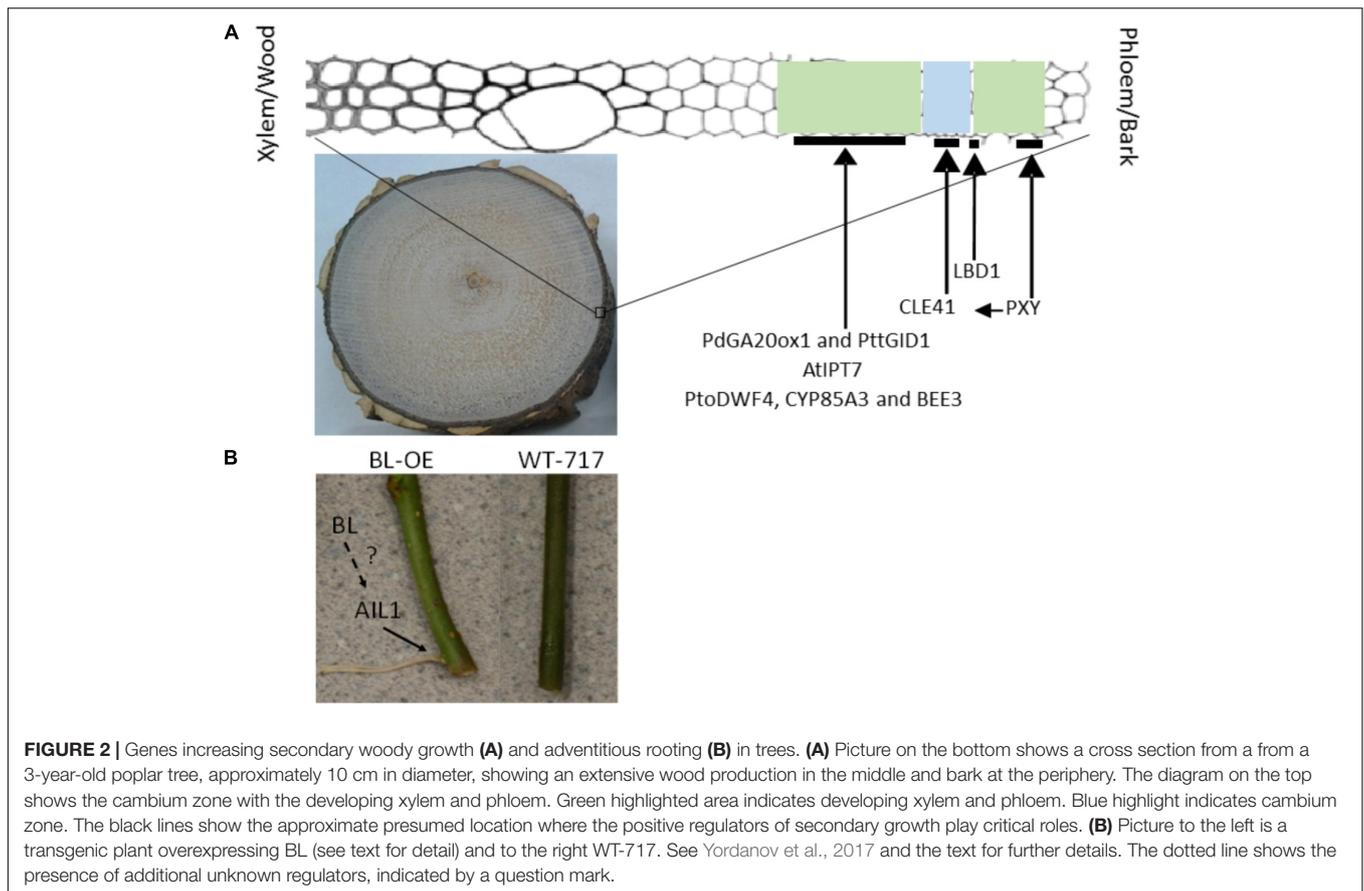
FIGURE 2 | Genes increasing secondary woody growth **(A)** and adventitious rooting **(B)** in trees. **(A)** Picture on the bottom shows a cross section from a from a 3-year-old poplar tree, approximately 10 cm in diameter, showing an extensive wood production in the middle and bark at the periphery. The diagram on the top shows the cambium zone with the developing xylem and phloem. Green highlighted area indicates developing xylem and phloem. Blue highlight indicates cambium zone. The black lines show the approximate presumed location where the positive regulators of secondary growth play critical roles. **(B)** Picture to the left is a transgenic plant overexpressing BL (see text for detail) and to the right WT-717. See Yordanov et al., 2017 and the text for further details. The dotted line shows the presence of additional unknown regulators, indicated by a question mark.

not result in changes in fiber length. The increased GA synthesis and signaling in the transgenics poplars, had negative effect on root development, decreased expression of defense-related genes and resulted in poor leaf development (Eriksson et al., 2000; Mauriat et al., 2014; Jeon et al., 2016). These negative pleiotropic effects resulting from the constitutive overexpression were mitigated by using a xylem-specific promoter (Jeon et al., 2016). The xylem specific expression resulted in similar increases of wood biomass (Jeon et al., 2016). However, tissue-specific upregulation of *PttGID1* using a different xylem-specific promoter did not result in any increases in woody biomass (Mauriat and Moritz, 2009), suggesting that different promoter::gene combination can have specific effects. Cisgenic modifications of several poplar GA 20-oxidase genes led to no pleiotropic effects and increased wood biomass and fiber length but the gains were more modest than these obtained with the constitutive and xylem-specific promoters (Han et al., 2011).

Cytokinins

The regulatory role(s) of cytokinins during secondary woody growth has been known (Nieminen et al., 2008). However, it was only recently demonstrated that modification of cytokinin biosynthesis can have a positive effect on secondary woody growth (Immanen et al., 2016). Transgenic poplars transformed with the Arabidopsis *AtIPT7* (key cytokinin biosynthetic gene)

driven by a xylem-specific promoter, showed significant (nearly 2-fold) increases in secondary growth and no negative pleiotropic effects (Immanen et al., 2016) (**Figure 2A**).

Brassinosteroids

Recent evidence suggests that both brassinosteroid (BR) biosynthesis and signaling has a positive effect on woody biomass production (Noh et al., 2015; Jin et al., 2017; Shen et al., 2018; **Figure 2**). Overexpression of key biosynthetic genes (*PtoDWF4* and *CYP85A3*) led to increased brassinosteroid concentrations and woody biomass (Jin et al., 2017; Shen et al., 2018). The productivity gains however, were much smaller than, these observed with the manipulations of GA and cytokinin biosynthesis/signaling. The increase in brassinosteroids led to longer fibers and no or little impact on cell wall chemistry (Jin et al., 2017; Shen et al., 2018). Similar results were obtained with manipulation of brassinosteroid signaling. Overexpression of a poplar ortholog of *BEE3* (*Brassinosteroid Enhanced Expression 3*), a transcription factor involved in BR signaling increased stem, leaf and root biomass. As with the enhancement of BR biosynthesis, gains in wood biomass were less than these observed with GA and cytokinin and ranged between 25 and 50%. The modifications of both BRs biosynthesis and signaling did not cause any negative pleiotropic effects, despite the strong constitutive promoters used in both studies.

Small Protein Signaling in the Cambium

Cambial cell division in Arabidopsis is controlled by a protein ligand receptor complex (Fisher and Turner, 2007; Hirakawa et al., 2008; Etchells and Turner, 2010). The ligand is the small CLE41 protein, produced in the phloem and transported into the cambium, where it interacts with the PXY receptor to stimulate cambium cell division (Fisher and Turner, 2007; Hirakawa et al., 2008; Etchells and Turner, 2010). Recently, constitutive expression of aspen orthologs of the ligand and receptor in transgenic poplar trees resulted in highly pleiotropic and negative effects on growth and tissue organization (Etchells et al., 2015). However, when the ligand and receptor were simultaneously upregulated in their native tissue domains employing tissue specific promoters, not only that the negative effects were completely mitigated, but also the double transgenic plants showed a nearly double increases in wood production (**Figure 2A**).

Secondary Phloem and Bark Development

Secondary growth also yields secondary phloem and bark. Using activation tagging in poplar, the first gene that regulates secondary phloem development was discovered (Yordanov et al., 2010). The gene encodes a transcription factor of the LATERAL ORGAN BOUNDARIES (LBD) gene family that is a positive regulator of secondary phloem development (Yordanov et al., 2010; Yordanov and Busov, 2011). Transgenic plants overexpressing the gene produced more, while dominant negative modification of the protein produced less secondary phloem (**Figure 2A**).

GENES PROMOTING ADVENTITIOUS ROOTING

Adventitious rooting (AR) is root formation from organs and tissues that typically do not produce roots. The process is most important in forestry and horticulture for clonal propagation and deployment of elite germplasm. The cellular and molecular events underlying AR has been reviewed elsewhere (Diaz-Sala, 2014; Legue et al., 2014; Pacurar et al., 2014). Here we focus on several genes that have been functionally characterized in trees and provide strong positive effects on AR formation.

Controls of Cell Proliferation Provide Points for AR Manipulation

AINTEGUMENTA (ANT) and ANT-like (AIL) genes are a group of eight AP2 transcription factors in Arabidopsis with important functions in regulation of meristem establishment and maintenance as well as organ growth and size (Horstman et al., 2014). One of the members of the *AIL* family from poplar (*AIL1*), showed induction during AR primordia activation (Rigal et al., 2012) and overexpression of the gene caused increase (**Figure 2B**), while RNAi downregulation decrease in the number of ARs. AIL1 transcriptionally regulates *Cyclin D3.1* by binding

to its promoter (Karlberg et al., 2011). Thus, AIL1 promotes AR at least in part via activation of cell proliferation.

The BIG LEAF/STERILE APETALA (BL/SAP) gene from poplar has a positive effect on AR formation when ectopically expressed (Yordanov et al., 2017; **Figure 2B**). BL is an F box protein that regulates leaf size in poplar and Arabidopsis through control of cell proliferation (Wang et al., 2016; Li et al., 2018). BL/SAP targets proteins for degradation that negatively regulate AIL genes (PLETHORA 1 and 2) (Horstman et al., 2014; Wang et al., 2016; Li et al., 2018). Thus, BL likely regulates AR formation through promoting degradation of a repressor(s)of the AIL-like genes (**Figure 2B**), which has a positive effect on cell proliferation and meristem organization.

Both *AIL1* and *BL*, when overexpressed have significant pleiotropic effects (Rigal et al., 2012; Yordanov et al., 2017) and to serve as biotechnological tool for increased AR formation, will require inducible or tissue-specific upregulation.

Gibberellins

GAs inhibit AR likely thought interfering with polar auxin transport (Mauriat et al., 2014). Increase and decrease in GA biosynthesis and signaling leads to decreased and increased AR (Busov et al., 2006; Gou et al., 2010; Elias et al., 2012). As mentioned earlier, GAs have strong positive effects on secondary woody growth and thus the decrease of AR may present an impediment for the clonal propagation of transgenics with increased GAs biosynthesis. Alternatively, decrease in bioactive GAs and block of signaling, which promotes AR formation, leads to various levels of dwarfism. Dwarfism is a desirable trait in fruit and ornamental tree crops and the increased AR formation would provide an additional benefit for the propagation of these genotypes. In forestry, however, extreme dwarfism may lead to loss in biomass productivity and thus, this effect can be either mitigated via increased girth growth using gene stacking with other transgenes that promote radial expansion (see above) or selection of semi-dwarfism genotypes (Elias et al., 2012).

Activation Tagging Discovery of AR-Involved Genes

Using activation tagging (AT), the poplar gene ETHYLENE RESPONSE FACTOR 3 (ERF003) was shown to have positive effect on AR formation (Trupiano et al., 2013). In addition to ERF003, several other AT mutants affected in AR and associated with ethylene signaling and biosynthesis were also discovered (Trupiano et al., 2013). These genes however, have not been recapitulated through re-transformation experiments and thus their involvement and utility in manipulation of AR formation is still tentative.

FUTURE OUTLOOK

Improvements in Transformation Technologies

Transformation is the golden standard for asserting gene function and preferred method of choice in delivering advanced

editing tools like CRIPSR/Cas9 system (Busov et al., 2005; Altpeter et al., 2016; Ran et al., 2017). However, transformation technologies are slow, inefficient, require cumbersome tissue culture processes and remain largely genotype-specific (Busov et al., 2005; Altpeter et al., 2016; Baltes et al., 2017), even in genera considered as 'easy-to-transform' like poplars. Thus, major strides in understanding and improving transformation technologies are needed (Altpeter et al., 2016).

Understanding Promoter Architecture

The need for research in isolation and engineering artificial promoters for precise targeting of transgenic manipulations has been known and well-recognized. However, research in this area has been lagging behind. The advances in gene editing and synthetic technologies would further necessitate better understanding promoter architecture in order to being able to effectively modify and design level and specificity of promoter activities.

Application of CRISPR Technology

New technological advances in gene editing technologies like CRISPR/Cas9 promise to revolutionize tree improvement (Tsai and Xue, 2015). The CRISPR/Cas9 was successfully implemented in a poplar tree (Fan et al., 2015; Zhou et al., 2015). CRISPR/Cas9 compared to RNAi produced stronger and more uniform phenotypic effects when the same gene was targeted (Zhou et al., 2015). Although now, the majority of the CRISPR/Cas9 applications involve generation of knock-outs via non-homologous end joining, a significant progress is also made in the application of CRISPR/Cas9 for gene editing through homologous recombination (Schaeffer and Nakata, 2015). However, the latter is still in developmental stages for plants. CRIPSR/Cas9 can also alleviate the regulatory burdens associated with field-testing because, in some countries CRISPR/Cas9-modified genotypes are considered as a non-GMO type of modification.

Using Induced and Natural Mutants

Application of natural or induced mutants in tree research has been rare. However, significant strides have been made in both approaches (Busov et al., 2010; Dardick et al., 2013). As described above, using activation tagging in poplar, genes important for secondary growth (Yordanov et al., 2010, 2014, 2017) and AR formation (Trupiano et al., 2013) were discovered. In addition to induced mutants, many natural tree mutants exist. The new sequencing technologies allow efficient mapping of the causative mutations, (Dardick et al., 2013). These approaches can be further used for identifying genes affecting various aspects of tree growth and development.

Understanding Integrative System Controls

A significant progress has been made in identification of individual genes and pathways regulating different traits. However, it has been long known at an organismal level that the various processes are highly coordinated at tissue and organismal level but also in response to various environmental cues. Identification of the coordinating genes, signals and mechanisms can lead to more integrative manipulation of one or several traits.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

REFERENCES

Agharkar, M., Lomba, P., Altpeter, F., Zhang, H., Kenworthy, K., and Lange, T. (2007). Stable expression of *AtGA2ox1* in a low-input turfgrass (*Paspalum notatum* Flugge) reduces bioactive gibberellin levels and improves turf quality under field conditions. *Plant Biotechnol. J.* 5, 791–801. doi: 10.1111/j.1467-7652.2007.00284.x

Altpeter, F., Springer, N. M., Bartley, L. E., Blechl, A. E., Brutnell, T. P., Citovsky, V., et al. (2016). Advancing crop transformation in the era of genome editing. *Plant Cell* 28, 1510–1520. doi: 10.1105/tpc.16.00196

Baltes, N. J., Gil-Humanes, J., and Voytas, D. F. (2017). Genome engineering and agriculture: opportunities and challenges. *Prog. Mol. Biol. Transl. Sci.* 149, 1–26. doi: 10.1016/bs.pmbts.2017.03.011

Barra-Jimenez, A., and Ragni, L. (2017). Secondary development in the stem: when Arabidopsis and trees are closer than it seems. *Curr. Opin. Plant Biol.* 35, 145–151. doi: 10.1016/j.pbi.2016.12.002

Busov, V., Meilan, R., Pearce, D. W., Rood, S. B., Ma, C., Tschaplinski, T. J., et al. (2006). Transgenic modification of gai or rgl1 causes dwarfing and alters gibberellins, root growth, and metabolite profiles in *Populus. Planta* 224, 288–299. doi: 10.1007/s00425-005-0213-9

Busov, V., Yordanov, Y., Gou, J., Meilan, R., Ma, C., Regan, S., et al. (2010). Activation tagging is an effective gene tagging system in *Populus. Tree Gen. Genom.* 7, 91–101. doi: 10.1007/s11295-010-0317-7

Busov, V. B., Brunner, A. M., Meilan, R., Filichkin, S., Ganio, L., Gandhi, S., et al. (2005). Genetic transformation: a powerful tool for dissection of adaptive traits in trees. *New Phytol.* 167, 9–18. doi: 10.1111/j.1469-8137.2005.01412.x

Castillejo, C., and Pelaz, S. (2008). The balance between CONSTANS and TEMPRANILLO activities determines FT expression to trigger flowering. *Curr. Biol.* 18, 1338–1343. doi: 10.1016/j.cub.2008.07.075

Dardick, C., Callahan, A., Horn, R., Ruiz, K. B., Zhebentyayeva, T., Hollender, C., et al. (2013). PpeTAC1 promotes the horizontal growth of branches in peach trees and is a member of a functionally conserved gene family found in diverse plants species. *Plant J.* 75, 618–630. doi: 10.1111/tpj.12234

Daviere, J. M., Wild, M., Regnault, T., Baumberger, N., Eisler, H., Genschik, P., et al. (2014). Class I TCP-DELLA interactions in inflorescence shoot apex determine plant height. *Curr. Biol.* 24, 1923–1928. doi: 10.1016/j.cub.2014.07.012

Demura, T., and Fukuda, H. (2007). Transcriptional regulation in wood formation. *Trends Plant Sci.* 12, 64–70. doi: 10.1016/j.tplants.2006.12.006

Diaz-Sala, C. (2014). Direct reprogramming of adult somatic cells toward adventitious root formation in forest tree species: the effect of the juvenile-adult transition. *Front. Plant Sci.* 5:310. doi: 10.3389/fpls.2014.00310

Ding, J., Bohlenius, H., Ruhl, M. G., Chen, P., Sane, S., Zambrano, J. A., et al. (2018). GIGANTEA-like genes control seasonal growth cessation in *Populus*. *New Phytol.* 218, 1491–1503. doi: 10.1111/nph.15087

Domagalska, M. A., and Leyser, O. (2011). Signal integration in the control of shoot branching. *Nat. Rev. Mol. Cell Biol.* 12, 211–221. doi: 10.1038/nrm3088

Elias, A. A., Busov, V. B., Kosola, K. R., Ma, C., Etherington, E., Shevchenko, O., et al. (2012). Green revolution trees: semidwarfism transgenes modify gibberellins, promote root growth, enhance morphological diversity, and reduce competitiveness in hybrid poplar. *Plant Physiol.* 160, 1130–1144. doi: 10.1104/pp.112.200741

Elo, A., Immanen, J., Nieminen, K., and Helariutta, Y. (2009). Stem cell function during plant vascular development. *Semin. Cell Dev. Biol.* 20, 1097–1106. doi: 10.1016/j.semcdb.2009.09.009

Eriksson, M. E., Israelsson, M., Olsson, O., and Moritz, T. (2000). Increased gibberellin biosynthesis in transgenic trees promotes growth, biomass production and xylem fiber length. *Nat. Biotechnol.* 18, 784–788. doi: 10.1038/77355

Etchells, J. P., Mishra, L. S., Kumar, M., Campbell, L., and Turner, S. R. (2015). Wood formation in trees is increased by manipulating PXY-regulated cell division. *Curr. Biol.* 25, 1050–1055. doi: 10.1016/j.cub.2015.02.023

Etchells, J. P., and Turner, S. R. (2010). The PXY-CLE41 receptor ligand pair defines a multifunctional pathway that controls the rate and orientation of vascular cell division. *Development* 137, 767–774. doi: 10.1242/dev.044941

Faiss, M., Zalubilova, J., Strnad, M., and Schmulling, T. (1997). Conditional transgenic expression of the ipt gene indicates a function for cytokinins in paracrine signaling in whole tobacco plants. *Plant J.* 12, 401–415. doi: 10.1046/j.1365-313X.1997.12020401.x

Fan, D., Liu, T., Li, C., Jiao, B., Li, S., Hou, Y., et al. (2015). Efficient CRISPR/Cas9-mediated targeted mutagenesis in Populus in the first generation. *Sci. Rep.* 5:12217. doi: 10.1038/srep12217

Fenning, T. M., and Gershenzon, J. (2002). Where will the wood come from? Plantation forests and the role of biotechnology. *Trends Biotechnol.* 20, 291–296. doi: 10.1016/s0167-7799(02)01983-2

Fisher, K., and Turner, S. (2007). PXY, a receptor-like kinase essential for maintaining polarity during plant vascular-tissue development. *Curr. Biol.* 17, 1061–1066. doi: 10.1016/j.cub.2007.05.049

Flachowsky, H., Szankowski, I., Waidmann, S., Peil, A., Trankner, C., and Hanke, M. V. (2012). The *MdTFL1* gene of apple (*Malus x domestica* Borkh.) reduces vegetative growth and generation time. *Tree Physiol.* 32, 1288–1301. doi: 10.1093/treephys/tps080

Foster, T. M., Ledger, S. E., Janssen, B. J., Luo, Z., Drummond, R. S. M., Tomes, S., et al. (2018). Expression of MdCCD7 in the scion determines the extent of sylleptic branching and the primary shoot growth rate of apple trees. *J. Exp. Bot.* 69, 2379–2390. doi: 10.1093/jxb/erx404

Gou, J., Strauss, S. H., Tsai, C. J., Fang, K., Chen, Y., Jiang, X., et al. (2010). Gibberellins regulate lateral root formation in Populus through interactions with auxin and other hormones. *Plant Cell* 22, 623–639. doi: 10.1105/tpc.109.073239

Groover, A. T. (2005). What genes make a tree a tree? *Trends Plant Sci.* 10, 210–214.

Groover, A. T., Nieminen, K., Helariutta, Y., and Mansfield, S. D. (2010). "Wood formation in Populus," in *Genetics and Genomics of Populus. Plant Genetics and Genomics: Crops and Models*, Vol. 8, eds S. Jansson, R. Bhalerao, and A. Groover (New York, NY: Springer), 201–224.

Han, K. M., Dharmawardhana, P., Arias, R. S., Ma, C., Busov, V., and Strauss, S. H. (2011). Gibberellin-associated cisgenes modify growth, stature and wood properties in Populus. *Plant Biotechnol. J.* 9, 162–178. doi: 10.1111/j.1467-7652.2010.00537.x

Hanano, S., and Goto, K. (2011). Arabidopsis TERMINAL FLOWER1 is involved in the regulation of flowering time and inflorescence development through transcriptional repression. *Plant Cell* 23, 3172–3184. doi: 10.1105/tpc.111.088641

Hanzawa, Y., Money, T., and Bradley, D. (2005). A single amino acid converts a repressor to an activator of flowering. *Proc. Natl. Acad. Sci. U.S.A.* 102, 7748–7753. doi: 10.1073/pnas.0500932102

Helariutta, Y., and Bhalerao, R. (2003). Between xylem and phloem: the genetic control of cambial activity in plants. *Plant Biol.* 5, 465–472. doi: 10.1055/s-2003-44780

Hirakawa, Y., Shinohara, H., Kondo, Y., Inoue, A., Nakanomyo, I., Ogawa, M., et al. (2008). Non-cell-autonomous control of vascular stem cell fate by a CLE peptide/receptor system. *Proc. Natl. Acad. Sci. U.S.A.* 105, 15208–15213. doi: 10.1073/pnas.0808444105

Hollender, C. A., and Dardick, C. (2015). Molecular basis of angiosperm tree architecture. *New Phytol.* 206, 541–556. doi: 10.1111/nph.13204

Hollender, C. A., Waite, J. M., Tabb, A., Raines, D., Chinnithambi, S., and Dardick, C. (2018). Alteration of *TAC1* expression in Prunus species leads to pleiotropic shoot phenotypes. *Hortic. Res.* 5:26. doi: 10.1038/s41438-018-0034-1

Horstman, A., Willemsen, V., Boutilier, K., and Heidstra, R. (2014). AINTEGUMENTA-LIKE proteins: hubs in a plethora of networks. *Trends Plant Sci.* 19, 146–157. doi: 10.1016/j.tplants.2013.10.010

Hsu, C. Y., Adams, J. P., Kim, H. J., No, K., Ma, C. P., Strauss, S. H., et al. (2011). FLOWERING LOCUS T duplication coordinates reproductive and vegetative growth in perennial poplar. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10756–10761. doi: 10.1073/pnas.1104713108

Hsu, C. Y., Liu, Y., Luthe, D. S., and Yuceer, C. (2006). Poplar FT2 shortens the juvenile phase and promotes seasonal flowering. *Plant Cell* 18, 1846–1861. doi: 10.1105/tpc.106.041038

Immanen, J., Nieminen, K., Smolander, O. P., Kojima, M., Alonso Serra, J., Koskinen, P., et al. (2016). Cytokinin and auxin display distinct but interconnected distribution and signaling profiles to stimulate cambial activity. *Curr. Biol.* 26, 1990–1997. doi: 10.1016/j.cub.2016.05.053

Janssen, B. J., Drummond, R. S., and Snowden, K. C. (2014). Regulation of axillary shoot development. *Curr. Opin. Plant Biol.* 17, 28–35. doi: 10.1016/j.pbi.2013.11.004

Jeon, H. W., Cho, J. S., Park, E. J., Han, K. H., Choi, Y. I., and Ko, J. H. (2016). Developing xylem-preferential expression of PdGA20ox1, a gibberellin 20-oxidase 1 from Pinus densiflora, improves woody biomass production in a hybrid poplar. *Plant Biotechnol. J.* 14, 1161–1170. doi: 10.1111/pbi.12484

Jin, Y. L., Tang, R. J., Wang, H. H., Jiang, C. M., Bao, Y., Yang, Y., et al. (2017). Overexpression of *Populus trichocarpa* CYP85A3 promotes growth and biomass production in transgenic trees. *Plant Biotechnol. J.* 15, 1309–1321. doi: 10.1111/pbi.12717

Karlberg, A., Bako, L., and Bhalerao, R. P. (2011). Short day-mediated cessation of growth requires the downregulation of AINTEGUMENTALIKE1 transcription factor in hybrid aspen. *PLoS Genet.* 7:e1002361. doi: 10.1371/journal.pgen.1002361

Klocko, A. L., Ma, C., Robertson, S., Esfandiari, E., Nilsson, O., and Strauss, S. H. (2016). FT overexpression induces precocious flowering and normal reproductive development in Eucalyptus. *Plant Biotechnol. J.* 14, 808–819. doi: 10.1111/pbi.12431

Kobayashi, Y., Kaya, H., Goto, K., Iwabuchi, M., and Araki, T. (1999). A pair of related genes with antagonistic roles in mediating flowering signals. *Science* 286, 1960–1962. doi: 10.1126/science.286.5446.1960

Kotoda, N., Iwanami, H., Takahashi, S., and Abe, K. (2006). Antisense expression of MdTFL1, a TFL1-like gene, reduces the juvenile phase in apple. *J. Am. Soc. Hortic. Sci.* 131, 74–81.

Legue, V., Rigal, A., and Bhalerao, R. P. (2014). Adventitious root formation in tree species: involvement of transcription factors. *Physiol. Plant.* 151, 192–198. doi: 10.1111/ppl.12197

Li, N., Liu, Z. P., Wang, Z. B., Ru, L. C., Gonzalez, N., Baekelandt, A., et al. (2018). STERILE APETALA modulates the stability of a repressor protein complex to control organ size in *Arabidopsis thaliana*. *PLoS Genet.* 14:e1007218. doi: 10.1371/journal.pgen.1007218

Lo, S. F., Yang, S. Y., Chen, K. T., Hsing, Y. L., Zeevaart, J. A. D., Chen, L. J., et al. (2008). A novel class of gibberellin 2-oxidases control semidwarfism, tillering, and root development in rice. *Plant Cell* 20, 2603–2618. doi: 10.1105/tpc.108.060913

Lofke, C., Zwiewka, M., Heilmann, I., Van Montagu, M. C., Teichmann, T., and Friml, J. (2013). Asymmetric gibberellin signaling regulates vacuolar trafficking of PIN auxin transporters during root gravitropism. *Proc. Natl. Acad. Sci. U.S.A.* 110, 3627–3632. doi: 10.1073/pnas.1300107110

Mauriat, M., and Moritz, T. (2009). Analyses of GA20ox- and GID1-over-expressing aspen suggest that gibberellins play two distinct roles in wood formation. *Plant J.* 58, 989–1003. doi: 10.1111/j.1365-313X.2009.03836.x

Mauriat, M., Petterle, A., Bellini, C., and Moritz, T. (2014). Gibberellins inhibit adventitious rooting in hybrid aspen and Arabidopsis by affecting auxin transport. *Plant J.* 78, 372–384. doi: 10.1111/tpj.12478

Mauriat, M., Sandberg, L. G., and Moritz, T. (2011). Proper gibberellin localization in vascular tissue is required to control auxin-dependent leaf development and bud outgrowth in hybrid aspen. *Plant J.* 67, 805–816. doi: 10.1111/j.1365-313X.2011.04635.x

McGarry, R. C., and Ayre, B. G. (2012). Manipulating plant architecture with members of the CETS gene family. *Plant Sci.* 18, 71–81. doi: 10.1016/j.plantsci.2012.03.002

Mizrachi, E., and Myburg, A. A. (2016). Systems genetics of wood formation. *Curr. Opin. Plant Biol.* 30, 94–100. doi: 10.1016/j.pbi.2016.02.007

Mohamed, R., Wang, C. T., Ma, C., Shevchenko, O., Dye, S. J., Puzey, J. R., et al. (2010). *Populus CEN/TFL1* regulates first onset of flowering, axillary meristem identity and dormancy release in *Populus. Plant J.* 62, 674–688. doi: 10.1111/j.1365-313X.2010.04185.x

Moreno-Cortes, A., Hernandez-Verdeja, T., Sanchez-Jimenez, P., Gonzalez-Melendi, P., Aragoncillo, C., and Allona, I. (2012). CsRAV1 induces sylleptic branching in hybrid poplar. *New Phytol.* 194, 83–90. doi: 10.1111/j.1469-8137.2011.04023.x

Moreno-Cortes, A., Ramos-Sanchez, J. M., Hernandez-Verdeja, T., Gonzalez-Melendi, P., Alves, A., Simoes, R., et al. (2017). Impact of RAV1-engineering on poplar biomass production: a short-rotation coppice field trial. *Biotechnol. Biofuels* 10:110. doi: 10.1186/s13068-017-0795-z

Muhr, M., Prufer, N., Paulat, M., and Teichmann, T. (2016). Knockdown of strigolactone biosynthesis genes in *Populus* affects *BRANCHED1* expression and shoot architecture. *New Phytol.* 212, 613–626. doi: 10.1111/nph.14076

Nieminen, K., Immanen, J., Laxell, M., Kauppinen, L., Tarkowski, P., Dolezal, K., et al. (2008). Cytokinin signaling regulates cambial development in poplar. *Proc. Natl. Acad. Sci. U.S.A.* 105, 20032–20037. doi: 10.1073/pnas.0805617106

Noh, S. A., Choi, Y. I., Cho, J. S., and Lee, H. (2015). The poplar basic helix-loop-helix transcription factor BEE3 - Like gene affects biomass production by enhancing proliferation of xylem cells in poplar. *Biochem. Biophys. Res. Commun.* 462, 64–70. doi: 10.1016/j.bbrc.2015.04.109

Pacurar, D. I., Perrone, I., and Bellini, C. (2014). Auxin is a central player in the hormone cross-talks that control adventitious rooting. *Physiol. Plant.* 151, 83–96. doi: 10.1111/ppl.12171

Paquette, A., and Messier, C. (2010). The role of plantations in managing the world's forests in the Anthropocene. *Front. Ecol. Environ.* 8, 27–34. doi: 10.1890/080116

Parmentier-Line, C. M., and Coleman, G. D. (2015). Constitutive expression of the Poplar FD-like basic leucine zipper transcription factor alters growth and bud development. *Plant Biotechnol. J.* 14, 260–270. doi: 10.1111/pbi.12380

Pin, P. A., and Nilsson, O. (2012). The multifaceted roles of FLOWERING LOCUS T in plant development. *Plant Cell Environ.* 35, 1742–1755. doi: 10.1111/j.1365-3040.2012.02558.x

Ran, Y., Liang, Z., and Gao, C. (2017). Current and future editing reagent delivery systems for plant genome editing. *Sci. China Life Sci.* 60, 490–505. doi: 10.1007/s11427-017-9022-1

Rigal, A., Yordanov, Y. S., Perrone, I., Karlberg, A., Tisserant, E., Bellini, C., et al. (2012). The AINTEGUMENTA LIKE1 homeotic transcription factor PtAIL1 controls the formation of adventitious root primordia in poplar. *Plant Physiol.* 160, 1996–2006. doi: 10.1104/pp.112.204453

Ruotsalainen, S. (2014). Increased forest production through forest tree breeding. *Scand. J. For. Res.* 29, 333–344. doi: 10.1080/02827581.2014.926100

Schaeffer, S. M., and Nakata, P. A. (2015). CRISPR/Cas9-mediated genome editing and gene replacement in plants: transitioning from lab to field. *Plant Sci.* 240, 130–142. doi: 10.1016/j.plantsci.2015.09.011

Shen, Y., Li, Y., Xu, D., Yang, C., Li, C., and Luo, K. (2018). Molecular cloning and characterization of a brassinosteroid biosynthesis-related gene PtoDWF4 from Populus tomentosa. *Tree Physiol.* doi: 10.1093/treephys/tpy027 [Epub ahead of print].

Spicer, R., and Groover, A. (2010). Evolution of development of vascular cambia and secondary growth. *New Phytol.* 186, 577–592. doi: 10.1111/j.1469-8137.2010.03236.x

Srinivasan, C., Dardick, C., Callahan, A., and Scorza, R. (2012). Plum (*Prunus domestica*) trees transformed with poplar FT1 result in altered architecture,

dormancy requirement, and continuous flowering. *PLoS One* 7:e40715. doi: 10.1371/journal.pone.0040715

Teichmann, T., and Muhr, M. (2015). Shaping plant architecture. *Front. Plant Sci.* 6:233. doi: 10.3389/fpls.2015.00233

Triozzi, P. M., Ramos-Sanchez, J. M., Hernandez-Verdeja, T., Moreno-Cortes, A., Allona, I., and Perales, M. (2018). Photoperiodic regulation of shoot apical growth in poplar. *Front. Plant Sci.* 9:1030. doi: 10.3389/fpls.2018.01030

Trupiano, D., Yordanov, Y., Regan, S., Meilan, R., Tschaplinski, T., Scippa, G. S., et al. (2013). Identification, characterization of an AP2/ERF transcription factor that promotes adventitious, lateral root formation in *Populus. Planta* 238, 271–282. doi: 10.1007/s00425-013-1890-4

Tsai, C. J., and Xue, L. J. (2015). CRISPRing into the woods. *GM Crops Food* 6, 206–215. doi: 10.1080/21645698.2015.1091553

Wang, Z., Li, N., Jiang, S., Gonzalez, N., Huang, X., Wang, Y., et al. (2016). SCF(SAP) controls organ size by targeting PPD proteins for degradation in *Arabidopsis thaliana. Nat. Commun.* 7:11192. doi: 10.1038/ncomms11192

Willige, B. C., Isono, E., Richter, R., Zourelidou, M., and Schwechheimer, C. (2011). Gibberellin regulates PIN-FORMED abundance and is required for auxin transport-dependent growth and development in *Arabidopsis thaliana. Plant Cell* 23, 2184–2195. doi: 10.1105/tpc.111.086355

Xu, D., Qi, X., Li, J., Han, X., Wang, J., Jiang, Y., et al. (2017). PzTAC and PzLAZY from a narrow-crown poplar contribute to regulation of branch angles. *Plant Physiol. Biochem.* 118, 571–578. doi: 10.1016/j.plaphy.2017.07.011

Yang, M., and Jiao, Y. (2016). Regulation of axillary meristem initiation by transcription factors and plant hormones. *Front. Plant Sci.* 7:183. doi: 10.3389/fpls.2016.00183

Yordanov, Y., Regan, S., and Busov, V. (2010). Members of the lateral organ boundaries domain (LBD) transcription factors family are involved in regulation of secondary growth in *Populus. Plant Cell* 22, 3662–3677. doi: 10.1105/tpc.110.078634

Yordanov, Y. S., and Busov, V. (2011). Boundary genes in regulation and evolution of secondary growth. *Plant Signal. Behav.* 6, 688–690. doi: 10.4161/psb.6.5.14973

Yordanov, Y. S., Ma, C., Strauss, S. H., and Busov, V. B. (2014). EARLY BUD-BREAK 1 (EBB1) is a regulator of release from seasonal dormancy in poplar trees. *Proc. Natl. Acad. Sci. U.S.A.* 111, 10001–10006. doi: 10.1073/pnas.1405621111

Yordanov, Y. S., Ma, C., Yordanova, E., Meilan, R., Strauss, S. H., and Busov, V. B. (2017). BIG LEAF is a regulator of organ size and adventitious root formation in poplar. *PLoS One* 12:e0180527. doi: 10.1371/journal.pone.0180527

Yu, B., Lin, Z., Li, H., Li, X., Li, J., Wang, Y., et al. (2007). TAC1, a major quantitative trait locus controlling tiller angle in rice. *Plant J.* 52, 891–898. doi: 10.1111/j.1365-313X.2007.03284.x

Zawaski, C., Kadmiel, M., Pickens, J., Ma, C., Strauss, S., and Busov, V. (2011). Repression of gibberellin biosynthesis or signaling produces striking alterations in poplar growth, morphology, and flowering. *Planta* 234, 1285–1298. doi: 10.1007/s00425-011-1485-x

Zhang, J., Nieminen, K., Serra, J. A., and Helariutta, Y. (2014). The formation of wood and its control. *Curr. Opin. Plant Biol.* 17, 56–63. doi: 10.1016/j.pbi.2013.11.003

Zhou, X., Jacobs, T. B., Xue, L. J., Harding, S. A., and Tsai, C. J. (2015). Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial Populus reveals 4-coumarate:CoA ligase specificity and redundancy. *New Phytol.* 208, 298–301. doi: 10.1111/nph.13470

Engineering Tree Seasonal Cycles of Growth Through Chromatin Modification

Daniel Conde[1†], Mariano Perales[1], Avinash Sreedasyam[2], Gerald A. Tuskan[3], Alba Lloret[4], María L. Badenes[4], Pablo González-Melendi[1,5], Gabino Ríos[4] and Isabel Allona[1,5]*

[1] Centro de Biotecnología y Genómica de Plantas, Instituto de Investigación y Tecnología Agraria y Alimentaria, Universidad Politécnica de Madrid, Madrid, Spain, [2] HudsonAlpha Institute for Biotechnology, Huntsville, AL, United States, [3] Oak Ridge National Laboratory, Center for Bioenergy Innovation, Oak Ridge, TN, United States, [4] Instituto Valenciano de Investigaciones Agrarias, Moncada, Spain, [5] Departamento de Biotecnología-Biología Vegetal, Escuela Técnica Superior de Ingeniería Agronómica, Alimentaria y de Biosistemas, Universidad Politécnica de Madrid, Madrid, Spain

In temperate and boreal regions, perennial trees arrest cell division in their meristematic tissues during winter dormancy until environmental conditions become appropriate for their renewed growth. Release from the dormant state requires exposure to a period of chilling temperatures similar to the vernalization required for flowering in *Arabidopsis*. Over the past decade, genomic DNA (gDNA) methylation and transcriptome studies have revealed signatures of chromatin regulation during active growth and winter dormancy. To date, only a few chromatin modification genes, as candidate regulators of these developmental stages, have been functionally characterized in trees. In this work, we summarize the major findings of the chromatin-remodeling role during growth-dormancy cycles and we explore the transcriptional profiling of vegetative apical bud and stem tissues during dormancy. Finally, we discuss genetic strategies designed to improve the growth and quality of forest trees.

Keywords: *Populus*, epigenetics, growth-dormancy, methylation, phenology, chromatin remodeling

INTRODUCTION

In temperate and boreal regions, a perennial plant's interannual life cycle comprises multiple vegetative growth, and dormancy cycles. To guarantee survival, trees synchronize their growth and flowering times with the most favorable climate conditions of the year by following instructive information of annual photoperiod, and temperature patterns (Ding and Nilsson, 2016; Singh et al., 2016). For example, prior to winter, cell division in meristematic tissues is arrested and a protective structure is formed, i.e., the apical bud, in which a quiescent shoot apical meristem (SAM) and embryonic leaves are sheltered during the winter. In several tree species, such as poplar (*Populus* sp.), photoperiod plays a major role in cell division arrest, and bud formation (Fennell and Hoover, 1991; Cooke et al., 2012; Petterle et al., 2013). Such trees are able to sense the shortening of day length and thus anticipate the winter period. In other tree species, such as apple, growth cessation, and bud formation are controlled by temperature (Tanino et al., 2010). Once endodormancy has been established, low non-lethal temperatures progressively lead to dormancy release (chilling requirement). Once fulfilled, dormancy is released while growth cessation is maintained via external signals (ecodormancy), mainly low temperatures (Ding and Nilsson, 2016).

Finally, spring growth-promoting temperatures produce bud break in vegetative buds, followed by vegetative growth.

These developmental processes require orchestration of specific temporal and spatial patterns of gene expression. Chromatin-modification-based regulation of gene expression during dormancy-growth cycles have been proposed to play a role in the organization of these patterns based on the identification of spatio-temporal patterns of epigenetic marks and the seasonal expression profiling of chromatin modification genes (Schrader et al., 2004; Ruttink et al., 2007; Santamaría et al., 2009; Karlberg et al., 2010; Conde et al., 2013, Conde et al., 2017a; Howe et al., 2015; Kumar et al., 2016). Epigenetic targets arise from covalent modifications of DNA and histones that will determine the accessibility of the transcription machinery to chromatin. Unlike in animals where epigenetics targets are established during embryonic development, in plants, epigenetic mechanisms also operate during post-embryonic developmental stages, contributing to plant developmental plasticity (Henderson and Jacobsen, 2007).

In this perspective article, we review the most recent evidences of DNA methylation and histone modification roles during annual growth-dormancy cycles in trees. In addition, we explore RNA-seq-based gene expression profiles in poplar vegetative apical bud and stem tissues, discovering seasonal expression patterns of genes involved in DNA methylation machinery. Finally, we discuss future strategies focused on chromatin remodeling for tree biotechnology applications.

DNA METHYLATION AND GROWTH-DORMANCY CYCLES

DNA Methylation Patterns During Winter Dormancy

Genomic DNA methylation refers to the addition of a methyl group to the carbon atom at the fifth position of a cytosine (5 mC). DNA methylation plays a major role in gene expression, genome protection and stability through transposon silencing, DNA recombination, and other biological processes (Teixeira and Colot, 2010; Mirouze et al., 2012; Saze et al., 2012). Variation in DNA methylation has impact on plants phenotypic plasticity (Bossdorf et al., 2008; Bräutigam et al., 2013; Kooke et al., 2015). Several studies have revealed DNA methylation patterns during growth-dormancy cycles, both in buds, and stems. In chestnut, Santamaría et al. (2009) quantified higher levels of gDNA methylation and lower H4 acetylation levels in vegetative dormant apical buds compared to actively growing apices. Accordingly, poplar stems showed higher levels of gDNA methylation and lower levels of acetylation of lysine K8 of histone H4 during winter dormancy compared to active growth (Conde et al., 2013). Kumar et al. (2016) found that DNA methylation levels in apple decreased gradually from flower bud dormancy to fruit set. This dynamics in apple buds was only observed when apple trees were grown in environmental conditions satisfying the chilling requirement for winter dormancy release (Kumar et al., 2016). A search of differentially methylated genes in

flower buds of almond by epi-Genotyping by sequencing (epi-GBS) led to a higher number of hypermethylated sequences in dormant buds when compared with dormancy-released samples (Prudencio et al., 2018). Recently, DNA methylation patterns during winter dormancy have been weekly examined in SAM tissue, from January to the time of vegetative bud break (Conde et al., 2017a). Results revealed a hypermethylation-hypomethylation wave formed by an initial stage of gDNA hypermethylation followed by a period of progressive 5 mC reduction to minimum levels before vegetative bud break and a 5 mC increase coincident with cell division reactivation (Conde et al., 2017a). Remarkably, a similar hypermethylation-hypomethylation wave has also been described in inflorescence SAM during the cold treatment of sugar beet (*Beta vulgaris*), suggesting comparable DNA methylation dynamics during vernalization and chilling requirements (Trap-Gentil et al., 2011; Conde et al., 2017a). Collectively, these findings suggest dynamic postembryonic deposition and removal of DNA methylation marks in SAM and stem tissues of woody perennials closely linked to the environmental factors.

DNA Methylation Machinery Profile During Winter Dormancy

To investigate in poplar how the DNA methylation machinery could create this winter hypermethylation-hypomethylation wave in vegetative SAM and stem tissues, we performed a RNA-seq mediated gene expression profiling on weekly collected vegetative apical buds and stems of hybrid poplar (*Populus tremula × alba* INRA clone 717 1B4), grown under natural conditions in Pozuelo de Alarcón, Madrid, over the period of January 13th to April 14th 2015 for apical buds, and from November 7th 2014 to April 9th 2015 for stems, coinciding with bud break. Weekly time points were grouped according to the Pearson correlation for samples within the groups and considered as a biological replicate. A detailed list of sample names, dates and groups is shown in the **Supplementary Table S1**. These analyses resulted in 6 groups for apical bud samples, from mid-winter to mid-spring, and 10 groups for stem samples, from late-fall to mid-spring. The expression data of this experiment can be found in Phytozome[1], under the expression tab for each gene.

Our RNAseq-based gene expression profiles revealed that the poplar homologs to *Arabidopsis* genes, involved in *de novo* DNA methylation machinery, such us *domains rearranged 2* (*DRM2*), are highly and constantly expressed from autumn to spring including winter dormancy in apical bud and stem tissues (**Figures 1A,B**). In contrast, a seasonal specific gene expression pattern was found for a plant specific 5-methyl-cytosine DEMETER-like demethylase (*DML10*). *DML10* showed a steady expression decline during early dormancy followed by a progressive increase in mRNA levels from mid-winter, with maximum expression observed at bud break in apical bud and stem tissues (**Figures 1A,B**). Similar results were reported in transcriptomics performed in poplar stem and lateral vegetative buds (Shim et al., 2014; Howe et al., 2015). In addition to that, poplar homologs to *methyltransferase 1* (*MET1*) and

---

[1] https://phytozome.jgi.doe.gov/pz/portal.html

FIGURE 1 | (A) Heat map showing normalized expression of genes coding for the enzymes involved in DNA methylation/demethylation during winter dormancy in hybrid poplar: *DML10* (Potri.010G234400), *CMT3* (Potri.001G009600), *MET1* (Potri.004G134000), *DRM2* (Potri.001G347000), *DRD1* (Potri.009G120700), *ARGO4* (Potri.001G219700), *DCL3* (Potri.018G047500), *RDR2* (Potri.015G073700), from mid-winter to mid-spring in apical bud, and from late-fall to mid-spring in stem. We added "early," "mid," or "late" to the name of groups when sampling dates were within the first, second, or third month of each season, following the Northern Meteorological Seasons dates. And "f1" or "f2," for the fortnight-1 and fortnight-2 within the month. **(B)** Line plot showing normalized expression during winter dormancy of *CMT3*, *DRM2,* and *DML10* in apical vegetative bud and stem tissues. **(C)** Diagram showing the pattern of overall DNA methylation during hybrid poplar dormancy (taken from Conde et al., 2017a) along with patterns of the genes involved in DNA methylation/demethylation during this period.

chromomethylase 3 (*CMT3*), that operate in CG and CHG contexts, were found induced just before the onset of bud break in apical bud and stem tissues (**Figures 1A,B**). According to Shim et al. (2014), *CMT3* is also relatively highly expressed during growth resumption and active growth, while *MET1* is more expressed during endodormancy and the start of ecodormancy in stem tissues. This difference in *MET1* expression could be explained by the different environmental conditions in which the two experiments were carried out.

Collectively, these gene expression patterns evoke a plausible scenario whereby DNA methylation levels gradually increase during winter dormancy, likely via the unbalance between *de novo* DNA methylation and demethylation activity induced by the downregulation of *DML10* during early dormancy (**Figure 1C**). Since 5 mC shows a maximal accumulation during winter dormancy, the progressive decline in 5 mC observed at the end of the dormancy period correlates with the induction of *DML10* mRNA observed here (**Figure 1C**). In addition, the 5 mC increase produced during bud break was concurrent with the induction of *MET1* and *CMT3*. These enzymes could contribute to the maintenance of heterochromatin and transposon methylation during

cell replication, once cell division has been reactivated in poplar SAM. Together, these findings highlight the main contribution of DML10 enzyme in generating winter specific 5 mC pattern.

In spite of the close agreement between DNA methylation levels and gene expression of methylation enzymes, we cannot discard putative developmental, and tissue-dependent effects on the transcriptional and post-transcriptional regulation of *DML10* and other epigenetic modifier genes, with potential impact on the methylation status of particular tissues and cells. In order to clarify this, tissue and cell-specific analyses would be very helpful.

DNA Methylation and Energy Status During Winter Dormancy

DNA methyltransferases use S-adenosyl methionine (SAMe) as a methyl group donor (Lyko, 2018). The end-product of this catalytic reaction is S-adenosyl homocysteine (SAH). SAH hydrolase (SAHH) breaks down SAH into adenosine and homocysteine, the precursor of methionine (Met), which in turn produces SAMe. SAHH activity is linked to cell metabolism, as SAHH activity is dependent of NAD$^+$

FIGURE 2 | Chromatin modifications produced in *DAM6* and *PpeS6PDH* genes during bud development in peach. Relative expression levels of *DAM6* and *PpeS6PDH* in flower buds are shown in the upper panel. Enriched histone modifications in dormant (blue area) and dormancy-released (green area) buds in *DAM6* and *PpeS6PDH* are depicted in the lower panel.
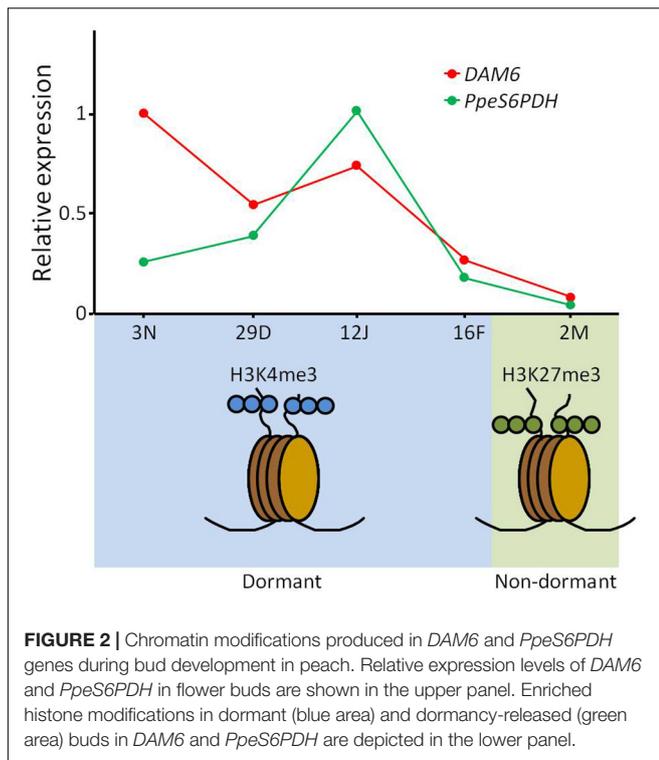
(Grillo and Colombatto, 2008). During glycolysis and TCA cycles, in which cells extract energy from glucose and pyruvate breakdown, molecules of NAD^+ are reduced into NADH, and hence a lower NAD^+/NADH ratio could diminish the activity of SAHH (Luo and Kuo, 2009). In apple, the activity of the TCA cycle enzyme isocitrate dehydrogenase (ICDH) was found to be low in dormant vegetative buds compared to non-dormant buds (Wang et al., 1991). Transcriptional profiling of *Paoenia ostii* highlighted the importance of glycolysis and TCA cycle induction for flower bud dormancy release (Gai et al., 2013). We suggest that SAHH activity could acts as the bridge between the energetic status and DNA methylation, contributing to the DNA methylation increase observed during winter dormancy, following the NAD^+/NADH ratio increment. We identified two poplar genes coding for SAHH enzymes: SAHH1 (Potri.001G320500) and SAHH17 (Potri.017G059400). Interestingly, both genes have been genetically associated with vegetative bud phenology during dormancy by Evans et al. (2014).

HISTONE MODIFICATION AND GROWTH-DORMANCY CYCLES

Gene-Specific Histone Marks in Buds

The modification of histones via different biochemical mechanisms and chromatin remodeling are key elements of plant development regulation and their response to environmental conditions (de la Paz Sanchez et al., 2015). In particular, histone modifications identified in *dormancy*

associated *MADS-box* (*DAM*) genes, the master regulators of vegetative and reproductive bud dormancy in Rosaceae and other perennial species, resemble chromatin dynamics of known regulators of vernalization, and seed dormancy and germination in *Arabidopsis* (Rios et al., 2014; Velappan et al., 2017). *DAM1-6* was first identified in peach as a family of six tandemly arrayed genes coding for similar MADS-Box transcription factors, which were partially deleted in the *evergrowing* non-dormant mutant of peach (Bielenberg et al., 2008). Related *DAM* genes involved in bud dormancy maintenance have been also described in leafy spurge (Horvath et al., 2010), Japanese apricot (Sasaki et al., 2011), pear (Niu et al., 2016), and apple (Wu et al., 2017) among other species. *DAM* genes have been postulated to modulate dormancy via transcriptional regulation of *FLOWERING LOCUS T* (Hsu et al., 2011) in vegetative buds of leafy spurge (Hao et al., 2015) and flower buds of pear (Niu et al., 2016), and also the biosynthesis of ABA in lateral flower buds of pear (Tuan et al., 2017). Concomitant with bud dormancy release, a decrease in the trimethylation of lysine 4 in histone H3 (H3K4me3) has been found in the chromatin of the *DAM1* gene in leafy spurge (Horvath et al., 2010), *DAM6* in peach (Leida et al., 2012), and *PpMADS13-1* in pear (Saito et al., 2015), suggesting that vegetative and flower buds share chromatin-related modifications across dormancy development in different species. Moreover, in peach, we observed a decrease in H3 acetylation on the ATG region of *DAM6*, and an increase in trimethylation of lysine 27 in H3 (H3K27me3) in a wider region of the gene (Leida et al., 2012). These chromatin changes are commonly associated with gene repression, and in fact, coincide with down-regulation of *DAM*-like genes, suggesting a mechanism for flower bud dormancy modulation through an arranged succession of epigenetic events in chromatin of *DAM* genes (Rios et al., 2014). In addition, small interference RNAs and microRNAs have been also postulated to regulate *DAM*-like expression and the floral dormancy transition in sweet cherry and pear, respectively (Niu et al., 2016; Rothkegel et al., 2017).

However, *DAM* genes are not the only known targets of chromatin modification during bud dormancy and development. The *early bud-break 1* (*EBB1*) gene encodes a putative APETALA2/Ethylene responsive transcription factor that reactivates vegetative meristem growth and bud-break after dormancy release in poplar (Yordanov et al., 2014), with conserved orthologs in other woody perennial species (Busov et al., 2016). Two regions on the promoter and ATG site of *PpEBB* gene from pear have been found differentially trimethylated at H3K4 in accordance with *PpEBB* up-regulation during the floral sprouting stage (Tuan et al., 2016). On the other hand, *PpeS6PDH* encodes a sorbitol-6-phosphate dehydrogenase involved in the synthesis of sorbitol in axillary flower dormant buds of peach (Lloret et al., 2017). *PpeS6PDH* expression was found to be silenced in dormancy-released buds concomitantly with an H3K4me3 decrease and H3K27me3 increase in a particular regulatory region of the gene near the translation start (**Figure 2**). We postulated that sorbitol exerts a role as cryoprotectant and compatible solute in dormant buds, and hence dormancy regulation by *DAM6* and abiotic stress

tolerance by *PpeS6PDH* could share a common mechanism for gene repression through concerted H3K27 trimethylation (Lloret et al., 2018).

Histone Modifiers During Winter Dormancy

Trimethylation of H3K27 is achieved by the polycomb repressive complex 2 (PRC2), containing components conserved in animals and plants (Alvarez-Venegas, 2010; Derkacheva and Hennig, 2014). In peach, several genes coding for such subunits of PRC2 complexes co-localize with quantitative trait loci for the chilling requirement and bloom date traits, providing genetic evidence of the role of these complexes in dormancy regulation (Zhebentyayeva et al., 2014). The PRC2 component gene *fertilization independent endosperm* (*FIE*) was sharply up-regulated under short-photoperiod treatments correlating with growth cessation and dormancy induction in poplar (Ruttink et al., 2007). Moreover, *FIE* suppression by RNAi prevented the establishment of dormancy in transgenic hybrid aspen (*Populus* × spp.), even though growth cessation and bud formation were not affected (Petterle, 2011). In peach buds, genome-wide stretches enriched in H3K27me3 were found associated with GA-repeat sequences (de la Fuente et al., 2015), suggesting that basic pentacysteine (BPC) factors able to bind GA-repeats could mediate the recruitment of PRC2 and thus H3K27me3 modification in flower bud dormancy dependent genes such as *DAM* family and *PpeS6PDH*, as was recently reported in *Arabidopsis* (Xiao et al., 2017).

Additional chromatin-related genes, such as the chromatin remodeler *PICKLE* (*PKL*), and putative modifiers involved in histone deacetylation (*HDA14* and *HDA08*), histone lysine methylation (*SUVR3*), and histone ubiquitination (*HUB2*) are also up-regulated during the transition to dormancy in *Populus* (Ruttink et al., 2007; Karlberg et al., 2010). Interestingly, down-regulation of *PKL* expression restores photoperiod-induced dormancy in *abi1-1* hybrid aspen mutants with a defective abscisic acid (ABA) response, suggesting that ABA promotes dormancy by repressing *PKL* (Tylewicz et al., 2018).

In axillary flower buds, dormancy changes are concomitant with flower developmental processes such as gametogenesis and organ development, which thus may account for part of the observed regulation of modifier genes and histone modifications. Strong changes in gene expression associated with microsporogenesis have been found in peach in parallel to dormancy release (Ríos et al., 2013), although no histone changes have been reported so far during this or other developmental processes in trees. However, in the model species *Arabidopsis* and rice, several epigenetic mechanisms involving chromatin remodeling and histone modification by PRC2 and other complexes modulate floral initiation and development at different steps (Guo et al., 2015), suggesting that processes other than dormancy may contribute to modify gene expression, DNA methylation, and histone modifications measurements in tree buds. Detailed tissue and organ-specific studies will be required in order to assess the contribution

of particular organs and processes to these biochemical and molecular observations.

BIOTECHNOLOGY OF THE EPIGENOME IN TREES

Modification of the plant epigenome contributes substantially to variation in plant growth, morphology, and plasticity (Johannes et al., 2008). In temperate and boreal trees, several lines of evidence point to environmental-guided DNA and histone modification profiles as critical regulators of chromatin function controlling the tempo of annual growth-dormancy cycles.

Recently, Kumar et al. (2016) reported that DNA demethylation in apple trees, which precedes floral bud break and fruit formation, only occurs under environments that fulfill the chilling requirement. Moreover, Conde et al. (2017a) observed that the induction of *DML10* expression before poplar bud break only happens if a chilling requirement has been fulfilled. These observations indicate that the downregulation of 5 mC observed during winter dormancy is closely linked to the transition from endodormancy to ecodormancy and is a precondition for growth resumption of vegetative and reproductive meristems. Therefore, modification of this DNA methylation pattern may have an impact on phenology as a biotechnological strategy to relax or tighten the chromatin state thus modifying the annual growth-dormancy cycle in trees.

Based on the hyper-hypomethylation wave of 5 mC reported during winter in poplar, it could be hypothesized that disruption of this pattern by creating hypermethylated poplar lines will delay growth resumption after dormancy. Conversely the generation of hypomethylated lines could result in rapid growth resumption following dormancy period. Candidate genes that alter winter dormancy DNA methylation/demethylation patterns can be inferred from our analyses. Thus, it could be possible to engineer poplars showing DNA hypermethylation through the upregulation of *DRM2* or silencing of *DML10* or *SAHH*, respectively. On the other hand, upregulation of DML10 or SAHH, or silencing of DRM2, respectively, could yield hypomethylated poplar lines. Accordingly, Conde et al. (2017a) described an RNAi strategy to reduce *DML* gene expression in hybrid poplars, noting that transgenic poplars featuring *DML10* downregulation showed significantly higher levels of DNA methylation, which delayed bud break. These RNAi *DML10* poplar lines showed negligible alteration of growth and development despite the specific effect mentioned. Equally, the overexpression of a chestnut *DML* (*Castanea sativa*) in hybrid poplar resulted in transgenic lines in which apical bud formation during dormancy establishment was accelerated showing no other visible alterations (Conde et al., 2017b). Hence, the consequences of epigenome engineering need to be tested gene by gene to check for possible pleiotropic phenotypes. Alternatively, a detailed knowledge of the tissue and time-dependent expression and activity of key modifier genes may provide useful information for the specific activation/repression of epigenetic regulators without undesirable effects on many other biological processes affected by them, with the help of adequate specific promoters.

The functional role of histone modifications during tree annual growth-dormancy cycles needs further clarification. The PRC2 complex seems to play a key role in dormancy regulation through the H3K27me3 modification on regulatory genes, although only the function of the PRC2 component FIE has been initially explored in poplar (Petterle, 2011). Tree orthologs of other subunits of this complex, such as *CURLY LEAF* (*CLF*) and *SWINGER* (*SWN*) could provide additional evidences on PCR2 participation in dormancy mechanisms and also serve as candidates for phenological manipulation. However, the putative targets of PCR2 activity may constitute more suitable and specific objectives for biotechnological approaches.

A genomic deletion of several *DAM* genes located in tandem causes a non-dormant phenotype in peach, highlighting the functional relevance of these genes in regulating dormancy (Bielenberg et al., 2008). As *DAM* genes have been postulated to integrate environmental signals, particularly chilling accumulation, by an epigenetic mechanism involving H3K27me3 and other histone modifications, targeted mutants on *DAM* *cis*-elements promoting the binding of histone modifier complexes could be employed to specifically modulate the response of *DAM* genes to chilling, with a potential use in manipulating the adaptability of stone fruit crops to changing climatic conditions. GA repeats on the large intron of *DAM6* may serve as tentative candidate *cis*-elements for this approach. In addition to GA repeats, the telobox *cis*-element has been shown to recruit the PRC2 complex in *Arabidopsis* through telomere-repeat-binding factors (TRBs) (Zhou et al., 2018). Thus, putative telobox motifs on *DAM* and other regulatory genes could also be modified in order to reprogram the environmental input on dormancy cycles.

Additional research is needed on the study of detailed characterization of the seasonal growth cycles involving frequently sampled intervals along the growth-dormancy-growth cycle. Detailed, subtle changes in chromatin modifications may be occurring outside the published data sets. In addition, new analytical techniques are emerging that may allow high-throughput characterization of DNA methylation and histone acetylation (e.g., ATAC-seq) should be applied to such periodically sampled tissues.

Other genes mentioned in this study involved in DNA methylation, RNA interference, chromatin remodeling, histone modification and transcriptional regulation of meristem growth, and dormancy are interesting candidates for biotechnological applications in tree phenology. The CRISPR/Cas9 system emerges as a promising technique due to its simplicity, design flexibility and high degree of efficiency. However, so far, few studies involving the use of this technique in tree species have been conducted, and no study has examined the impacts of epigenetics mark modifications (Fan et al., 2015; Zhou et al., 2015). Future studies designed to edit specific epigenetic regulator genes will unravel the impacts of particular epigenetic modifications on the annual cycle of trees.

AUTHOR CONTRIBUTIONS

DC, MP, AL, MB, PG-M, GR, and IA participated in the discussions described here. DC, MP, AS, GT, and IA implicated in the RNA-seq analysis. DC, MP, GT, GR, and IA wrote the manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.00412/full#supplementary-material

REFERENCES

Alvarez-Venegas, R. (2010). Regulation by polycomb and trithorax group proteins in *Arabidopsis*. *Arabidopsis Book* 8:e0128. doi: 10.1199/tab.0128

Bielenberg, D. G., Wang, Y., Li, Z., Zhebentyayeva, T., Fan, S., Reighard, G. L., et al. (2008). Sequencing and annotation of the evergrowing locus in peach [*Prunus persica* (L.) Batsch] reveals a cluster of six MADS-box transcription factors as candidate genes for regulation of terminal bud formation. *Tree Genet. Genomes* 4, 495–507. doi: 10.1007/s11295-007-0126-9

Bossdorf, O., Richards, C. L., and Pigliucci, M. (2008). Epigenetics for ecologists. *Ecol. Lett.* 11, 106–115. doi: 10.1111/j.1461-0248.2007.01130.x

Bräutigam, K., Vining, K. J., Lafon-Placette, C., Fossdal, C. G., Mirouze, M., Marcos, J. G., et al. (2013). Epigenetic regulation of adaptive responses of forest tree species to the environment. *Ecol. Evol.* 3, 399–415. doi: 10.1002/ece3.461

Busov, V., Carneros, E., and Yakovlev, I. (2016). EARLY BUD-BREAK1 (EBB1) defines a conserved mechanism for control of bud-break in woody perennials. *Plant Signal. Behav.* 11:e1073873. doi: 10.1080/15592324.2015.1073873

Conde, D., González-Melendi, P., and Allona, I. (2013). Poplar stems show opposite epigenetic patterns during winter dormancy and vegetative growth. *Trees Struct. Funct.* 27, 311–320. doi: 10.1007/s00468-012-0800-x

Conde, D., Le Gac, A.-L., Perales, M., Dervinis, C., Kirst, M., Maury, S., et al. (2017a). Chilling responsive DEMETER-LIKE DNA demethylase mediates in poplar bud break. *Plant Cell Environ.* 40, 2236–2249. doi: 10.1111/pce.13019

Conde, D., Moreno, A., Christopher, C., José, D., Sánchez, M. R., Kirst, M., et al. (2017b). Overexpression of DEMETER, a DNA demethylase, promotes early

apical bud maturation in poplar. *Plant Cell Environ.* 40, 2806–2819. doi: 10. 1111/pce.13056

Cooke, J. E. K., Eriksson, M. E., and Junttila, O. (2012). The dynamic nature of bud dormancy in trees: environmental control and molecular mechanisms. *Plant Cell Environ.* 35, 1707–1728. doi: 10.1111/j.1365-3040.2012.02552.x

de la Fuente, L., Conesa, A., Lloret, A., Badenes, M. L., and Ríos, G. (2015). Genome-wide changes in histone H3 lysine 27 trimethylation associated with bud dormancy release in peach. *Tree Genet. Genomes* 11:45. doi: 10.1007/s11295-015-0869-7

de la Paz Sanchez, M., Aceves-García, P., Petrone, E., Steckenborn, S., Vega-León, R., Álvarez-Buylla, E. R., et al. (2015). The impact of Polycomb group (PcG) and Trithorax group (TrxG) epigenetic factors in plant plasticity. *New Phytol.* 208, 684–694. doi: 10.1111/nph.13486

Derkacheva, M., and Hennig, L. (2014). Variations on a theme: polycomb group proteins in plants. *J. Exp. Bot.* 65, 2769–2784. doi: 10.1093/jxb/ert410

Ding, J., and Nilsson, O. (2016). Molecular regulation of phenology in trees-because the seasons they are a-changin'. *Curr. Opin. Plant Biol.* 29, 73–79. doi: 10.1016/j.pbi.2015.11.007

Evans, L. M., Slavov, G. T., Rodgers-Melnick, E., Martin, J., Ranjan, P., Muchero, W., et al. (2014). Population genomics of *Populus trichocarpa* identifies signatures of selection and adaptive trait associations. *Nat. Genet.* 46, 1089–1096. doi: 10.1038/ng.3075

Fan, D., Liu, T., Li, C., Jiao, B., Li, S., Hou, Y., et al. (2015). Efficient CRISPR/Cas9-mediated targeted mutagenesis in Populus in the first generation. *Sci. Rep.* 5:12217. doi: 10.1038/srep12217

Fennell, A., and Hoover, E. (1991). Photoperiod influences growth bud dormancy and cold acclimation in *Vitis labruscana* and *Vitis riparia*. *J. Am. Soc. Hortic. Sci.* 116, 270–273. doi: 10.21273/jashs.116.2.270

Gai, S., Zhang, Y., Liu, C., Zhang, Y., and Zheng, G. (2013). Transcript profiling of *Paeonia ostii* during artificial chilling induced dormancy release identifies activation of GA pathway and carbohydrate metabolism. *PLoS One* 8:e55297. doi: 10.1371/journal.pone.0055297

Grillo, M. A., and Colombatto, S. (2008). S-adenosylmethionine and its products. *Amino Acids* 34, 187–193. doi: 10.1007/s00726-007-0500-9

Guo, S., Sun, B., Looi, L. S., Xu, Y., Gan, E. S., Huang, J., et al. (2015). Co-ordination of flower development through epigenetic regulation in two model species: rice and *Arabidopsis*. *Plant Cell Physiol.* 56, 830–842. doi: 10.1093/pcp/pcv037

Hao, X., Chao, W., Yang, Y., and Horvath, D. (2015). Coordinated expression of FLOWERING LOCUS T and DORMANCY ASSOCIATED MADS-BOX-like genes in leafy spurge. *PLoS One* 10:e0126030. doi: 10.1371/journal.pone.0126030

Henderson, I. R., and Jacobsen, S. E. (2007). Epigenetic inheritance in plants. *Nature* 447, 418–424. doi: 10.1038/nature05917

Horvath, D. P., Sung, S., Kim, D., Chao, W., and Anderson, J. (2010). Characterization, expression and function of DORMANCY ASSOCIATED MADS-BOX genes from leafy spurge. *Plant Mol. Biol.* 73, 169–179. doi: 10.1007/s11103-009-9596-5

Howe, G. T., Horvath, D. P., Dharmawardhana, P., Priest, H. D., Mockler, T. C., and Strauss, S. H. (2015). Extensive transcriptome changes during natural onset and release of vegetative bud dormancy in Populus. *Front. Plant Sci.* 6:989. doi: 10.3389/fpls.2015.00989

Hsu, C. Y., Adams, J. P., Kim, H., No, K., Ma, C., and Strauss, S. H. (2011). FLOWERING LOCUS T duplication coordinates reproductive and vegetative growth in perennial poplar. *Proc. Natl. Aacd. Sci. U.S.A.* 108, 10756–10761. doi: 10.1073/pnas.1104713108

Johannes, F., Colot, V., and Jansen, R. C. (2008). Epigenome dynamics: a quantitative genetics perspective. *Nat. Rev. Genet.* 9, 883–890. doi: 10.1038/nrg2467

Karlberg, A., Englund, M., Petterle, A., Molnar, G., Sjödin, A., Bako, L., et al. (2010). Analysis of global changes in gene expression during activity-dormancy cycle in hybrid aspen apex. *Plant Biotechnol.* 27, 1–16. doi: 10.5511/plantbiotechnology.27.1

Kooke, R., Johannes, F., Wardenaar, R., Becker, F., Etcheverry, M., Colot, V., et al. (2015). Epigenetic basis of morphological variation and phenotypic plasticity in *Arabidopsis thaliana*. *Plant Cell* 27, 337–348. doi: 10.1105/tpc.114.133025

Kumar, G., Rattan, U. K., and Singh, A. K. (2016). Chilling-mediated DNA methylation changes during dormancy and its release reveal the importance of epigenetic regulation during winter dormancy in apple (*Malus* x *domestica* Borkh.). *PLoS One* 11:e0149934. doi: 10.1371/journal.pone.0149934

Leida, C., Conesa, A., Llácer, G., Badenes, M. L., and Ríos, G. (2012). Histone modifications and expression of DAM6 gene in peach are modulated during bud dormancy release in a cultivar-dependent manner. *New Phytol.* 193, 67–80. doi: 10.1111/j.1469-8137.2011.03863.x

Lloret, A., Badenes, M. L., and Ríos, G. (2018). Modulation of dormancy and growth responses in reproductive buds of temperate trees. *Front. Plant Sci.* 9:1368. doi: 10.3389/fpls.2018.01368

Lloret, A., Martínez-Fuentes, A., Agustí, M., Badenes, M. L., and Ríos, G. (2017). Chromatin-associated regulation of sorbitol synthesis in flower buds of peach. *Plant Mol. Biol.* 95, 507–517. doi: 10.1007/s11103-017-0669-6

Luo, J., and Kuo, M.-H. (2009). Linking nutrient metabolism to epigenetics. *Cell Sci. Rev.* 6. 1.

Lyko, F. (2018). The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. *Nat. Rev. Genet.* 19, 81–92. doi: 10.1038/nrg.2017.80

Mirouze, M., Lieberman-Lazarovich, M., Aversano, R., Bucher, E., Nicolet, J., Reinders, J., et al. (2012). Loss of DNA methylation affects the recombination landscape in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 109, 5880–5885. doi: 10.1073/pnas.1120841109

Niu, Q., Li, J., Cai, D., Qian, M., Jia, H., Bai, S., et al. (2016). Dormancy-associated MADS-box genes and microRNAs jointly control dormancy transition in pear (*Pyrus pyrifolia* white pear group) flower bud. *J. Exp. Bot.* 67, 239–257. doi: 10.1093/jxb/erv454

Petterle, A. (2011). *ABA and Chromatin Remodeling Regulate the Activity-Dormancy Cycle in Hybrid Aspen*. Ph.D. thesis, Swedish University of Agricultural Sciences, Uppsala.

Petterle, A., Karlberg, A., and Bhalerao, R. P. (2013). Daylength mediated control of seasonal growth patterns in perennial trees. *Curr. Opin. Plant Biol.* 16, 301–306. doi: 10.1016/j.pbi.2013.02.006

Prudencio, Á. S., Werner, O., Martínez-García, P. J., Dicenta, F., Ros, R. M., and Martínez-Gómez, P. (2018). DNA methylation analysis of dormancy release in almond (*Prunus dulcis*) flower buds using epi-genotyping by sequencing. *Int. J. Mol. Sci.* 19:E3542. doi: 10.3390/ijms19113542

Rios, G., Leida, C., Conejero, A., and Badenes, M. L. (2014). Epigenetic regulation of bud dormancy events in perennial plants. *Front. Plant Sci.* 5:247. doi: 10.3389/fpls.2014.00247

Ríos, G., Tadeo, F. R., Leida, C., and Badenes, M. L. (2013). Prediction of components of the sporopollenin synthesis pathway in peach by genomic and expression analyses. *BMC Genomics* 14:40. doi: 10.1186/1471-2164-14-40

Rothkegel, K., Sánchez, E., Montes, C., Greve, M., Tapia, S., Bravo, S., et al. (2017). DNA methylation and small interference RNAs participate in the regulation of MADS-box genes involved in dormancy in sweet cherry (*Prunus avium* L.). *Tree Physiol.* 37, 1739–1751. doi: 10.1093/treephys/tpx055

Ruttink, T., Arend, M., Morreel, K., Storme, V., Rombauts, S., Fromm, J., et al. (2007). A molecular timetable for apical bud formation and dormancy induction in poplar. *Plant Cell* 19, 2370–2390. doi: 10.1105/tpc.107.052811

Saito, T., Bai, S., Imai, T., Ito, A., Nakajima, I., and Moriguchi, T. (2015). Histone modification and signalling cascade of the dormancy-associated MADS-box gene, PpMADS13-1, in Japanese pear (*Pyrus pyrifolia*) during endodormancy. *Plant Cell Environ.* 38, 1157–1166. doi: 10.1111/pce.12469

Santamaría, M., Hasbún, R., Valera, M., Meijón, M., Valledor, L., Rodríguez, J. L., et al. (2009). Acetylated H4 histone and genomic DNA methylation patterns during bud set and bud burst in *Castanea sativa*. *J. Plant Physiol.* 166, 1360–1369. doi: 10.1016/j.jplph.2009.02.014

Sasaki, R., Yamane, H., Ooka, T., Jotatsu, H., Kitamura, Y., Akagi, T., et al. (2011). Functional and expressional analyses of PmDAM genes associated with endodormancy in Japanese apricot. *Plant Physiol.* 157, 485–497. doi: 10.1104/pp.111.181982

Saze, H., Tsugane, K., Kanno, T., and Nishimura, T. (2012). DNA methylation in plants: relationship to small RNAs and histone modifications, and functions in transposon inactivation. *Plant Cell Physiol.* 53, 766–784. doi: 10.1093/pcp/pcs008

Schrader, J., Moyle, R., Bhalerao, R., Hertzberg, M., Lundeberg, J., Nilsson, P., et al. (2004). Cambial meristem dormancy in trees involves extensive remodelling of the transcriptome. *Plant J.* 40, 173–187. doi: 10.1111/j.1365-313X.2004.02199.x

Shim, D., Ko, J.-H., Kim, W.-C., Wang, Q., Keathley, D. E., and Han, K.-H. (2014). A molecular framework for seasonal growth-dormancy regulation in perennial plants. *Hortic. Res.* 1:14059. doi: 10.1038/hortres.2014.59

Singh, R. K., Svystun, T., Aldahmash, B., Jonsson, A. M., and Bhalerao, R. P. (2016). Photoperiod- and temperature-mediated control of phenology in trees – a molecular perspective. *New Phytol.* 213, 511–524. doi: 10.1111/nph.14346

Tanino, K. K., Kalcsits, L., Silim, S., Kendall, E., and Gray, G. R. (2010). Temperature-driven plasticity in growth cessation and dormancy development in deciduous woody plants: a working hypothesis suggesting how molecular and cellular function is affected by temperature during dormancy induction. *Plant Mol. Biol.* 73, 49–65. doi: 10.1007/s11103-010-9610-y

Teixeira, F. K., and Colot, V. (2010). Repeat elements and the *Arabidopsis* DNA methylation landscape. *Heredity* 105, 14–23. doi: 10.1038/hdy.2010.52

Trap-Gentil, M. V., Hébrard, C., Lafon-Placette, C., Delaunay, A., Hagège, D., Joseph, C., et al. (2011). Time course and amplitude of DNA methylation in the shoot apical meristem are critical points for bolting induction in sugar beet and bolting tolerance between genotypes. *J. Exp. Bot.* 62, 2585–2597. doi: 10.1093/jxb/erq433

Tuan, P. A., Bai, S., Saito, T., Imai, T., Ito, A., and Moriguchi, T. (2016). Involvement of EARLY BUD-BREAK, an AP2/ERF transcription factor gene, in bud break in Japanese pear (*Pyrus pyrifolia* Nakai) lateral flower buds: expression, histone modifications and possible target genes. *Plant Cell Physiol.* 57, 1038–1047. doi: 10.1093/pcp/pcw041

Tuan, P. A., Bai, S., Saito, T., Ito, A., and Moriguchi, T. (2017). *Dormancy-associated MADS-Box* (*DAM*) and the abscisic acid pathway regulate pear endodormancy through a feedback mechanism. *Plant Cell Physiol.* 58, 1378–1390. doi: 10.1093/pcp/pcx074

Tylewicz, S., Petterle, A., Marttila, S., Miskolczi, P., Azeez, A., Singh, R. K., et al. (2018). Photoperiodic control of seasonal growth is mediated by ABA acting on cell-cell communication. *Science* 360, 212–215. doi: 10.1126/science.aan8576

Velappan, Y., Signorelli, S., and Considine, M. J. (2017). Cell cycle arrest in plants: what distinguishes quiescence, dormancy and differentiated G1? *Ann. Bot.* 120, 495–509. doi: 10.1093/aob/mcx082

Wang, S. Y., Jiao, H. J., and Faust, M. (1991). Changes in metabolic enzyme activities during thidiazuron-induced lateral budbreak of apple. *Hortscience* 26, 171–173. doi: 10.21273/hortsci.26.2.171

Wu, R., Tomes, S., Karunairetnam, S., Tustin, S. D., Hellens, R. P., Allan, A. C., et al. (2017). SVP-like MADS box genes control dormancy and budbreak in apple. *Front. Plant Sci.* 8:477. doi: 10.3389/fpls.2017.00477

Xiao, J., Jin, R., Yu, X., Shen, M., Wagner, J. D., Pai, A., et al. (2017). *Cis* and *trans* determinants of epigenetic silencing by Polycomb repressive complex 2 in *Arabidopsis*. *Nat. Genet.* 49, 1546–1552. doi: 10.1038/ng.3937

Yordanov, Y. S., Ma, C., Strauss, S. H., and Busov, V. B. (2014). EARLY BUD-BREAK 1 (EBB1) is a regulator of release from seasonal dormancy in poplar trees. *Proc. Natl. Acad. Sci. U.S.A.* 111, 10001–10006. doi: 10.1073/pnas.1405621111

Zhebentyayeva, T. N., Fan, S., Chandra, A., Bielenberg, D. G., Reighard, G. L., Okie, W. R., et al. (2014). Dissection of chilling requirement and bloom date QTLs in peach using a whole genome sequencing of sibling trees from an F2 mapping population. *Tree Genet. Genomes* 10, 35–51. doi: 10.1007/s11295-013-0660-6

Zhou, X., Jacobs, T. B., Xue, L. J., Harding, S. A., and Tsai, C. J. (2015). Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial *Populus* reveals 4-coumarate: CoA ligase specificity and redundancy. *New Phytol.* 208, 298–301. doi: 10.1111/nph.13470

Zhou, Y., Wang, Y., Krause, K., Yang, T., Dongus, J. A., Zhang, Y., et al. (2018). Telobox motifs recruit CLF/SWN-PRC2 for H3K27me3 deposition via TRB factors in *Arabidopsis*. *Nat. Genet.* 50, 638–644. doi: 10.1038/s41588-018-0109-9

Environmentally Sensitive Molecular Switches Drive Poplar Phenology

Jay P. Maurya[1†], Paolo M. Triozzi[2†], Rishikesh P. Bhalerao[1]* and Mariano Perales[2]*

[1] Umeå Plant Science Centre, Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, Umeå, Sweden, [2] Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid-Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Madrid, Spain

Boreal and temperate woody perennials are highly adapted to their local climate, which delimits the length of the growing period. Moreover, seasonal control of growth-dormancy cycles impacts tree productivity and geographical distribution. Therefore, traits related to phenology are of great interest to tree breeders and particularly relevant in the context of global warming. The recent application of transcriptional profiling and genetic association studies to poplar species has provided a robust molecular framework for investigating molecules with potential links to phenology. The environment dictates phenology by modulating the expression of endogenous molecular switches, the identities of which are currently under investigation. This review outlines the current knowledge of these molecular switches in poplar and covers several perspectives concerning the environmental control of growth-dormancy cycles. In the process, we highlight certain genetic pathways which are affected by short days, low temperatures and cold-induced signaling.

Keywords: poplar, adaptive response, cold response, circadian clock, short day, low ambient temperature, bud set, winter dormancy

INTRODUCTION

When the photoperiod falls below the critical day length, poplars undergo growth cessation, culminating in bud set and the acquisition of cold hardiness. The fact that components of light signaling, the circadian clock and orthologs of Arabidopsis flowering time regulators have all been implicated in this stage suggests the interplay between environmental signals and diurnal gene expression in dormancy switch. Thereafter, the sequential induction of ethylene and abscisic acid signaling pathways promotes bud maturation, cessation of meristematic activity and the establishment of dormancy. Once dormant, meristem becomes insensitive to growth-promoting signals. Release from dormancy strongly depends on the accumulation of a defined number of chilling hours, a mechanism for which the molecular basis is unknown. Under spring conditions, the low temperature (LT)-mediated activation of growth promoters reactivates meristem growth. Thus, poplar phenology is controlled by the orchestrated activity of molecular switches following exposure to environmental cues. This mini-review provides insight into the genetic network underlying poplar bud set, dormancy establishment and bud break with a focus on photoperiod- and temperature-dependent signaling. Moreover, we show evidence that bud set is highly sensitive to low ambient temperatures and identify candidate genes that may participate in this response.

SHORT DAYS PROMOTE ACCLIMATION TO WINTER CONDITIONS DURING BUD SET

Irrefutable experimental evidence for the short day (SD) requirement to cold acclimation comes from studies of photoperiod-insensitive oat *PHYA* overexpressing poplars Junttila and Kaurin, 1990; Olsen et al., 1997; Welling et al., 2002). When transgenic and wild type plants grown during SDs were subjected to freezing conditions, poplars overexpressing oat *PHYA* did not develop cold hardiness (Olsen et al., 1997; Welling et al., 2002). In contrast, neither transgenic nor wild type poplars grown during long days were able to survive freezing temperatures Junttila and Kaurin, 1990. This indicates that exposure to SDs acclimates poplars to cold weather conditions, and possibly other environmental stresses, which suggests that SDs may activate adaptation pathways.

The transcriptional and metabolomic profiling of poplar shoot apices and stem tissues grown under LD and SD conditions has been used to investigate the molecular signatures underlying SD-driven adaptive responses. The results revealed that SD conditions significantly alter the transcription of certain genes (Rohde et al., 2007; Ruttink et al., 2007; Karlberg et al., 2010; Hoffman et al., 2010; Resman et al., 2010). For example, a reduction in day length affects the transcription of several light signaling- and circadian clock-regulated genes (Ruttink et al., 2007). Moreover, SD also induces genes related to dehydration and cold adaptation, such as *LATE EMBRYOGENESIS-ABUNDANT* (*LEA*), *DEHYDRIN* (*DHN*), *COLD REGULATED GENES* (*COR*), and a set of transcription factors that mediate responses to abiotic stimuli, among others (Ruttink et al., 2007; Karlberg et al., 2010). Continued SD exposure stimulates ethylene signaling and abscisic acid (ABA) hormonal burst (Rohde et al., 2002; Ruttink et al., 2007; Karlberg et al., 2010). Ethylene and ABA are abiotic stress-related phytohormones required for adaptive responses in plants and are also functionally associated to the regulation of many aspects of bud development in poplar and birch (Rohde et al., 2002; Ruonala et al., 2006; Shi and Yang, 2014; Müller and Munné-Bosch, 2015; Tylewicz et al., 2015; Trivedi et al., 2016; Tylewicz et al., 2018). Thus, acclimation to winter conditions is mediated by the temporal orchestration of SD-induced transcriptional responses that partly originate from light signaling and circadian clock dependent pathways. As a result, poplars with modified circadian clock sensitivity showed different tolerance to freezing temperatures (Ibañez et al., 2010).

TEMPERATURE MODULATES BUD SET

The effect of temperature on bud set has been poorly investigated at the molecular level. Earlier physiological studies have shown that bud set is a thermosensitive process in trees (reviewed in Tanino et al., 2010). Furthermore, recent phenotyping studies using clonally replicated poplars grown under natural conditions show that the timing of bud set depends on the local climate (Rohde et al., 2011; Evans et al., 2014). Rohde et al. (2011) showed that sub-optimal growth temperatures delay the timing of bud set. Accordingly, we observed that the timing of bud set in hybrid aspen (*Populus tremula x P. alba*) grown under controlled conditions is highly sensitivity to small changes in ambient temperature (**Figures 1A,B**). Bud set was delayed by approximately 7 days in plants grown at 18°C relative to plants grown at 21°C, and this difference was amplified when the temperature was further reduced to 15°C (**Figures 1A,B**). This indicates that ambient temperature can impact the SD responses controlling bud set timing. Thus, it can be hypothesized that the activation or repression of thermosensitive genes underlies the regulation of pathways that control bud set. Numerous biological processes are governed by the endogenous clock (Greenham and McClung, 2015); thus, bud set in response to low ambient temperature may additionally be mediated by the modulation of clock-related pathways. For instance, the transcription of cold tolerance genes, e.g., *C-REPEAT BINDING FACTOR (CBF)* genes, follow circadian rhythms in Arabidopsis (Fowler et al., 2005; Dong et al., 2011).

To identify poplar candidate genes that could affect the timing of bud set under low ambient temperature, we investigated how diurnal photothermocycles affect bud set-associated single nucleotide polymorphism (SNP) genes (Mockler et al., 2007; Filichkin et al., 2011; Evans et al., 2014). Particularly, we analyzed the diurnal oscillation (cut-off 0.8) of bud set-associated SNP genes in two different conditions, photothermocycles (LDHC; Day 25°C/Night 12°C) and photocycles (LDHH; Day 25°C/Night 25°C), in *Populus thricocarpa* using the Diurnal database[1] (Mockler et al., 2007; Filichkin et al., 2011). A total of 134 genes associated with bud set showed robust diurnal transcription patterns over a 48 h period under LDHC conditions but not under constant temperature (LDHH), which suggests that the lower night temperature promotes their rhythmic expression (**Figures 1C–E**). Within the clusters of genes affected by photothermal cycles, we identified diurnal expression in the poplar ortholog of Arabidopsis *PHOSPHOLIPASE D DELTA* (*PLDDELTA*; Potri.007G060300), which is involved in phospholipid metabolism, freezing tolerance and stomatal closure (**Figure 1D**; Chen et al., 2008a,b; Distéfano et al., 2012; Uraji et al., 2012). Moreover, the poplar ortholog of *C-REPEAT-BINDING FACTOR 4* (*CBF4*; Potri.012G134100), which mediates the response to decreased temperatures in Arabidopsis (Wang and Hua, 2009), is also induced by cold temperatures in poplar and birch (**Figure 1E**; Benedict et al., 2006; Welling and Palva, 2008). These examples indicate that the rhythmic expression of genes associated with bud set – stimulated by diurnal oscillations in temperature – may be required for cold acclimation. In contrast, we identified 172 genes with bud set-associated SNPs that showed diurnal transcription patterns under constant temperatures (LDHH) but not under photothermocycles (LDHC). This suggests that reduced night-time temperatures can undermine the rhythmic expression of certain genes (**Figures 1C,F,G**). Within the clusters of

---

[1] http://diurnal.mocklerlab.org

FIGURE 1 | Bud set-associated genes in poplar are sensitive to low ambient temperature. **(A)** Bud set scores for hybrid poplar (*Populus tremula x P. alba*) plants grown under SD and three different temperatures (21, 18, and 15°C) for 9 weeks. **(B)** Comparative analysis of hybrid poplar bud stage progression in plants grown under SD and three different temperatures (21, 18, and 15°C) over 4 weeks. Asterisks denote significant differences between hybrid poplar plant groups (*n* = 8) grown at different temperatures (Kruskal–Wallis test followed by pairwise Wilcox test; *P < 0.05, **P < 0.01). **(C)** Venn diagram representing the intersection between genes with SNPs associated with bud set and that show robust diurnal oscillations (cut-off = 0.8) under light/dark cycles (LD) and constant temperature (25°C, HH) or photothermocycles (25°C during the day, 12°C during the night, HC). 172 genes (lilac) show robust diurnal rhythms only under LDHH, 134 genes (pink) show diurnal rhythms only under LDHC and 135 genes (fuchsia) oscillate under both conditions, respectively. **(D,E)** Clusters of genes that show diurnal expression within 0–11 h under LDHC **(D**, mean = pink) and within 12–23 h under LDHC **(E**, mean = pink). The line chart demonstrates how photothermocycles (LDHC) promote diurnal oscillation of poplar PDLDELTA **(D)** and CBF4 genes **(E)**. **(F,G)** Clusters of genes that show diurnal expression within 0–11 h under LDHH **(F**, mean = lilac) and within 12–23 h under LDHH **(G**, mean = lilac). The line chart demonstrates how photothermocycles (LDHC) disrupt the diurnal oscillation of poplar CCR1 **(F)** and PUR7 genes **(G)**.

genes with diurnal expression disrupted by photothermocycles, we highlight the poplar ortholog of Arabidopsis *COLD, CIRCADIAN RHYTHM, AND RNA BINDING 1* (*CCR1*; Potri.009G116400) gene (**Figure 1F**). CCR1 contains an RNA recognition motif (RRM), which promotes alternative splicing that is coupled to degradation by nonsense-mediated decay (NMD) (Schöning et al., 2008). *CCR1* is also regulated by both cold conditions and circadian rhythms in Arabidopsis (Carpenter et al., 1994). Furthermore, we found that the diurnal expression of the poplar ortholog of Arabidopsis *PURIN 7* (*PUR7*; Potri.017G051500), which is required to generate purine dependent cofactors in tissues under high rates of cell division, was disrupted by low night-time temperatures (Senecoff et al., 1996). These examples indicate that the impairment of key, diurnally regulated nucleic acid metabolism processes by decreased night-time temperatures could be important to timely bud set in poplar. Future

functional studies will hopefully elucidate the genetic network involved in bud set regulation when poplars are exposed to LTs.

COLD DISRUPTION OF CIRCADIAN CLOCK AND BUD SET

The circadian clock creates endogenous, 24 h rhythms to help plants and animals anticipate daily and seasonal environmental changes (Schultz and Kay, 2003). The circadian clock controls physiology, growth and development, as temporal transcriptional profiles have revealed that more than 30% of genes in Arabidopsis and poplar show circadian rhythms (Harmer et al., 2000; Covington et al., 2008; Michael et al., 2008; Hoffman et al., 2010). Recent research has revealed a role for the circadian clock in the genetic network regulating growth-dormancy cycles. It

FIGURE 2 | Photoperiodic and temperature control of bud set, dormancy establishment and bud break. Short days promote bud set via activation of *Flowering locus D1* (*FLD1*) and *Abscisic acid insensitive 3* (*ABI3*) (Tylewicz et al., 2015), the ethylene and ABA pathways (Rohde et al., 2002; Ruonala et al., 2006) as well as the circadian clock pathway (Hoffman et al., 2010; Ibañez et al., 2010). The timing of bud set is sensitive to low ambient temperature, which alters the expression of the circadian pathway (**Figure 1**). ABA plays an essential role in dormancy establishment by promoting plasmodesmata closure via PKL repression (Tylewicz et al., 2018) and activating SVL growth repressive gene (Singh et al., 2018). Increased exposure to chilling temperatures is necessary to the downregulation of ABA and degradation of growth repressors such as CENL1 (Karlberg et al., 2010; Mohamed et al., 2010; Tylewicz et al., 2018). Dormancy release increases gibberellin levels (Rinne et al., 2011), mediates plasmodesmata opening (Tylewicz et al., 2018) and activates bud break-promoting factors such as FT1 (Hsu et al., 2011; Rinne et al., 2011), EBB1 (Yordanov et al., 2014), and DML10 (Conde et al., 2017). Long days and warm temperatures restore shoot apical growth.

is widely accepted that circadian rhythms play a central role in the photoperiodic mechanism that controls poplar shoot apical growth (reviewed in Triozzi et al., 2018). However, the implication of the biological clock in the control of growth cessation and bud set has recently emerged (Ibañez et al., 2010; Kozarewa et al., 2010; Ding et al., 2018). The circadian rhythms of several pathways involved in growth cessation and bud set can persist 8–10 weeks after exposure to continuous SDs (Triozzi et al., personal communication). This is supported by the finding that the downregulation of poplar clock-related genes *LATE ELONGATED HYPOCOTYL* (*LHY*) and *TIMING OF CAB EXPRESSION 1* (*TOC1*) delays growth cessation and bud set (Ibañez et al., 2010). Furthermore, it has been firstly shown that chestnut clock-related genes display arrhythmic expression under winter conditions (Ramos et al., 2005; Ibañez et al., 2008). Moreover, chestnut and poplar clock-related genes respond to cold temperatures (4°C), showing high and constant expression irrespective of photoperiod (Ramos et al., 2005; Ibañez et al., 2010). It has been suggested that circadian clock disruption may facilitate bud set under cold temperatures (Johansson et al., 2015). Accordingly, cold-induced disruption of circadian rhythms caused wide transcriptional rearrangement of cold response genes in Arabidopsis (Bieniawska et al., 2008). Additionally, when exposed to freezing temperatures, *LHY*-RNAi poplars showed far more severe stem injuries than control plants. This indicates that the circadian clock is pivotal in the development of cold hardiness during bud set (Ibañez et al., 2010). Nevertheless,

the functional implications of cold-induced disruption of clock-regulated pathways during bud set needs further investigation.

DORMANCY ESTABLISHMENT

Prolonged SD exposure after growth cessation and bud set results in dormancy establishment in various plants (Heide, 1974; Espinosa-Ruiz et al., 2004; Ruttink et al., 2007; Maurya and Bhalerao, 2017). However, until recently, our knowledge of the molecular basis of bud dormancy was rudimentary. Conclusive evidences for dormancy establishment has come from studies on plant hormones ethylene and ABA, both of which are required for apical bud formation (**Figure 2**). For example, ethylene-insensitive dominant mutant *etr1-1* birch plants failed to develop apical buds under SD conditions (Ruonala et al., 2006). Ethylene has also been suggested to participate in dormancy induction in plants such as leafy spurge and *Chrysanthemum* (Sumitomo et al., 2008; Doğramacı et al., 2013. Nevertheless, genetic evidences for how ethylene affects bud dormancy are missing.

Abscisic acid levels increase following SD exposure (Ruttink et al., 2007; Karlberg et al., 2010). Interestingly, the apical buds of *etr1-1* mutant birches failed to accumulate ABA under SD conditions, which may suggest that defects in these plants may stem from the inability to increase ABA levels. Moreover, under SD conditions, *ABI3oe* plants display apical buds with

defects during bud maturation (Rohde et al., 2002; Ruttink et al., 2007). Although poplar *ABI3* has not been demonstrated to be ABA responsive or contribute to the ABA response *per se*, this finding may suggest a role for ABA, or at least ABI3, in bud maturation. Other research suggests that molecules associated with the opening and closing of plasmodesmata are involved in dormancy establishment (Rinne et al., 2011). It has been proposed that plasmodesmata closure in shoot apical meristem (SAM) is important for dormancy establishment (Rinne et al., 2011; Singh et al., 2017). Growth-promoting signals such a florigen move symplastically through plasmodesmata to SAM to promote growth under appropriate conditions, yet molecular and genetic evidence for this mechanism was only recently published (Singh et al., 2017; Tylewicz et al., 2018).

Abscisic acid-insensitive hybrid aspen plants overexpressing the dominant negative *abi1-1* allele failed to establish dormancy under short photoperiods (SPs) (Tylewicz et al., 2018). Both WT and *abi1-1* plants underwent growth cessation and apical bud formation when grown under SPs. However, when transferred from 11 weeks of SP to growth-promoting long photoperiods (LPs), WT plants did not start growing, which suggests that these plants were in a dormant state. However, *abi1-1* lines restarted growth 11–15 days after being transferred to LPs. The transcriptomic analysis of WT and *abi1-1* plants after 0, 6, and 10 weeks of SPs suggests that 1000s of genes are differentially regulated. Interestingly, many of these genes are associated with plasmodesmata closure and opening. Plasmodesmata closure-related genes such as *GERMIN-LIKE 10*, *REMORIN-LIKE 1* and *2* and *CALLOSE SYNTHASE 1* were upregulated in the apices of WT plants whereas genes related to plasmodesmata opening, such as *GH17_39,* were downregulated. The transcript levels of these genes differed between *abi1-1* and WT plants. Microscopic analyses of WT and *abi1-1* plants suggest that 83 and less than 3% of plasmodesmata in these plants, respectively, were closed after 10W of SPs (Tylewicz et al., 2018). Flowering Locus T (FT) protein acts as a mobile signal that can move from leaves to apex (Corbesier et al., 2007) and promotes growth of poplar trees even under SP (Böhlenius et al., 2006). Furthermore, when rootstocks of Flowering Locus T1 (FT1) plants were grafted to the scions of 10W SP grown WT and *abi1-1* plants, it was able to activate the growth of *abi1-1* plants even in SP conditions but not in WT scions.

Many of the transcripts associated with the plasmodesmata closure and opening also responded to the SPs (Tylewicz et al., 2018). Overexpression of plasmodesmata-located protein 1 (PDLP1) in *abi1-1* plants leads to dormancy establishment, as *abi1-1/PDLP1*oe plants transferred from SPs to LPs did not show growth reactivation. Certain genes involved in chromatin remodeling, such as *FERTILIZATION INDEPENDENT ENDOSPERM (FIE)* and *PICKLE (PKL),* have been shown to be upregulated after SD exposure (Ruttink et al., 2007). FIE is a component of the polycomb repressive complex 2 (PRC2), while PKL is its antagonist. PRC2 plays important roles in keeping the genes in the repressed state (Margueron and Reinberg, 2011). Transcript level of *PKL* was upregulated

in *abi1-1* plants, and its downregulation in *abi1-1* background developed dormancy by closing plasmodesmata (Tylewicz et al., 2018).

Taken together, these results suggest that SP influences ABA levels, which then differentially regulate plasmodesmata opening and closure-related genes to induce dormancy. Once the plasmodesmata are closed, growth-promoting signals cannot enter the SAM due to the formation of dormancy sphincters through callose deposition in plasmodesmata.

DORMANCY RELEASE AND BUD BURST

As mentioned earlier, dormancy establishment makes the SAM impermeable to growth-promoting signals, even under LD conditions. Release from dormancy requires that buds are exposed to LT for a prolonged period (Saure, 1985; Hannerz et al., 2003; Brunner et al., 2014; Fu et al., 2015). How temperature regulates dormancy release remains a poorly studied topic. Interestingly, the exposure of chilled buds to warm temperature is sufficient to induce dormancy release and subsequent bud burst. By monitoring the dormancy release time in plants with altered basal level of certain genes, we can anticipate their role in this process. Since plasmodesmata closure is required for dormancy establishment (Tylewicz et al., 2018), their opening could be expected to result in dormancy release. It has been suggested that LT treatment leads to the opening of plasmodesmata (Rinne et al., 2011), but this does not have proper experimental proof until now. However, some of the genes needed for the removal of plasmodesmatal dormancy sphincters by degrading the deposited callose, such as 1,3-β-glucanase (glucan hydrolase family 17 [GH17]), are upregulated after chilling treatment (Rinne et al., 2011).

Gibberellic acids (GAs) and FT are positive regulators of growth; as such, LT induces the transcription of *FT1* and genes implicated in GA metabolism, those that encode members of the GA3 and GA20 oxidases (Rinne et al., 2011). GAs may promote *GH17s* expression to reopen the plasmodesmata and thus restart the symplastic, growth-promoting cell-to-cell signaling within the SAM (**Figure 2**). However, genetic and experimental proof of this mechanism is still lacking. *CONSTANS (CO)* expression was shown to be upregulated almost threefold after 2 weeks of chilling. This suggests that the CO/FT module may be initiated by LT, yet *CO* expression remains at a similar level throughout the cold period and is only further enhanced when chilled buds sense LDs and warmer temperatures (Rinne et al., 2011). However, the expression of *CENL1/TFL1* (*CENTRORADIALIS-LIKE1/TERMINAL FLOWER 1*), a negative regulator of growth, remains at low levels in chilled buds throughout the LT period (**Figure 2**; Rinne et al., 2011). Experiments with *CENL1*oe and *CENL*-RNAi lines showed delayed and early bud burst, respectively, relative to wild type plants (Mohamed et al., 2010). This suggests that CENL1 represses dormancy release and bud burst. In chilled buds, CENL1 expression is only upregulated once the buds are transferred to warmer temperatures.

The process of dormancy release in trees is comparable to vernalization in Arabidopsis (Chouard, 1960). Many of the MAD-box transcription factors – termed DORMANCY ASSOCIATED MADS-BOX (DAM) – are known to be induced in the dormant buds of many plants, and are similar to short vegetative protein (SVP) and AGAMOUS-LIKE (AGL) transcription factors which control flowering in Arabidopsis (Ríos et al., 2014). Similar to *Flowering Locus C* (*FLC*) in Arabidopsis, these *DAM* genes are also epigenetically regulated by cold conditions. Cold-induced epigenetic silencing of floral repressor *FLC* is required during vernalization to induce flowering in Arabidopsis (Amasino, 2004; Bastow et al., 2004). The finding that expression of *DAM* genes goes down after cold treatment suggests that these genes are repressors of dormancy release (Shim et al., 2014; Howe et al., 2015). However, the exact roles of *DAM* genes can only be elucidated through future genetic studies.

Like dormancy release, our knowledge about bud burst is also very limited in perennial plants. A large number of genes are differentially regulated during dormancy release (Shim et al., 2014; Howe et al., 2015), but we do not have enough genetic evidence to conclusively describe their roles in dormancy release. Very recently, a genetic network mediating the dormancy release and bud break has been described (Singh et al., 2018). This study demonstrated that short vegetative protein-like (SVL), a tree ortholog of Arabidopsis SVP, acts as a central regulator of dormancy release and bud burst in poplar. It negatively regulates dormancy release and bud burst by promoting and repressing the expressions of negative and positive regulators of growth, respectively, after cold treatment (Singh et al., 2018). The expression of *Early Bud-Break 1* (*EBB1*), a putative APETALA2/Ethylene responsive factor, also increases after cold periods during dormancy release and before bud burst. However, *EBB1* transcripts are undetectable during the dormancy period. Transgenic lines with *EBB1* overexpression and downregulation show early and delayed bud burst relative to wild type plants, respectively. This suggests that EBB1 is a positive regulator of bud burst (**Figure 2**; Yordanov et al., 2014).

CONCLUDING REMARKS

The effect of temperature on tree phenology is an important topic in the context of climate change. Extended seasonal growth and shifts in latitudinal distribution demonstrate how plants are already adapting to increasing global temperatures. Although previous research has attempted to elucidate how temperature influences growth cessation and bud set in trees, very little is still known about the molecular mechanisms involved in these phenomena. For this reason, understanding how environmentally sensitive molecular switches regulate the perception and transduction of temperature signals in woody plants will be crucial to predicting how plants will adapt to warmer environments and designing appropriate breeding programs for this scenario.

AUTHOR CONTRIBUTIONS

PT and MP analyzed the data. JM, PT, RB, and MP contributed to the writing of the manuscript.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Amasino, R. (2004). Vernalization, competence, and the epigenetic memory of winter. *Plant Cell* 16, 2553–2559. doi: 10.1105/tpc.104.161070

Bastow, R., Mylne, J. S., Lister, C., Lippman, Z., Martienssen, R. A., and Dean, C. (2004). Vernalization requires epigenetic silencing of FLC by histone methylation. *Nature* 427, 164–167. doi: 10.1038/nature02269

Benedict, C., Skinner, J. S., Meng, R., Chang, Y., Bhalerao, R., Huner, N. P. A., et al. (2006). The CBF1-dependent low temperature signalling pathway, regulon and increase in freeze tolerance are conserved in Populus spp. *Plant Cell Environ.* 29, 1259–1272. doi: 10.1111/j.1365-3040.2006.01505.x

Bieniawska, Z., Espinoza, C., Schlereth, A., Sulpice, R., Hincha, D. K., and Hannah, M. A. (2008). Disruption of the Arabidopsis circadian clock is responsible for extensive variation in the cold-responsive transcriptome. *Plant Physiol.* 147, 263–279. doi: 10.1104/pp.108.118059

Böhlenius, H., Huang, T., Charbonnel-Campaa, L., Brunner, A. M., Jansson, S., Strauss, S. H., et al. (2006). CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. *Science* 312, 1040–1043. doi: 10.1126/science.1126038

Brunner, A. M., Evans, L. M., Hsu, C. Y., and Sheng, X. (2014). Vernalization and the chilling requirement to exit bud dormancy: shared or separate regulation? *Front. Plant Sci.* 5:732. doi: 10.3389/fpls.2014.00732

Carpenter, C. D., Kreps, J. A., and Simon, A. E. (1994). Genes encoding glycine-rich *Arabidopsis thaliana* proteins with RNA-binding motifs are influenced by cold treatment and an endogenous circadian rhythm. *Plant Physiol.* 104, 1015–1025. doi: 10.1104/pp.104.3.1015

Chen, Q. F., Xiao, S., and Chye, M. L. (2008a). Arabidopsis ACBP6 is an acyl-CoA-binding protein associated with phospholipid metabolism. *Plant Signal. Behav.* 3, 1019–1020. doi: 10.4161/psb.6762

Chen, Q. F., Xiao, S., and Chye, M. L. (2008b). Overexpression of the Arabidopsis 10-kilodalton Acyl-coenzyme a-binding protein acbp6 enhances freezing tolerance. *Plant Physiol.* 148, 304–315. doi: 10.1104/pp.108.123331

Chouard, P. (1960). Vernalization and its relations to dormancy. *Annu. Rev. Plant Physiol.* 11, 191–238. doi: 10.1146/annurev.pp.11.060160.001203

Conde, D., Le Gac, A. L., Perales, M., Dervinis, C., Kirst, M., Maury, S., et al. (2017). Chilling-responsive demeter-like DNA demethylase mediates in poplar bud break. *Plant Cell Environ.* 40, 2236–2249. doi: 10.1111/pce.13019

Corbesier, L., Vincent, C., Jang, S., Fornara, F., Fan, Q., Searle, I., et al. (2007). FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. *Science* 316, 1030–1033. doi: 10.1126/science.1141752

Covington, M. F., Maloof, J. N., Straume, M., Kay, S. A., and Harmer, S. L. (2008). Global transcriptome analysis reveals circadian regulation of key pathways in plant growth and development. *Genome Biol.* 9:R130. doi: 10.1186/gb-2008-9-8-130

Ding, J., Böhlenius, H., Rühl, M. G., Chen, P., Sane, S., Zambrano, J. A., et al. (2018). Gigantea-like genes control seasonal growth cessation in populus. *New Phytol.* 218, 1491–1503. doi: 10.1111/nph.15087

Distéfano, A. M., Scuffi, D., García-Mata, C., Lamattina, L., and Laxalt, A. M. (2012). Phospholipase Dδ is involved in nitric oxide-induced stomatal closure. *Planta* 236, 1899–1907. doi: 10.1007/s00425-012-1745-4

Doğramacı, M., Foley, M. E., Chao, W. S., Christoffers, M. J., and Anderson, J. V. (2013). Induction of endodormancy in crown buds of leafy spurge (*Euphorbia esula* L.) implicates a role for ethylene and cross-talk between photoperiod and temperature. *Plant Mol. Biol.* 81, 577–593. doi: 10.1007/s11103-013-0026-3

Dong, M. A., Farre, E. M., and Thomashow, M. F. (2011). Circadian clock-associated 1 and late elongated hypocotyl regulate expression of the c-repeat binding factor (cbf) pathway in Arabidopsis. *Proc. Natl. Acad. Sci. U.S.A.* 108, 7241–7246. doi: 10.1073/pnas.1103741108

Espinosa-Ruiz, A., Saxena, S., Schmidt, J., Mellerowicz, E., Miskolczi, P., Bako, L., et al. (2004). Differential stage-specific regulation of cyclin-dependent kinases during cambial dormancy in hybrid aspen. *Plant J.* 38, 603–615. doi: 10.1111/j.1365-313X.2004.02070.x

Evans, L. M., Slavov, G. T., Rodgers-Melnick, E., Martin, J., Ranjan, P., Muchero, W., et al. (2014). Population genomics of Populus trichocarpa identifies signatures of selection and adaptive trait associations. *Nat. Genet.* 46, 1089–1096. doi: 10.1038/ng.3075

Filichkin, S. A., Breton, G., Priest, H. D., Dharmawardhana, P., Jaiswal, P., Fox, S. E., et al. (2011). Global profiling of rice and poplar transcriptomes highlights key conserved Circadian-controlled pathways and cis-regulatory modules. *PLoS One* 6:e16907. doi: 10.1371/journal.pone.0016907

Fowler, S. G., Cook, D., and Thomashow, M. F. (2005). Low temperature induction of Arabidopsis CBF1, 2, and 3 is gated by the circadian clock. *Plant Physiol.* 137, 961–968. doi: 10.1104/pp.104.058354

Fu, Y. H., Piao, S., Vitasse, Y., Zhao, H., Boeck, H. J. D., Liu, Q., et al. (2015). Increased heat requirement for leaf flushing in temperate woody species over 1980–2012: effects of chilling, precipitation and insolation. *Global Change Biol.* 21, 2687–2697. doi: 10.1111/gcb.12863

Greenham, K., and McClung, C. R. (2015). Integrating circadian dynamics with physiological processes in plants. *Nat. Rev. Genet.* 16, 598–610. doi: 10.1038/nrg3976

Hannerz, M., Ekberg, I., and Norell, L. (2003). Variation in chilling requirements for completing bud rest between provenances of Norway spruce. *Silvae Genet.* 52, 161–168.

Harmer, S. L., Hogenesch, J. B., Straume, M., Chang, H. S., Han, B., Zhu, T., et al. (2000). Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. *Science* 290, 2110–2113. doi: 10.1126/science.290.5499.2110

Heide, O. M. (1974). Growth and dormancy in Norway spruce ecotypes (*Picea abies*) I. Interaction of photoperiod and temperature. *Physiol. Plant.* 30, 1–12. doi: 10.1111/j.1399-3054.1974.tb04983.x

Hoffman, D. E., Jonsson, P., Bylesjö, M., Trygg, J., Antti, H., Eriksson, M. E., et al. (2010). Changes in diurnal patterns within the Populus transcriptome and metabolome in response to photoperiod variation. *Plant Cell Environ.* 33, 1298–1313. doi: 10.1111/j.1365-3040.2010.02148.x

Howe, G. T., Horvath, D. P., Dharmawardhana, P., Priest, H. D., Mockler, T. C., and Strauss, S. H. (2015). Extensive transcriptome changes during natural onset and release of vegetative bud dormancy in Populus. *Front. Plant Sci.* 6:989. doi: 10.3389/fpls.2015.00989

Hsu, C.-Y., Adams, J. P., Kim, H., No, K., Ma, C., Strauss, S. H., et al. (2011). Flowering locus t duplication coordinates reproductive and vegetative growth in perennial poplar. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10756–10761. doi: 10.1073/pnas.1104713108

Ibañez, C., Kozarewa, I., Johansson, M., Ogren, E., Rohde, A., and Eriksson, M. E. (2010). Circadian clock components regulate entry and affect exit of seasonal dormancy as well as winter hardiness in Populus trees. *Plant Physiol.* 153, 1823–1833. doi: 10.1104/pp.110.158220

Ibañez, C., Ramos, A., Acebo, P., Contreras, A., Casado, R., Allona, I., et al. (2008). Overall alteration of circadian clock gene expression in the chestnut cold response. *PLoS One* 3:e3567. doi: 10.1371/journal.pone.0003567

Johansson, M., Ramos-Saìnchez, J. M., Conde, D., Ibaìnþez, C., Takata, N., and Allona, I. (2015). "Role of the circadian clock in cold acclimation and winter dormancy in perennial plants," in *Advances in Plant Dormancy*, ed. J. V. Anderson (Basel: Springer International Publishing), 51–74.

Junttila, O., and Kaurin, Å (1990). Environmental control of cold acclimation in salix pentandra. *Scand. J. For. Res.* 5, 195–204. doi: 10.1080/02827589009382605

Karlberg, A., Englund, M., Petterle, A., Molnar, G., Sjödin, A., and Bako, L. (2010). Analysis of global changes in gene expression during activity-dormancy cycle in hybrid aspen apex. *Plant Biotechnol.* 27, 1–16. doi: 10.5511/plantbiotechnology.27.1

Kozarewa, I., Ibanez, C., and Johansson, M. (2010). Alteration of PHYA expression change circadian rhythms and timing of bud set in Populus. *Plant Mol. Biol.* 73, 143–156. doi: 10.1007/s11103-010-9619-2

Margueron, R., and Reinberg, D. (2011). The polycomb complex PRC2 and its mark in life. *Nature* 469, 343–349. doi: 10.1038/nature09784

Maurya, J. P., and Bhalerao, R. P. (2017). Photoperiod- and temperature-mediated control of growth cessation and dormancy in trees: a molecular perspective. *Ann. Bot.* 120, 351–360. doi: 10.1093/aob/mcx061

Michael, T. P., Mockler, T. C., Breton, G., McEntee, C., Byer, A., and Trout, J. D. (2008). Network discovery pipeline elucidates conserved time-of-day-specific cis-regulatory modules. *PLoS Genet.* 4:e14. doi: 10.1371/journal.pgen.0040014

Mockler, T. C., Michael, T. P., Priest, H. D., Shen, R., Sullivan, C. M., and Givan, S. A. (2007). The diurnal project: diurnal and circadian expression profiling, model-based pattern matching, and promoter analysis. *Cold Spring Harb. Symp. Quant. Biol.* 72, 353–363. doi: 10.1101/sqb.2007.72.006

Mohamed, R., Wang, C. T., Ma, C., Shevchenko, O., Dye, S. J., Puzey, J. R., et al. (2010). Populus CEN/TFL1 regulates first onset of flowering, axillary meristem identity and dormancy release in Populus. *Plant J.* 62, 674–688. doi: 10.1111/j.1365-313X.2010.04185.x

Müller, M., and Munné-Bosch, S. (2015). Ethylene response factors: a key regulatory hub in hormone and stress signaling. *Plant Physiol.* 169, 32–41. doi: 10.1104/pp.15.00677

Olsen, J. E., Junttila, O., Nilsen, J., Eriksson, M. E., Martinussen, I., Olsson, O., et al. (1997). Ectopic expression of oat phytochrome A in hybrid aspen changes critical daylength for growth and prevents cold acclimatization. *Plant J.* 12, 1339–1350. doi: 10.1046/j.1365-313x.1997.12061339.x

Ramos, A., Pérez-Solís, E., Ibáñez, C., Casado, R., Collada, C., Gómez, L., et al. (2005). Winter disruption of the circadian clock in chestnut. *Proc. Natl. Acad. Sci. U.S.A.* 102, 7037–7042. doi: 10.1073/pnas.0408549102

Resman, L., Howe, G., Jonsen, D., Englund, M., Druart, N., Schrader, J., et al. (2010). Components acting downstream of short day perception regulate differential cessation of cambial activity and associated responses in early and late clones of hybrid poplar. *Plant Physiol.* 154, 1294–1303. doi: 10.1104/pp.110.163907

Rinne, P. L., Welling, A., Vahala, J., Ripel, L., Ruonala, R., Kangasjarvi, J., et al. (2011). Chilling of dormant buds hyperinduces flowering locus t and recruits GA-inducible 1,3-betaglucanases to reopen signal conduits and release dormancy in Populus. *Plant Cell* 23, 130–146. doi: 10.1105/tpc.110.081307

Ríos, G., Leida, C., Conejero, A., and Badenes, M. L. (2014). Epigenetic regulation of bud dormancy events in perennial plants. *Front. Plant Sci.* 5:247. doi: 10.3389/fpls.2014.00247

Rohde, A., Bastien, C., and Boerjan, W. (2011). Temperature signals contribute to the timing of photoperiodic growth cessation and bud set in poplar. *Tree Physiol.* 31, 472–482. doi: 10.1093/treephys/tpr038

Rohde, A., Prinsen, E., De Rycke, R., Engler, G., Van Montagu, M., and Boerjan, W. (2002). PtABI3 impinges on the growth and differentiation of embryonic leaves during bud set in poplar. *Plant Cell* 14, 1885–1901. doi: 10.1105/tpc.003186

Rohde, A., Ruttink, T., Hostyn, V., Sterck, L., Van Driessche, K., and Boerjan, W. (2007). Gene expression during the induction, maintenance, and release of dormancy in apical buds of poplar. *J. Exp. Bot.* 58, 4047–4060. doi: 10.1093/jxb/erm261

Ruonala, R., Rinne, P. L. H., Baghour, M., Moritz, T., Tuominen, H., and Kangasjärvi, J. (2006). Transitions in the functioning of the shoot apical meristem in birch (*Betula pendula*) involve ethylene. *Plant J.* 46, 628–640. doi: 10.1111/j.1365-313X.2006.02722.x

Ruttink, T., Arend, M., Morreel, K., Storme, V., Rombauts, S., Fromm, J., et al. (2007). A molecular timetable for apical bud formation and dormancy induction in poplar. *Plant Cell* 19, 2370–2390. doi: 10.1105/tpc.107.052811

Saure, M. (1985). Dormancy release in deciduous fruit trees. *Hortic. Rev.* 7, 239–300. doi: 10.1002/9781118060735.ch6

Schöning, J. C., Streitner, C., Meyer, I. M., Gao, Y., and Staiger, D. (2008). Reciprocal regulation of glycine-rich RNA-binding proteins via an interlocked feedback loop coupling alternative splicing to nonsense-mediated decay in Arabidopsis. *Nucleic Acids Res.* 36, 6977–6987. doi: 10.1093/nar/gkn847

Schultz, T. F., and Kay, S. A. (2003). Circadian clocks in daily and seasonal control of development. *Science* 301, 326–328. doi: 10.1126/science.1085935

Senecoff, J. F., McKinney, E. C., and Meagher, R. B. (1996). De novo purine synthesis in *Arabidopsis thaliana*. II. The PUR7 gene encoding 5'-phosphoribosyl-4-(N-succinocarboxamide)-5-aminoimidazole synthetase is expressed in rapidly dividing tissues. *Plant Physiol.* 112, 905–917. doi: 10.1104/pp.112.3.905

Shi, Y., and Yang, S. (2014). "ABA regulation of the cold stress response in plants," in *Abscisic Acid: Metabolism, Transport and Signaling*, ed. D. P. Zhang (Dordrecht: Springer), 337–363. doi: 10.1007/978-94-017-9424-4_17

Shim, D., Ko, J. H., Kim, W. C., Wang, Q., Keathley, D. E., and Han, K. H. (2014). A molecular framework for seasonal growth–dormancy regulation in perennial plants. *Hortic. Res.* 1:14059. doi: 10.1038/hortres.2014.59

Singh, R. K., Maurya, J. P., Azeez, A., Miskolczi, P., Tylewicz, S., Stojkoviè, K., et al. (2018). A genetic network mediating the control of bud break in hybrid aspen. *Nat. Commun.* 9:4173. doi: 10.1038/s41467-018-06696-y

Singh, R. K., Svystun, T., AlDahmash, B., Jönsson, A. M., and Bhalerao, R. P. (2017). Photoperiod- and temperature-mediated control of phenology in trees – a molecular perspective. *New Phytol.* 213, 511–524. doi: 10.1111/nph.14346

Sumitomo, K., Narumi, T., Satoh, S., and Hisamatsu, T. (2008). Involvement of the ethylene response pathway in dormancy induction in chrysanthemum. *J. Exp. Bot.* 59, 4075–4082. doi: 10.1093/jxb/ern247

Tanino, K. K., Kalcsits, L., Silim, S., Kendall, E., and Gray, G. R. (2010). Temperature-driven plasticity in growth cessation and dormancy development in deciduous woody plants: a working hypothesis suggesting how molecular and cellular function is affected by temperature during dormancy induction. *Plant Mol. Biol.* 73, 49–65. doi: 10.1007/s11103-010-9610-y

Triozzi, P. M., Ramos-Sánchez, J. M., Hernández-Verdeja, T., Moreno-Cortés, A., Allona, I., and Perales, M. (2018). Photoperiodic control of shoot apical growth in poplar. *Front. Plant Sci.* 9:1030. doi: 10.3389/fpls.2018.01030

Trivedi, D. K., Gill, S. S., and Tuteja, N. (2016). "Abscisic acid (ABA): biosynthesis, regulation, and role in abiotic stress tolerance," in *Abiotic Stress Response in Plants*, eds N. Tuteja and S. S. Gill (Hoboken, NJ: Wiley), doi: 10.1002/9783527694570.ch15

Tylewicz, S., Petterle, A., Marttila, S., Miskolczi, P., Azeez, A., Singh, R. K., et al. (2018). Photoperiodic control of seasonal growth is mediated by ABA acting on cell-cell communication. *Science* 360, 212–215. doi: 10.1126/science.aan8576

Tylewicz, S., Tsuji, H., Miskolczi, P., Petterle, A., Azeez, A., Jonsson, K., et al. (2015). Dual role of tree florigen activation complex component FD in photoperiodic growth control and adaptive response pathways. *Proc. Natl. Acad. Sci. U.S.A.* 112, 3140–3145. doi: 10.1073/pnas.1423440112

Uraji, M., Katagiri, T., Okuma, E., Ye, W., Hossain, M. A., Masuda, C., et al. (2012). Cooperative function of pld and pld 1 in abscisic acid-induced stomatal closure in Arabidopsis. *Plant Physiol.* 159, 450–460. doi: 10.1104/pp.112.195578

Wang, Y., and Hua, J. (2009). A moderate decrease in temperature induces COR15a expression through the CBF signaling cascade and enhances freezing tolerance. *Plant J.* 60, 340–349. doi: 10.1111/j.1365-313X.2009.03959.x

Welling, A., Moritz, T., Palva, E. T., and Junttila, O. (2002). Independent activation of cold acclimation by low temperature and short photoperiod in hybrid aspen. *Plant Physiol.* 129, 1633–1641. doi: 10.1104/pp.003814

Welling, A., and Palva, E. T. (2008). Involvement of CBF transcription factors in winter hardiness in birch. *Plant Physiol.* 147, 1199–1211. doi: 10.1104/pp.108.117812

Yordanov, Y. S., Ma, C., Strauss, S. H., and Busov, V. B. (2014). Early bud-break 1 (EBB1) is a regulator of release from seasonal dormancy in poplar trees. *Proc. Natl. Acad. Sci. U.S.A.* 111, 10001–10006. doi: 10.1073/pnas.1405621111

Check for updates

Hardwood Tree Genomics: Unlocking Woody Plant Biology

Gerald A. Tuskan[1]*, Andrew T. Groover[2], Jeremy Schmutz[3,4], Stephen Paul DiFazio[5], Alexander Myburg[6], Dario Grattapaglia[7,8], Lawrence B. Smart[9], Tongming Yin[10], Jean-Marc Aury[11], Antoine Kremer[12], Thibault Leroy[12,13], Gregoire Le Provost[12], Christophe Plomion[12], John E. Carlson[14], Jennifer Randall[15], Jared Westbrook[16], Jane Grimwood[3], Wellington Muchero[1], Daniel Jacobson[1] and Joshua K. Michener[1]

[1] Center for Bioenergy Innovation, Biosciences Division, Oak Ridge National Laboratory (DOE), Oak Ridge, TN, United States, [2] Pacific Southwest Research Station, USDA Forest Service, Davis, CA, United States, [3] HudsonAlpha Institute for Biotechnology, Huntsville, AL, United States, [4] Joint Genome Institute, Walnut Creek, CA, United States, [5] Department of Biology, West Virginia University, Morgantown, WV, United States, [6] Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute, University of Pretoria, Pretoria, South Africa, [7] Embrapa Recursos Genéticos e Biotecnologia, Brasília, Brazil, [8] Universidade Católica de Brasília, Brasília, Brazil, [9] Horticulture Section, School of Integrative Plant Science, Cornell University, Geneva, NY, United States, [10] The Key Laboratory for Poplar Improvement of Jiangsu Province, Nanjing Forestry University, Nanjing, China, [11] Commissariat à l'Energie Atomique, Genoscope, Institut de Biologie François-Jacob, Evry, France, [12] BIOGECO, INRA, Université de Bordeaux, Cestas, France, [13] ISEM, CNRS, IRD, EPHE, Université de Montpellier, Montpellier, France, [14] Schatz Center for Tree Molecular Genetics, Department of Ecosystem Science and Management, Pennsylvania State University, University Park, PA, United States, [15] Department of Entomology, Plant Pathology and Weed Science, New Mexico State University, Las Cruces, NM, United States, [16] The American Chestnut Foundation, Asheville, NC, United States

Woody perennial angiosperms (i.e., hardwood trees) are polyphyletic in origin and occur in most angiosperm orders. Despite their independent origins, hardwoods have shared physiological, anatomical, and life history traits distinct from their herbaceous relatives. New high-throughput DNA sequencing platforms have provided access to numerous woody plant genomes beyond the early reference genomes of *Populus* and *Eucalyptus*, references that now include willow and oak, with pecan and chestnut soon to follow. Genomic studies within these diverse and undomesticated species have successfully linked genes to ecological, physiological, and developmental traits directly. Moreover, comparative genomic approaches are providing insights into speciation events while large-scale DNA resequencing of native collections is identifying population-level genetic diversity responsible for variation in key woody plant biology across and within species. Current research is focused on developing genomic prediction models for breeding, defining speciation and local adaptation, detecting and characterizing somatic mutations, revealing the mechanisms of gender determination and flowering, and application of systems biology approaches to model complex regulatory networks underlying quantitative traits. Emerging technologies such as single-molecule, long-read sequencing is being employed as additional woody plant species, and genotypes within species, are sequenced, thus enabling a comparative ("evo-devo") approach to understanding the unique biology of large woody plants. Resource availability, current genomic and genetic applications, new discoveries and predicted future developments are illustrated and discussed for poplar, eucalyptus, willow, oak, chestnut, and pecan.

Keywords: tree habit, somatic mutations, evolutionary ecology, quantitative genetics, adaptive traits, comparative genomics

INTRODUCTION

Forest trees have historically been classified as softwood or hardwoods for practical purposes. Botanically, softwood trees are gymnosperms and hardwood trees are angiosperms (i.e., flowering plants) and, as the names indicate, there are differences in reproductive biology and overall habit between these two groups. Around the globe, hardwood trees underpin complex ecosystems, sequester carbon (e.g., rainforests), provide raw materials for the forest products industry, are a source of primary energy in developing countries and are increasingly utilized as a source of renewable bioenergy, biomaterials, and other bioproducts (Ragauskas et al., 2014; FAO, 2016).

Woody perennial angiosperms, i.e., hardwood trees, are amazingly diverse in terms of morphology, life habits, physiology, among many other traits. Hardwood trees show extensive variation in wood anatomy, leaf morphology, whole-tree architecture, secondary metabolism and numerous adaptive traits (Groover and Crook, 2017). The classification of "tree" or arboreal form reflects several distinctive features: e.g., extensive wood formation produced by a vascular or secondary cambium, persistent interannual growth, single or multiple stems exceeding two meters in height, accompanied by perennial lateral branches. Individually, these defining features can however occur outside of the classification as a tree, e.g., some annuals that can produce woody tissue or achieve heights greater than 2 m. Exception aside, hardwood trees can establish extensive root systems that enable vigorous interannual growth, store energy and nutrients overwinter in stem and root tissues, and in many scenarios, create dominant ecosystem overstories.

The ancestral habit for angiosperms is believed to be reflected in the most basal extant angiosperm, *Amborella trichopoda* (Soltis et al., 2008; Albert et al., 2013), which is a small tree native to New Caledonia. During angiosperm evolution, woody perennial habit has been accentuated in some lineages, lost in others (e.g., monocots), and lost and then regained in still others (Spicer and Groover, 2010). To a large extent, these evolutionary and developmental changes are contained in the genomes of the extant hardwood tree species.

Reflecting their evolutionary history, hardwood trees genomes display divergent chromosomal architectures, with multiple ancestral whole-genome duplications and rearrangements, gene family expansions, and many segmental deletions across lineages (Tuskan et al., 2006; Soltis et al., 2008; Dai et al., 2014; Myburg et al., 2014; Plomion et al., 2018). Understanding comparative connections among genomes and the associated characteristics defining hardwood trees remains a substantial research challenge. In this review we will cover the current state of the science in five major groups of hardwood trees outlining the current assembly and annotation metrics, as well as presenting on-going application of these resources. We will highlight on-line resources for these genomes and provide comparative resources for other angiosperm species. We will provide projections of near-term technologies and analytics that will increase the availability and research value of genomes of hardwood trees and conclude with a summary of several unanswered questions related to macro- and micro-evolution of woody perennials that can be addressed using genomic approaches in future.

STATE OF THE SCIENCE

Populus

When the *Populus trichocarpa* genome, 'Nisquallly-1,' was released in 2006 it was the first plant genome to use shotgun sequencing assembly approaches, the first genome to create chromosome-level assemblies based on genetic mapping, the first woody perennial genome to be assembled and annotated, as well as the first metagenomic assembly of associated endophytic bacteria and fungi (Tuskan et al., 2006). Today, v3.2 contains 423 Mb [of the ~485 Mb genome] in 1,446 scaffolds with 2,585 gaps[1]. The scaffold N50 metric is 8 Mb and contig N50 is 205 kb, with roughly 98% of the genome in 181 scaffolds. Trained on deep RNAseq data from the Gene Atlas project[2], there are 41,335 predicted loci and 73,013 protein coding transcripts. Methylation maps are available for 10 tissue types, including bud tissue, callus tissue, female catkins, internode explants, male catkins, phloem, xylem, and roots (Vining et al., 2012). Over 28 million single nucleotide polymorphisms (SNPs) from resequenced data representing over 880 native *P. trichocarpa* genotypes covering the core distribution of the species range are publicly available (Geraldes et al., 2013[3]). This SNP resource has been used to characterize geographic structure and linkage disequilibrium (Slavov et al., 2012), detect signatures of selection across the genome (Evans et al., 2014) and identify genetic loci associated with various phenotypes based on genome-wide association approaches (Porth et al., 2013; Evans et al., 2014; McKown et al., 2014; Muchero et al., 2015; Fahrenkrog et al., 2017; Liu et al., 2018). The resequenced data is currently being assembled into a *Populus* pan genome, with preliminary data suggesting that there may be as many as 20,000 additional gene models that are not included in the current *P. trichocarpa* reference genome annotation (Pinosio et al., 2016). The first draft assemblies of *Populus deltoides*, 'WV94' and *Populus tremula* × *alba* hybrid, '717-1B4,' have been sequenced, assembled and annotated[1]. The assembly for WV94 is approximately 446 Mb in 1,375 scaffolds, with scaffold N50 of 21.7 Mb and contig N50 of 590 kb and there are 44,853 protein coding loci. CRISPR-related PAM (protospacer adjacent motif) sites for *P. tremula* × *alba* 717 have been published are on line at http://aspendb.uga.edu/s717. Highly contiguous, *de novo* genome assemblies have also been produced for *P. euphratica* (Ma et al., 2014) and *P. tremula*[4], which greatly widen the phylogenetic sampling of the genus. Ongoing resource development work in *Populus* includes efforts to expand the tissue type and experimental conditions in the Gene Atlas, enlarge the number resequenced genotypes from the southern and eastern portions of the range, and release of V4.0 of the *P. trichocarpa* genome based on long-read

---

sequencing and dense genetic maps derived from resequencing 1000 progeny from 49 half-sib families. Dozens of genome-wide association studies continue across a broad array of phenotypes and efforts to estimate the somatic mutation rate found in old-growth *P. trichocarpa* are currently underway. Methods to integrate SNP, gene expression, transcription factor binding, gene dosage, methylation, metabolite expression, phenome, and co-evolutionary relationships have been developed to generate a systems biology view of the molecular and regulatory interactions that lead to organismal scale traits (Henry et al., 2015; Liu et al., 2015; Joubert et al., 2017, 2018; Weighill et al., 2018). New algorithms are being developed to use signal processing techniques as well as explainable artificial intelligence to build better systems biology models, which include methods to discover and model genome-wide epistasis and pleiotropy. Efforts are also underway to gain a comprehensive view of chromatin structure and transcription factor binding sites (Rao et al., 2014).

Eucalyptus

Eucalyptus species and hybrids represent the most widely cultivated hardwood biomass crop globally and *Eucalyptus grandis* (rose gum) clonal genotype 'BRASUZ1,' was selected as the reference for the genus. Published in 2014, the *E. grandis* genome was the first for the rosid order Myrtales and for the family Myrtaceae. Key features of the genome included a very high proportion (34%) and number of genes (over 12,500) in tandem duplicate arrays, an ancient (109 Mya) genome-wide duplication event and high diversity of gene families encoding plant specialized metabolites such as phenylpropanoids and terpenes important for plant defense and pharmaceutical uses such as eucalyptus oils (Mewalal et al., 2017). BRASUZ1 was one of the last plant genomes produced exclusively with Sanger technology combined with extensive BAC-end coverage, resulting in a high-quality assembly (currently V2.0[1]) comprising 691 Mb in 4,943 scaffolds with a scaffold N50 of 57.5 Mb, contig N50 of 67 kb and 94% of the estimated 640 Mb genome (Grattapaglia and Bradshaw, 1994) in 11 pseudomolecules containing 288 scaffolds longer than 50 kb. A total of 36,349 protein coding loci are annotated and the assembly has served as reference for extensive transcriptomics studies targeting the regulation and biosynthesis of lignocellulosic biomass (Mizrachi et al., 2014; Carocha et al., 2015; Hussey et al., 2015; Soler et al., 2015) and in comparative analyses with *Arabidopsis* (Davin et al., 2016) and *Populus* (Hefer et al., 2015; Pinard et al., 2015). Other research targets have included terpene biosynthesis (Myburg et al., 2014; Kulheim et al., 2015), reproductive biology (Vining et al., 2015), plant defense (Christie et al., 2015; Mangwanda et al., 2015; Oates et al., 2015; Meyer et al., 2016) and abiotic interactions (Plett et al., 2015; Spokevicius et al., 2017). High-throughput transcriptome sequencing in 100s of segregating interspecific hybrids of *E. grandis* × *E. urophylla* have been used to perform the first comprehensive systems genetics analysis of wood development in *Eucalyptus* (Mizrachi et al., 2017). The *E. grandis* genome has also served as a reference for whole-genome analysis of the causes of inbreeding depression (Hedrick et al., 2016) and development of a commercial,

multi-species 60K SNP genotyping chip tagging 96% of the genome with 1 SNP every 12–20 kb (Silva-Junior et al., 2015). The EuCHIP60K has been a key resource to investigate genome-wide recombination, linkage disequilibrium, and nucleotide diversity (Silva-Junior and Grattapaglia, 2015), carry out genome-wide association (Resende et al., 2016; Muller et al., 2017) and drive genomic selection in different *Eucalyptus* species (Duran et al., 2017; Resende et al., 2017; Tan et al., 2017, 2018). *Eucalyptus* transcriptomics resources are available at https://eucgenie.org/. Ongoing and future efforts is focusing on understanding genome diversity and evolution in this species-rich genus and other members of the Myrtaceae (Grattapaglia et al., 2012), including sister genera such as *Corymbia* for which a genome assembly, soon to be released, revealed conservation but dynamic evolution of terpene genes relative to *Eucalyptus* (Butler et al., 2018).

Salix

Whole-genome analysis in willow (*Salix* spp.), including development of reference genome assemblies, have focused on four species: *Salix purpurea* L. (Section Helix), *S. suchowensis* Cheng Section Helix), *S. wilsonii* Seemen (Section Wilsonianae) and *S. viminalis* L. (Section Vimen). The genome of a female *S. purpurea* '94006' was assembled at the J. Craig Venter Institute from Illumina sequence generated at the DOE Joint Genome Institute, including mate-pair library reads, and scaffolds were ordered using a genetic map generated from an F$_2$ mapping population genotyped using genotyping-by-sequencing at Cornell University[1], placing ∼70% of the genome (276 Mb) in chromosome-scale pseudomolecules. The *S. suchowensis* genome was sequenced using a combination of Roche 454 technology and Illumina/HiSeq 2000 reads (Dai et al., 2014). Additional genome sequencing projects are currently underway in *S. purpurea*, *S. wilsonii,* and *S. viminalis* using deep PacBio sequencing, which should result in much more contiguous and complete genomes. Sequence information and gene assemblies for multiple *Salix* species are now available at TreeGenes[5].

The genome sequencing projects have revealed three notable differences between the *Populus* and *Salix* genomes. First, the *Salix* genomes are clearly smaller than the *Populus* genomes. Using 17-mer analysis, the *S. suchowensis* genome size was estimated at ∼425 Mb (Dai et al., 2014), while the *S. purpurea* genome was estimated to be ∼379 Mb by 25-mer analysis, both of which are slightly smaller than the estimate of ∼429 and 450 Mb from flow cytometry of each species, respectively. Notwithstanding, all of these estimates are smaller than the genome size of *P. trichocarpa* (∼485 Mb). The *Salix* genomes also show evidence of more extensive fractionation following the whole-genome Salicoid duplication that is shared by *Salix* and *Populus* (Tuskan et al., 2006), resulting in a reduced number of predicted genes and overall smaller genome size (Dai et al., 2014). Second, it is clear that the large chromosome 1 in *Populus* corresponds to portions of chromosome 1 and chromosome 16 in *Salix*, suggesting a series of major chromosomal fission and fusion events occurred after the Salicoid duplication leading to the divergence of these lineages (Berlin et al., 2010; Hou

---

[5]https://www.treegenesdb.org/

et al., 2016). Finally, the sex determination locus is located on chromosome 15 in *Salix*, while it is on chromosome 19 in all *Populus* species studied to date, supporting the hypothesis that there has been greater genome fractionation in *Salix* following the Salicoid duplication event. Furthermore, most *Populus* species have an XY sex determination system (Tuskan et al., 2012; Geraldes et al., 2015), while all sequenced *Salix* species have a ZW system (Hou et al., 2015; Pucholt et al., 2015; Chen et al., 2016).

Given the apparently dynamic nature of sex determination in this dioecious family, rigorous efforts to understand this trait have been undertaken in recent years. Extensive differential gene expression has been documented in shoot tips of *S. purpurea*, which display an abundance of female-biased gene expression (Carlson et al., 2017). Expression analysis in *S. viminalis* suggests a predominance of male-biased gene expression, perhaps due to masculinization of the Z chromosome (Pucholt et al., 2017). Analysis of differentially expressed genes in unflushed buds of *S. suchowensis* identified 806 differentially expressed genes between males and females (Liu et al., 2015).

There are also extensive genetic mapping resources to support efforts to relate genotype to phenotype and accelerate willow breeding using molecular tools. In addition to the F_2 map used to anchor the *S. purpurea* genome assemblies (Kopp et al., 2002), more than a dozen linkage mapping populations have been developed for *S. viminalis* (Karp et al., 2011), as well as an association mapping population (Hallingbäck et al., 2016). New mapping populations have also been developed based on an approach of using *S. purpurea* male '94001' and female '94006' reference individuals as common parents crossed with a number of other species, including *S. viminalis*, *S. suchowensis*, *S. integra*, *S. koriyanagi*, and *S. udensis* at Cornell University. At Nanjing Forestry University, an effort has been employed to establish recombinant inbred lines for *S. suchowensis*, with an inbred F_4 pedigree currently established in the field.

Quercus

A highly contiguous haploid genome of a heterozygous pedunculate oak (*Quercus robur*) was recently generated based on Illumina synthetic long-reads and 454 (Roche) sequences (Plomion et al., 2018). The assembly contains 1,409 scaffolds totaling 814 Mb (N50, 1.34 Mb)[6]. Overall 871 scaffolds, representing 96% of the physical size of the genome, were anchored to the 12 chromosomes. Based on RNA-seq and protein homologies, 25,808 genes were predicted. Two other draft genomes, while of lower quality, have also been published (Sork et al., 2016; Schmid-Siegert et al., 2017). As in *Eucalyptus*, one of the main features of the oak genome is its remarkably high level of proximal tandem duplication (35.6% of the gene models). The tight relationships between duplicated genes and lineage-specific selection already reported in other species was found to be particularly exacerbated in oak, as three quarters of the genes in expanded orthogroups were also found in tandem duplications. Another interesting feature of long-lived tree species rests in the accumulation of somatic mutations. The presence and transmission of somatic mutations was recently demonstrated

in oak (Schmid-Siegert et al., 2017; Plomion et al., 2018), which raises interesting questions related to the role of somatic mutation in the evolution of long-lived species with a potentially high mutational load.

One of the major challenges in plant biology is identifying and characterizing the genes responsible for variation in ecologically and economically important traits. Genetic architecture of traits related to growth, phenology, water metabolism, and defense against pathogens has been explored in different full-sib oak pedigrees (e.g., Brendel et al., 2008) and natural populations (e.g., Alberto et al., 2013). Linkage and association mapping revealed many underlying loci with low to moderate contributions to the trait of interest. The characterization of the oak transcriptome is more recent. Gene repertoires were first constructed (Ueno et al., 2010; Pereira-Leal et al., 2014; Cokus et al., 2015; Lesur et al., 2015) and RNA-seq data were then used to (1) analyze molecular plasticity under abiotic stress (e.g., drought, Spieß et al., 2012) and biotic interactions (e.g., with insects, Kersten et al., 2013) and (2) identify genes and gene networks underlying developmental mechanisms (e.g., bud phenology, Ueno et al., 2013 and acorn development, Miguel et al., 2015) and evolutionary processes including reproductive isolation (Le Provost et al., 2012, 2016) and local adaptation (Gugger et al., 2016a). Recent studies provide evidence for epigenetic variation in phenotypic plasticity and evolutionary processes in oaks (Gugger et al., 2016b), establishing new avenues for research into the role of epigenetics in trait plasticity for these long-lived species.

Shortlisted among the 'botanical horror stories' (Rieseberg et al., 2006), the *Quercus* genus constitutes an ideal taxon to investigate the dynamics of lineage diversification along a wide and fluid continuum of speciation. Several efforts to illuminate oak evolutionary histories using population genomic approaches have recently emerged (Ortego et al., 2016, 2018; Leroy et al., 2017, 2018). In European white oaks for example, the best-supported scenario of divergence is consistent with a long-term pervasive effect of interspecific gene flow, with the exception of some narrow genomic regions responsible for reproductive isolation. Further work will be required to identify genes targeted by intrinsic or ecological selection, but early attempts appear promising and could enable predictions of the evolutionary responses of oaks to climate change (Gugger et al., 2016a; Rellstab et al., 2016). An exciting perspective in population genomics is presented by allochronic approaches that compare ancient and modern DNA to derive evolutionary changes at targeted genomic regions. Wagner et al. (2018) have suggested ancient DNA could be retrieved from waterlogged archeological or fossil samples, thus enabling the incremental reconstruction of evolutionary trajectories.

CHESTNUT, PECAN, AND OTHER EARLY DRAFT ANGIOSPERM GENOMES

Other reference genome projects for hardwoods are underway, including efforts to produce reference sequences for chestnut and pecans. The chestnut project aims to produce reference genomes of both *Castanea mollissima* (Chinese chestnut), the

---

[6]http://www.oakgenome.fr/

donor of resistance to *Cryphonectria parasitica* (chestnut blight) and *Castanea dentata* (American chestnut) the species destroyed by the blight in the early 1900s. The Forest Health Initiative (Nelson et al., 2014) supported a project to sequence the genome of the Chinese chestnut donor parent tree 'Vanuxem' used by The American Chestnut Foundation's (TACF). The first draft of the Chinese chestnut genome assembly was released in January 2014 at: https://www.hardwoodgenomics.org/chinese-chestnut-genome, consisting of 41,260 scaffolds averaging 39.6 kb in length and covering 98% (~724 Mb) of the genome. Targeted regions of mapped QTL resistance loci from various clones have also been produced (Staton et al., 2015). In addition, work is underway to prepare chromosome-length sequences for the Chinese chestnut genome. Long, single-molecule PacBio data are being used to merge contigs, followed by anchoring of scaffolds to high-density linkage maps for the cv. Vanuxem. Fluorescent *in situ* hybridization is being used to validate the chromosome-level assemblies and subsequently to identify structural rearrangements among genotypes and species (Islam-Faridi et al., 2016). The Chinese chestnut reference genome is being utilized to develop genomic selection models for blight resistance in American chestnut backcross populations (Westbrook, 2018). The project to sequence the American chestnut has produced a contig assembly, based on PacBio sequencing, of 8.1 Mb contig N50 and efforts are now underway to use new technologies (see below) and genetic mapping to produce chromosome-scale assemblies for other *C. mollissima* and *C. dentata* genotypes.

Initial efforts at generating a *Carya illinoinensis* (pecan) reference genome, released in 2013 and based on Illumina short-read sequencing (Jenkins et al., 2015) resulted in a contig N50 of 6.5 kb and scaffold N50 of 11.2 kb. The Mexican accession '87MX3-2.11' from the USDA-ARS facility in Somerville, Texas was selected as the reference genome based on a reduced heterozygosity. Efforts are underway to generate PacBio-based genomes for pecans that include 87MX3-2 and three production cultivars[7]. The current version of the pre-chromosome assembly for 87MX3-2 covers 705.6 Mb of the genome with a contig N50 of 2.5 Mb with 36,489 genes annotated[1]. Multiple early efforts are also underway to sequenced the European ash (*Fraxinus excelsior*) and American ash (*F. americana*)[8,9], as well as several birch genomes (*Betula nana*, *B. pubescens,* and *B. pendula*) (Wang et al., 2013). Finally, several angiosperm genomes are in early draft form and include species of ash[5] and walnut[9], as well as several comparative angiosperms from horticultural species[10].

A LOOK TO THE NEAR FUTURE

The continued development of new technologies that can be applied to generate high-quality hardwood reference genomes has been fostered by single-molecule, real-time

sequencing (SMRT) and related assembly algorithms from PacBio (Eid et al., 2009). These long reads, which average 10–15 kb in length, but are frequently 20–50 kb in length, can be used to resolve complex genomic repeats and improve contiguity of genomes. An early example of a small, inbred, grass genome (*Oropetium thomaeum*) achieved a contig N50 of 2.4 Mb, while capturing almost the entire genome space (244 out of 245 Mb) (VanBuren et al., 2015). The PacBio long reads were also used to produce a reference for the large, complex sunflower (*Helianthus annuus*) genome (Badouin et al., 2017) at 3.6 Gb with a contig N50 of 524 kb. An alternative reference to the *Zea mays* B73 supports using PacBio Iso-Seq sequencing for collecting full-length cDNAs to enhance the reference annotation and a genome (Jiao et al., 2017), which despite large-scale nested transposon activity, was completed at a contig N50 of 1.18 Mb. Because of the low accuracy rate of individual PacBio reads, these reads must be combined with an alternative technology, such as Illumina short reads, to polish the consensus sequence. For hardwood tree genomes, the longer reads facilitate the generation of outbred reference sequences to sort haplotypes and produce a single reference despite high heterozygosity rates (see *Q. robur*, *P. deltoides* WV94 and *Castanea dentata* described above). In addition to long-read sequencing, which has brought us back to the completeness of Sanger-based sequencing, new tools are emerging to achieve accurate chromosome-scale contiguity. These tools for genomic mapping and assembly include improved single-molecule mapping from BioNano (Staňková et al., 2016), *in vitro* reconstituted chromatin (Putnam et al., 2016) and binned sequencing approaches such as 10x genomics (Coombe et al., 2016). The most promising of these technologies is Hi-C for ordering large contigs from genome assemblies. Originally developed to perform "contact mapping" in human cell lines to show genes adjacent to promoter or regulatory elements (Suhas, 2014), it has been repurposed as a general solution to determining order in genome assemblies (Dudchenko et al., 2017). A recent example, the 4.8 Gb genome of *Hordeum vulgare*, has shown the ability to use Hi-C to precisely place 95% of contigs on pseudomolecules (Mascher et al., 2017). All of these tools are actively being applied to improve hardwood genomic references as a foundation for accurate population and functional analysis.

These new technologies will enable greater within and across species comparisons. An example of unique tree biology, that can be explored, is the characterization of somatic mutations. Due to their perennial habit, trees can accumulate somatic mutations in alternate vegetative lineages (as noted in the *Populus* and *Quercus* sections above). However, the effect of somatic variation on the generation of new genetic variation at the population-level and/or during reproduction remains largely unknown. Since, in plants, germlines are not segregated from the vegetative lineages, somatic mutations can be transmitted to the next generation. The frequency with which a given somatic mutation propagates to the next generation depends both on the overall fitness of the tree and the probability of the somatic sector giving rise to flowers and gametes. Because branches/tissue types do not make equal contributions to the resulting gene pool, somatic mutations may alter the adaptive balance between branches. This

---

[7] pecantoolbox.nmsu.edu

[8] http://www.ashgenome.org/

[9] http://hardwoodgenomics.org

[10] https://www.rosaceae.org/

mosaic genetic architecture raises the possibility that selection acts simultaneously on both the branch and tree (Hadany, 2001; Clarke, 2011). Understanding the role of multi-level selection within a single tree will require answering several questions. That is, to what extent is the effect of a somatic mutation in a single branch shared with the entire tree (Folse and Roughgarden, 2011)? And, if a branch acquires a broadly-beneficial mutation, such as one conferring resistance to herbivory (e.g., Edwards et al., 1990), does the resistant branch gain a larger fitness benefit than the remainder of the tree? Conversely, can a mutation increase the fitness of a single branch despite imposing cost on the rest of the tree? For example, a mutation that increases flowering in a branch might increase seed production from that branch while depressing production from the tree as a whole by acting as a metabolic sink (Walbot, 1985). Finally, should these conflicts occur, do trees have mechanisms to mediate resource allocation among branches, and, if so, are these mechanisms common between independent tree lineages? New techniques (described above) for finding or constructing whole-genome genetic mosaic and chimeric trees with phenotypically-relevant somatic mutations are now allowing these questions to be addressed.

In summary, the assembly and annotation of multiple hardwood tree genomes has facilitated an increase in (1) functional characterization of genes and gene networks related to tree habit, (2) GWAS and genomic selection investigations and applications, and (3) comparative genomic efforts among various tree species. The rapidly expanding new technologies will add even greater number of hardwood species to these efforts. The power of comparative genomics will increase our understanding of how these highly dynamic genomes have evolved and resulted in the amazing array of phenotypic diversity found among and within hardwood species. To broaden the raft of available hardwood genomes, resources should be directed toward additional candidate hardwood genera, including but is not limited to *Liriodendron*, *Liquidambar*, *Swietenia*, and *Acer*. In sum total, most hardwoods are undomesticated long-lived plants that provide many ecological and commercial benefits, whose management, conservation and domestication for economic and ecological purposes will benefit from a set of rich genomics resources. Ultimately such resources will favorably impact the pressing problems of climate change, soil and water conservation, bioenergy and biomaterials production, and maintenance of heathy ecosystem functions. As a result, we may finally answer the questions of: (1) why has the tree habit evolved repeatedly in the angiosperms and (2) what is the connection between the genomes and the defining characteristics of long-lived perennial plants?

AUTHOR CONTRIBUTIONS

All authors contributed equally to the formatting and writing contained in this Mini-review.

FUNDING

REFERENCES

Albert, V. A., Barbazuk, W. B., Der, J. P., Leebens-Mack, J., Ma, H., Palmer, J. D., et al. (2013). The *Amborella* genome and the evolution of flowering plants. *Science* 342:1241089. doi: 10.1126/science.1241089

Alberto, F. J., Derory, J., Boury, C., Frigério, J. M., Zimmermann, N. E., and Kremer, A. (2013). Imprints of natural selection along environmental gradients in phenology-related genes of *Quercus petraea*. *Genetics* 195, 495–512. doi: 10.1534/genetics.113.153783

Badouin, H., Gouzy, J., Grassa, C. J., Murat, F., Staton, S. E., Cottret, L., et al. (2017). The sunflower genome provides insights into oil metabolism, flowering and asterid evolution. *Nature* 546, 148–152. doi: 10.1038/nature22380

Berlin, S., Lagercrantz, U., von Arnold, S., Öst, T., and Rönnberg-Wästljung, A. C. (2010). High-density linkage mapping and evolution of paralogs and orthologs in *Salix* and *Populus*. *BMC Genomics* 11:129. doi: 10.1186/1471-2164-11-129

Brendel, O., Le Thie, D., Scotti-Saintagne, C., Bodénès, C., Kremer, A., and Guehl, J. M. (2008). Detection of quantitative trait loci controlling water use efficiency and related traits in *Quercus robur* L. *Tree Genet. Genomes* 4, 263–278. doi: 10.1007/s11295-007-0107-z

Butler, J. B., Freeman, J. S., Potts, B. M., Vaillancourt, R. E., Grattapaglia, D., Silva-Junior, O. B., et al. (2018). Annotation of the *Corymbia* terpene synthase gene family shows broad conservation but dynamic evolution of physical clusters relative to *Eucalyptus*. *Heredity* 121, 87–104. doi: 10.1038/s41437-018-0058-1

Carlson, C. H., Choi, Y., Chan, A. P., Serapiglia, M. J., Town, C. D., and Smart, L. B. (2017). Dominance and sexual dimorphism pervade the *Salix purpurea* l. transcriptome. *Genome Biol. Evol.* 9, 2377–2394. doi: 10.1093/gbe/evx174

Carocha, V., Soler, M., Hefer, C., Cassan-Wang, H., Fevereiro, P., Myburg, A. A., et al. (2015). Genome-wide analysis of the lignin toolbox of *Eucalyptus grandis*. *New Phytol.* 206, 1297–1313. doi: 10.1111/nph.13313

Chen, Y., Wang, T., Fang, L., Li, X., and Yin, T. (2016). Confirmation of single-locus sex determination and female heterogamety in willow based on linkage analysis. *PLoS One* 11:e0147671. doi: 10.1371/journal.pone.0147671

Christie, N., Tobias, P. A., Naidoo, S., and Kulheim, C. (2015). The *Eucalyptus grandis* NBS-LRR gene family: physical clustering and expression hotspots. *Front. Plant. Sci.* 6:1238. doi: 10.3389/fpls.2015.01238

Clarke, E. (2011). "Plant individuality and multilevel selection theory," in *Major Transitions in Evolution Revisited*, eds B. Calcott and K. Sterelny (Cambridge, MA: The MIT Press).

Cokus, S. J., Gugger, P. F., and Sork, V. L. (2015). Evolutionary insights from de novo transcriptome assembly and SNP discovery in California white oaks. *BMC Genomics* 16:552. doi: 10.1186/s12864-015-1761-4

Coombe, L., Warren, R. L., Jackman, S. D., Yang, C., Vandervalk, B. P., Moore, R. A., et al. (2016). Assembly of the complete Sitka spruce chloroplast genome using 10x genomics gemcode sequencing data. *PLoS One* 11:e0163059. doi: 10.1371/journal.pone.0163059

Dai, X., Hu, Q., Cai, Q., Feng, K., Ye, N., Tuskan, G. A., et al. (2014). The willow genome and divergent evolution from poplar after the common genome duplication. *Cell Res.* 24, 1274–1277. doi: 10.1038/cr.2014.83

Davin, N., Edger, P. P., Hefer, C. A., Mizrachi, E., Schuetz, M., Smets, E., et al. (2016). Functional network analysis of genes differentially expressed during xylogenesis in soc1ful woody *Arabidopsis* plants. *Plant J.* 86, 376–390. doi: 10.1111/tpj.13157

Dudchenko, O., Batra, S. S., Omer, A. D., Nyquist, S. K., Hoeger, M., Durand, N. C., et al. (2017). De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science* 356:92. doi: 10.1126/science.aal3327

Duran, R., Isik, F., Zapata-Valenzuela, J., Balocchi, C., and Valenzuela, S. (2017). Genomic predictions of breeding values in a cloned *Eucalyptus globulus* population in Chile. *Tree Genet. Genomes* 13:74. doi: 10.1007/s11295-017-1158-4

Edwards, P., Wanjura, W., Brown, W., and Dearn, J. (1990). Mosaic resistance in plants. *Nature* 347:434. doi: 10.1038/347434a0

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., et al. (2009). Real-time DNA sequencing from single polymerase molecules. *Science* 323, 133–138. doi: 10.1126/science.1162986

Evans, L. M., Slavov, G. T., Rodgers-Melnick, E., Martin, J., Ranjan, P., Muchero, W., et al. (2014). Population genomics reveals signatures of selection coupled with adaptive trait associations in the model tree, *Populus trichocarpa*. *Nat. Genet.* 46, 1089–1096. doi: 10.1038/ng.3075

Fahrenkrog, A. M., Neves, L. G., Resende, M. F. R. Jr, Gustavo de los Campos AIV, Dervinis, C., et al. (2017).Genome-wide association study reveals putative regulators of bioenergy traits in *Populus* deltoides. *New Phytol.* 213, 799–811. doi: 10.1111/nph.14154

FAO (2016). *State of the World's Forests 2016*. Available at: http://www.fao.org/publications/sofo/2016/en/

Folse, H. J. III, and Roughgarden, J. (2011). Direct benefits of genetic mosaicism and intraorganismal selection: modeling coevolution between a long-lived tree and a short-lived herbivore. *Evolution* 66, 1091–1113. doi: 10.1111/j.1558-5646.2011.01500.x

Geraldes, A., Hannemann, J., Grassa, C., Farzaneh, N., Porth, I., McKown, A., et al. (2013). A 34K SNP genotyping array for *Populus trichocarpa*: design, application to the study of natural populations and transferability to other *Populus* species. *Mol. Ecol. Resour.* 13, 306–323. doi: 10.1111/1755-0998.12056

Geraldes, A., Hefer, C. A., Capron, A., Kolosova, N., Martinez-Nuñez, F., Soolanayakanahally, R. Y., et al. (2015). Recent Y chromosome divergence despite ancient origin of dioecy in poplars (*Populus*). *Mol. Ecol.* 24, 3243–3256. doi: 10.1111/mec.13126

Grattapaglia, D., Vaillancourt, R. E., Shepherd, M., Thumma, B. R., Foley, W., Külheim, C., et al. (2012). Progress in *Myrtaceae* genetics and genomics: *Eucalyptus* as the pivotal genus. *Tree Genet. Genomes* 8, 463–508. doi: 10.1007/s11295-012-0491-x

Grattapaglia, D., and Bradshaw, H. D. (1994). Nuclear DNA content of commercially important *Eucalyptus* species and hybrids. *Can. J. For. Res.* 24, 1074–1078. doi: 10.1139/x94-142

Groover, A., and Crook, Q., (eds) (2017). *Comparative and Evolutionary Genomics of Angiosperm Trees*. Berlin: Springer Publishing. 366. doi: 10.1007/978-3-319-49329-9

Gugger, P. F., Cokus, S. J., and Sork, V. L. (2016a). Association of transcriptome-wide sequence variation with climate gradients in valley oak (*Quercus lobata*). *Tree Genet. Genomes* 12:15. doi: 10.1007/s11295-016-0975-1

Gugger, P. F., Fitz-Gibbon, S., Pellegrini, M., and Sork, V. L. (2016b). Species-wide patterns of DNA methylation variation in *Quercus lobata* and their association with climate gradients. *Mol. Ecol.* 25, 1665–1680. doi: 10.1111/mec.13563

Hadany, L. (2001). A conflict between two evolutionary levels in trees. *J. Theor. Biol.* 208, 507–521. doi: 10.1006/jtbi.2000.2236

Hallingbäck, H. R., Fogelqvist, J., Powers, S. J., Turrion-Gomez, J., Rossiter, R., Amey, J., et al. (2016). Association mapping in *Salix viminalis* L. (Salicaceae) – identification of candidate genes associated with growth and phenology. *GCB Bioenergy* 8, 670–685. doi: 10.1111/gcbb.12280

Hedrick, P., Hellesten, U., and Grattapaglia, D. (2016). Examining the cause of high inbreeding depression: analysis of whole-genome sequence data in 28 selfed progeny of *Eucalyptus grandis*. *New Phytol.* 209, 600–611. doi: 10.1111/nph.13639

Hefer, C. A., Mizrachi, E., Myburg, A. A., Douglas, C. J., and Mansfield, S. D. (2015). Comparative interrogation of the developing xylem transcriptomes of two wood-forming species: *Populus trichocarpa* and *Eucalyptus grandis*. *New Phytol.* 206, 1391–1405. doi: 10.1111/nph.13277

Henry, I. M., Zinkgraf, M. S., Groover, A. T., and Comai, L. (2015). A system for dosage-based functional genomics in poplar. *Plant Cell* 27, 2370–2383.

Hou, J., Ye, N., Dong, Z., Lu, M., Li, L., and Yin, T. (2016). Major chromosomal rearrangements distinguish willow and poplar after the ancestral "Salicoid" genome duplication. *Genome Biol. Evol.* 8, 1868–1875. doi: 10.1093/gbe/evw127

Hou, J., Ye, N., Zhang, D., Chen, Y., Fang, L., Dai, X., et al. (2015). Different autosomes evolved into sex chromosomes in the sister genera of *Salix* and *Populus*. *Sci. Rep.* 5:9076. doi: 10.1038/srep09076

Hussey, S. G., Saidi, M. N., Hefer, C. A., Myburg, A. A., and Grima-Pettenati, J. (2015). Structural, evolutionary and functional analysis of the NAC domain protein family in *Eucalyptus*. *New Phytol.* 206, 1337–1350. doi: 10.1111/nph.13139

Islam-Faridi, N., Majid, M. A., Zhebentyayeva, T., Georgi, L. L., Fan, S., Hebard, V., et al. (2016). FISH confirmation of a reciprocal translocation in chestnut. *Cytogenet. Genome Res.* 148, 144–144.

Jenkins, J., Wilson, B., Grimwood, J., Schmutz, J., and Grauke, L. (2015). Towards a reference pecan genome sequence. *Acta Hortic.* 1070, 101–108. doi: 10.17660/ActaHortic.2015.1070.11

Jiao, Y., Peluso, P., Shi, J., Liang, T., Stitzer, M. C., Wang, B., et al. (2017). Improved maize reference genome with single-molecule technologies. *Nature* 546, 524–527. doi: 10.1038/nature22971

Joubert, W., Nance, J., Climer, S., Weighill, D., and Jacobson, D. (2017). Parallel accelerated custom correlation coefficient calculations for genomics applications. *arXiv* [Preprint]. arXiv:1705.08213

Joubert, W., Nance, J., Weighill, D., and Jacobson, D. (2018). Parallel accelerated vector similarity calculations for genomics applications. *Parallel Comput.* 75, 130–145. doi: 10.1016/j.parco.2018.03.009

Karp, A., Hanley, S. J., Trybush, S. O., Macalpine, W., Pei, M., and Shield, I. (2011). Genetic improvement of willow for bioenergy and biofuels. *J. Integr. Plant Biol.* 53, 151–165. doi: 10.1111/j.1744-7909.2010.01015.x

Kersten, B., Ghirardo, A., Schnitzler, J. P., Kanawati, B., Schmitt-Kopplin, P., Fladung, M., et al. (2013). Integrated transcriptomics and metabolomics decipher differences in the resistance of pedunculate oak to the herbivore *Tortrix viridana* L. *BMC Genomics* 14:737. doi: 10.1186/1471-2164-14-737

Kopp, F., Smart, B., Maynard, A., Tuskan, A., and Abrahamson, P. (2002). Predicting within-family variability in juvenile height growth of *Salix* based upon similarity among parental AFLP fingerprints. *Theor. Appl. Genet.* 105, 106–112.

Kulheim, C., Padovan, A., Hefer, C., Krause, S. T., Kollner, T. G., Myburg, A. A., et al. (2015). The *Eucalyptus terpene* synthase gene family. *BMC Genomics* 16:450. doi: 10.1186/s12864-015-1598-x

Le Provost, G., Sulmon, C., Frigerio, J. M., Bodenes, C., Kremer, A., and Plomion, C. (2012). Role of waterlogging genes in shaping interspecific differentiation between two sympatric oak species. *Tree Physiol.* 32, 119–134. doi: 10.1093/treephys/tpr123

Le Provost, G., Lesur, I., Lalanne, C., Da Silva, C., Labadie, K., Aury, J. M., et al. (2016). Implication of the suberin pathway in adaptation to waterlogging and hypertrophied lenticels formation in pedunculate oak (*Quercus robur* L.) *Tree Physiol.* 36, 1330–1342. doi: 10.1093/treephys/tpw056

Leroy, T., Roux, C., Villate, L., Bodénès, C., Romiguier, J., Paiva, J. A. P., et al. (2017). Extensive recent secondary contacts between four European white oak species. *New Phytol.* 214, 865–878. doi: 10.1111/nph.14413

Leroy, T., Rougemont, Q., Bodénès, C., Lalanne, C., Belser, C., Labadie, K., et al. (2018). Secondary contacts between European white oaks reveal genes underlying reproductive isolation. *bioRxiv* [Preprint]. doi: 10.1101/246637

Lesur, I., Le Provost, G., Bento, P., Da Silva, C., Leplé, J. C., Murat, F., et al. (2015). The oak gene expression Atlas: insights into Fagaceae genome evolution and the discovery of genes regulated during bud dormancy release. *BMC Genomics* 16:112. doi: 10.1186/s12864-015-1331-9

Liu, J., Ye, M., Zhu, S., Jiang, L., Sang, M., Gan, J., et al. (2018). Two-stage identification of SNP effects on dynamic poplar. *Plant J.* 93, 286–296. doi: 10.1111/tpj.13777

Liu, L., Ramsay, T., Zinkgraf, M., Sundell, D., Street, N. R., Filkov, V., et al. (2015). A resource for characterizing genome-wide binding and putative target genes of transcription factors expressed during secondary growth and wood formation in *Populus*. *Plant J.* 82, 887–898.

Ma, T., Wang, J., Zhou, G., Yue, Z., Hu, Q., Chen, Y., et al. (2014). Genomic insights into salt adaptation in a desert poplar. *Nat. Commun.* 5:2797. doi: 10.1038/ncomms3797

Mangwanda, R., Myburg, A. A., and Naidoo, S. (2015). Transcriptome and hormone profiling reveals *Eucalyptus grandis* defence responses against *Chrysoporthe austroafricana*. *BMC Genomics* 16:319. doi: 10.1186/s12864-015-1529-x

Mascher, M., Gundlach, H., Himmelbach, A., Beier, S., Twardziok, S. O., Wicker, T., et al. (2017). A chromosome conformation capture ordered sequence of the barley genome. *Nature* 544:427. doi: 10.1038/nature22043

McKown, A. D., Klápštì, J., Guya, R. D., Geraldes, A., Portha, I., Hannemanne, J., et al. (2014). Genome-wide association implicates numerous genes underlying ecological trait variation in natural populations of *Populus trichocarpa*. *New Phytol.* 203, 535–553. doi: 10.1111/nph.12815

Mewalal, R., Rai, D. K., Kainer, D., Chen, F., Külheim, C., Peter, G. F., et al. (2017). Plant-derived terpenes: a feedstock for specialty biofuels. *Trends Biotechnol.* 35, 227–240. doi: 10.1016/j.tibtech.2016.08.003

Meyer, F. E., Shuey, L. S., Naidoo, S., Mamni, T., Berger, D. K., Mybur, A. A., et al. (2016). Dual RNA-sequencing of *Eucalyptus nitens* during *Phytophthora cinnamomi* challenge reveals pathogen and host factors influencing compatibility. *Front. Plant Sci.* 7:191. doi: 10.3389/fpls.2016.00191

Miguel, A., Vega Bartol, J., Marum, L., Chaves, I., Santo, T., Leitão, J., et al. (2015). Characterization of the cork oak dynamics during acorn development. *BMC Plant Biol.* 15:258. doi: 10.1186/s12870-015-0534-1

Mizrachi, E., Maloney, V. J., Silberbauer, J., Hefer, C. A., Berger, D. K., Mansfield, S. D., et al. (2014). Investigating the molecular underpinnings underlying morphology and changes in carbon partitioning during tension wood formation in *Eucalyptus*. *New Phytol.* 206, 1351–1363. doi: 10.1111/nph.13152

Mizrachi, E., Verbeke, L., Christie, N., Fierro, A. C., Mansfield, S. D., Davis, M. F., et al. (2017). Network-based integration of systems genetics data reveals pathways associated with lignocellulosic biomass accumulation and processing. *Proc. Natl. Acad. Sci. U.S.A.* 114, 1195–1200. doi: 10.1073/pnas.1620119114

Muchero, W., Guo, J., DiFazio, S. P., Chen, J. G., Ranjan, P., Slavov, G. T., et al. (2015). High-resolution genetic mapping of allelic variants associated with cell wall chemistry in *Populus*. *BMC Genomics* 16:24. doi: 10.1186/s12864-015-1215-z

Muller, B. S. F., Neves, L. G., Almeida-Filho, J. E., Resende, M. F. R. J., Munoz Del, Valle, P., et al. (2017). Genomic prediction in contrast to a genome-wide association study in explaining heritable variation of complex growth traits in breeding populations of *Eucalyptus*. *BMC Genomics* 18:524. doi: 10.1186/s12864-017-3920-2

Myburg, A. A., Grattapaglia, D., Tuskan, G. A., Hellsten, U., Hayes, R. D., Grimwood, J., et al. (2014). The genome of *Eucalyptus grandis*. *Nature* 510, 356–362. doi: 10.1038/nature13308

Nelson, C. D., Powell, W. A., Maynard, C. A., Baier, K. M., Newhouse, A., Merkle, S. A., et al. (2014). The forest health initiative, American chestnut (*Castanea dentata*) as a model for forest tree restoration: biological research program. *Acta Hortic.* 1019, 179–189.

Oates, C. N., Kulheim, C., Myburg, A. A., Slippers, B., and Naidoo, S. (2015). The transcriptome and terpene profile of *Eucalyptus grandis* reveals mechanisms of defense against the insect pest, *Leptocybe invasa*. *Plant Cell Physiol.* 56, 1418–1428. doi: 10.1093/pcp/pcv064

Ortego, J., Gugger, P. F., and Sork, V. L. (2018). Genomic data reveal cryptic lineage diversification and introgression in Californian golden cup oaks (section *Protobalanus*). *New Phytol.* 2, 804–818. doi: 10.1111/nph.14951

Ortego, J., Gutter, P. F., and Sork, V. L. (2016). Impacts of human-induced environmental disturbances on hybridization between two ecologically differentiated Californian oak species. *New Phytol.* 213, 942–955. doi: 10.1111/nph.14182

Pereira-Leal, J. B., Abreu, I. A., Alabaça, C. S., Almeida, M. H., Almeida, P., Almeida, T., et al. (2014). A comprehensive assessment of the transcriptome of cork oak (*Quercus suber*) through EST sequencing. *BMC Genomics* 15:371. doi: 10.1186/1471-2164-15-371

Pinard, D., Mizrachi, E., Hefer, C. A., Kersting, A. R., Joubert, F., Douglas, C. J., et al. (2015). Comparative analysis of plant carbohydrate active enzymes and their role in xylogenesis. *BMC Genomics* 16:402. doi: 10.1186/s12864-015-1571-8

Pinosio, S., Giacomello, S., Faivre-Rampant, P., Taylor, G., Jorge, V., Le Paslier, M. C., et al. (2016). Characterization of the poplar pan-genome by genome-wide identification of structural variation. *Mol. Biol. Evol.* 33, 2706–2719. doi: 10.1093/molbev/msw161

Plett, J. M., Kohler, A., Khachane, A., Keniry, K., Plett, K. L., Martin, F., et al. (2015). The effect of elevated carbon dioxide on the interaction between *Eucalyptus grandis* and diverse isolates of *Pisolithus* sp. is associated with a complex shift in the root transcriptome. *New Phytol.* 206, 1423–1436. doi: 10.1111/nph.13103

Plomion, C., Aury, J. M., Amselem, J., Leroy, T., Murat, F., Duplessis, S., et al. (2018). Oak genome reveals facets of long lifespan. *Nat. Plants* 4, 440–452. doi: 10.1038/s41477-018-0172-3

Porth, I., Klapšte, J., Skyba, O., Geraldes, A., Hannemann, J., McKown, A. D., et al. (2013). Genome-wide association mapping for wood characteristics in *Populus* identifies an array of candidate SNPs. *New Phytol.* 200, 710–726. doi: 10.1111/nph.12422

Pucholt, P., Hallingbäck, H. R., and Berlin, S. (2017). Allelic incompatibility can explain female biased sex ratios in dioecious plants. *BMC Genomics* 18:251. doi: 10.1186/s12864-017-3634-5

Pucholt, P., Rönnberg-Wästljung, A.-C., and Berlin, S. (2015). Single locus sex determination and female heterogamety in the basket willow (*Salix viminalis* L.). *Heredity* 114, 575–583. doi: 10.1038/hdy.2014.125

Putnam, N. H., O'Connell, B. L., Stites, J. C., Rice, B. J., Blanchette, M., Calef, R., et al. (2016). Chromosome-scale shotgun assembly using an in vitro method for long-range linkage. *Genome Res.* 26:342. doi: 10.1101/gr.193474.115

Ragauskas, A. J., Beckham, G., Biddy, M., Chandra, R., Davis, M., Davison, B., et al. (2014). Lignin valorization in the biorefinery. *Science* 344, 709–714.

Rao, S. S., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., et al. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159, 1665–1680. doi: 10.1016/j.cell.2014.11.021

Rellstab, C., Zoller, S., Walthert, L., Lesur, I., Pluess, A. R., Graf, R., et al. (2016). Signatures of local adaptation in candidate genes of oaks (*Quercus* spp.) with respect to present and future climatic conditions. *Mol. Ecol.* 25, 5907–5924. doi: 10.1111/mec.13889

Resende, R. T., Resende, M. D. V., Silva, F. F., Azevedo, C. F., Takahashi, E. K., Silva-Junior, et al. (2017). Assessing the expected response to genomic selection of individuals and families in *Eucalyptus* breeding with an additive-dominant model. *Heredity* 119, 245–255. doi: 10.1038/hdy.2017.37

Resende, R. T., Resende, M. D. V., Silva, F. F., Azevedo, C. F., Takahashi, E. K., Silva-Junior, O. B., et al. (2016). Regional heritability mapping and genome-wide association identify loci for complex growth, wood and disease resistance traits in *Eucalyptus*. *New Phytol.* 213, 1287–1300. doi: 10.1111/nph.14266

Rieseberg, L. H., Wood, T. E., and Baack, E. J. (2006). The nature of plant species. *Nature* 440, 524–527. doi: 10.1038/nature04402

Schmid-Siegert, E., Sarkar, N., Iseli, C., Calderon, S., Gouhier-Darimont, C., Chrast, J., et al. (2017). Low number of fixed somatic mutations in a long-lived oak tree. *Nat. Plants* 3, 926–929. doi: 10.1038/s41477-017-0066-9

Silva-Junior, O. B., Faria, D. A., and Grattapaglia, D. (2015). A flexible multi-species genome-wide 60K SNP chip developed from pooled resequencing of 240 *Eucalyptus* tree genomes across 12 species. *New Phytol.* 206, 1527–1540. doi: 10.1111/nph.13322

Silva-Junior, O. B., and Grattapaglia, D. (2015). Genome-wide patterns of recombination, linkage disequilibrium and nucleotide diversity from pooled resequencing and single nucleotide polymorphism genotyping unlock the evolutionary history of *Eucalyptus grandis*. *New Phytol.* 208, 830–845. doi: 10.1111/nph.13505

Slavov, G. T., DiFazio, S. P., Martin, J., Schackwitz, W., Muchero, W., Rodgers-Melnick, E., et al. (2012). Genome resequencing reveals multiscale geographic structure and extensive linkage disequilibrium in the forest tree *Populus trichocarpa*. *New Phytol.* 196, 713–725. doi: 10.1111/j.1469-8137.2012.04258.x

Soler, M., Camargo, E. L., Carocha, V., Cassan-Wang, H., San Clemente, H., Savelli, B., et al. (2015). The *Eucalyptus grandis* R2R3-MYB transcription factor family: evidence for woody growth-related evolution and function. *New Phytol.* 206, 1364–1377. doi: 10.1111/nph.13039

Soltis, D. E., Albert, V. A., Leebens-Mack, J., Palmer, J. D., Wing, R. A., Ma, H., et al. (2008). The *Amborella* genome: an evolutionary reference for plant biology. *Genome Biol.* 9:402.

Sork, V. L., Fitz-Gibbon, S. T., Puiu, D., Crepeau, M., Gugger, P. F., Sherman, R., et al. (2016). First draft assembly and annotation of the genome of a California endemic oak *Quercus lobata* Née (Fagaceae). *G3* 6, 3485–3495. doi: 10.1534/g3. 116.030411

Spicer, R., and Groover, A. (2010). The evolution of development of the vascular cambium and secondary growth. *New Phytol.* 186, 577–592.

Spieß, N., Oufir, M., Matušíková, I., Stierschneider, M., Kopecky, D., Homolka, A., et al. (2012). Ecophysiological and transcriptomic responses of oak (*Quercus robur*) to long-term drought exposure and rewatering. *Environ. Exp. Bot.* 77, 117–126. doi: 10.1016/j.envexpbot.2011.11.010

Spokevicius, A. V., Tibbits, J., Rigault, P., Nolin, M. A., Muller, C., and Merchant, A. (2017). Medium term water deficit elicits distinct transcriptome responses in *Eucalyptus* species of contrasting environmental origin. *BMC Genomics* 18:284. doi: 10.1186/s12864-017-3664-z

Staňková, H., Hastie, A. R., Chan, S., Vrána, J., Tulpová, Z., Kubaláková, M., et al. (2016). BioNano genome mapping of individual chromosomes supports physical mapping and sequence assembly in complex plant genomes. *Plant Biotechnol. J.* 14, 1523–1531. doi: 10.1111/pbi.12513

Staton, M., Zhebentyayeva, T., Olukolu, B., Fang, G. C., Nelson, D., Carlson, J. E., et al. (2015). Substantial genome synteny preservation among woody angiosperm species: comparative genomics of Chinese chestnut (*Castanea mollissima*) and plant reference genomes. *BMC Genomics* 16:744. doi: 10.1186/ s12864-015-1942-1

Tan, B., Grattapaglia, D., Martins, G. S., Ferreira, K. Z., Sundberg, B., and Ingvarsson, P. K. (2017). Evaluating the accuracy of genomic prediction of growth and wood traits in two *Eucalyptus* species and their F1 hybrids. *BMC Genomics* 17:110. doi: 10.1186/s12870-017-1059-6

Tan, B., Grattapaglia, D., Wu, H. X., and Ingvarsson, P. K. (2018). Genomic relationships reveal significant dominance effects for growth in hybrid *Eucalyptus*. *Plant Sci.* 267, 84–93. doi: 10.1016/j.plantsci.2017.11.011

Tuskan, G. A., DiFazio, S. P., Faivre-Rampant, P., Gaudet, M., Harfouche, A., Jorge, V., et al. (2012). The obscure events contributing to the evolution of an incipient sex chromosome in *Populus*: a retrospective working hypothesis. *Tree Genet. Genomes* 8, 559–571. doi: 10.1007/s11295-012-0495-6

Tuskan, G. A., DiFazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313, 1596–1604. doi: 10.1126/science.1128691

Ueno, S., Klopp, C., Leplè, J. C., Derory, J., Noirot, C., Leger, V., et al. (2013). Transcriptional profiling of bud dormancy and release in oak by next-generation sequencing. *BMC Genomics* 14:236. doi: 10.1186/1471-2164- 14-236

Ueno, S., Le Provost, G., Léger, V., Klopp, C., Noirot, C., Frigerio, J. M., et al. (2010). Bioinformatic analysis of ESTs collected by Sanger and pyrosequencing

methods for a keystone forest tree species: oak. *BMC Genomics* 11:650. doi: 10.1186/1471-2164-11-650

VanBuren, R., Bryant, D., Edger, P. P., Tang, H. B., Burgess, D., Challabathula, D., et al. (2015). Single-molecule sequencing of the desiccation-tolerant grass *Oropetium thomaeum*. *Nature* 527, 508–511. doi: 10.1038/nature 15714

Vining, K. J., Pomraning, K. R., Wilhelm, L. J., Priest, H. D., Pellegrini, M., Mockler, T. C., et al. (2012). Dynamic DNA cytosine methylation in the *Populus trichocarpa* genome: tissue-level variation and relationship to gene expression. *BMC Genomics* 13:27. doi: 10.1186/1471-2164-13-27

Vining, K. J., Romanel, E., Jones, R. C., Klocko, A., Alves-Ferreira, M., Hefer, C. A., et al. (2015). The floral transcriptome of *Eucalyptus grandis*. *New Phytol.* 206, 406–1422. doi: 10.1111/nph.13077

Wagner, S., Lagane, F., Seguin-Orlando, A., Schubert, M., Leroy, T., Guichoux, E., et al. (2018). High-throughput DNA sequencing of ancient wood. *Mol. Ecol.* 27, 1138–1154. doi: 10.1111/mec.14514

Walbot, V. (1985). On the life strategies of plants and animals. *Trends Genet.* 1, 165–169. doi: 10.1016/0168-9525(85)90071-X

Wang, N., Thomson, M., Bodles, W. J. A., Crawford, R. M. M., Hunt, H. V., Featherstone, A. W., et al. (2013). Genome sequence of dwarf birch (*Betula nana*) and cross-species RAD markers. *Mol. Ecol.* 22, 3098–3111.

Weighill, D., Jones, P., Shah, M., Ranjan, P., Muchero, W., Schmutz, J., et al. (2018). Pleiotropic and epistatic network-based discovery: integrated networks for target gene discovery. *Front. Energy Res.* 6:30. doi: 10.3389/fenrg.2018. 00030

Westbrook, J. (2018). *Merging Backcross Breeding and Transgenic Blight Resistance to Accelerate Restoration of the American Chestnut: The American Chestnut Foundation's Breeding and Selection Plan 2015–2025.* Asheville, NC: The American Chestnut Foundation. Available at: www.acf.org

The Road to Resistance in Forest Trees

Sanushka Naidoo[1]*, Bernard Slippers[1], Jonathan M. Plett[2], Donovin Coles[2] and Caryn N. Oates[1]

[1]Division of Genetics, Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa, [2]Hawkesbury Institute for the Environment, Western Sydney University, Richmond, NSW, Australia

In recent years, forests have been exposed to an unprecedented rise in pests and pathogens. This, coupled with the added challenge of climate change, renders forest plantation stock vulnerable to attack and severely limits productivity. Genotypes resistant to such biotic challenges are desired in plantation forestry to reduce losses. Conventional breeding has been a main avenue to obtain resistant genotypes. More recently, genetic engineering has become a viable approach to develop resistance against pests and pathogens in forest trees. Tree genomic resources have contributed to advancements in both these approaches. Genome-wide association studies and genomic selection in tree populations have accelerated breeding tools while integration of various levels of omics information facilitates the selection of candidate genes for genetic engineering. Furthermore, tree associations with non-pathogenic endophytic and subterranean microbes play a critical role in plant health and may be engineered in forest trees to improve resistance in the future. We look at recent studies in forest trees describing defense mechanisms using such approaches and propose the way forward to developing superior genotypes with enhanced resistance against biotic stress.

Keywords: genomics, breeding, candidate genes, endophytes, genetic engineering

INTRODUCTION

Forest trees in their native range are introduced and domesticated as plantation stock encounter various biotic and abiotic factors that influence their persistence and productivity. In recent years, introduced pests and pathogens have been increasing at an alarming rate (Santini et al., 2013; Wingfield et al., 2015; Hurley et al., 2016). In some cases, these invasions threaten tree species with local extinction, such as the fungal pathogen *Cryphonectria parasitica*, that is responsible for the devastating chestnut blight that has all but eliminated the American chestnut from North America and led to novel transgenic strategies to save this tree (Newhouse et al., 2014). In recent years, *Austropuccinia psidii* has been introduced into various regions of the world and threaten some Myrtaceae with extinction (Roux et al., 2013; Granados et al., 2017; McTaggart et al., 2017). Additionally, various pine plantations have suffered losses from pitch canker disease caused by *Fusarium circinatum* (Gordon and Reynolds, 2017). Invasive pests can also cause major loss to forestry operations. For example, introduced *Eucalyptus* plantations have experienced devastation by the gall wasp, *Leptocybe invasa* (Dittrich-Schröder et al., 2012). At the same time, global climate change will exacerbate disease and pest incidence

(Sturrock et al., 2011; Das et al., 2016), as is the case with the devastating outbreaks of the mountain pine beetle (Six et al., 2018). The rapid increase in both biotic stress factors in many forests and forest plantations require a matching increase in options to mitigate their impact.

Plant defenses are complex and multiple signaling pathways may be induced following recognition of the attacker through effectors or molecular patterns specific to the invader (Jones and Dangl, 2006; reviewed in Naidoo et al., 2014). Resistance genes interact with effectors to induce effector-triggered immunity (ETI), accompanied by a hypersensitive response and containment of the invader. In the absence of such specific pathogen recognition, a basal level of defense may be induced during pathogenesis, although this may be too weak to fight off the invading microbe. Plant biotic stress interactions have been dissected in model herbaceous plants and have extended our knowledge of the sophisticated defense systems to different invaders. A challenge in forest trees is that the underlying resistance mechanisms are not always the same as described or predicted in model systems necessitating study of defense pathways and strategies utilized by tree species directly. Resistance may be polygenic (quantitative) or the result of a single major gene (qualitative; reviewed in Kovalchuk et al., 2013). The latter type of resistance may become ineffective over time due to rapidly evolving pathogens. Durable resistance that persists over time and across environments is needed in long-lived forest trees.

Long-term solutions to biotic stress interactions are clearly needed, and an important component is the inclusion of strategies that utilize plant-encoded genetic resistance to a particular stress. Genomic resources for forest trees have supported tree breeding and selection efforts in the form of genome-wide association studies (GWAS) and genomic selection (GS). In addition, genomic resources have improved our ability to infer defense pathways and genes important for producing superior forest trees with resilience against biotic stress. In this minireview, we will highlight recent discoveries in tree systems employing such approaches and propose a way forward for the insightful selection of candidate genes for genetic engineering and improvement of trees.

Breeding for Resistance

Conventional tree breeding, based on phenotype selection, has been a successful means of attaining resistance against pests and pathogens (**Figure 1A**). Sniezko and Koch (2017) review some examples of successful resistance breeding programs around the world. An example from Africa is the hybridization of *Eucalyptus grandis* with *Eucalyptus urophylla* which has improved resistance against the fungal pathogen, *Chrysoporthe austroafricana* (Wingfield, 2003).

Several resistance loci have been identified in forest trees, e.g. *Ppr1* (*Puccinia psidii* resistance 1, Junghans et al., 2003) and Fusiform rust based on mapping using molecular markers. Despite this, the marker-aided selection (MAS) approach in forest systems was limited as most studies reported quantitative trait loci (QTL) discovery but were not validated for the application of MAS itself (reviewed in Sniezko and Koch, 2017)

and has recently been superseded by genomic selection (GS). Grattapaglia et al. (2018) provide a critical review of GS in this special issue. In GS, associations are made between phenotypes and genotypes of a structured training population. A predictive model is made incorporating as many desirable traits as possible. The model is refined based on additional populations and used to predict desired phenotypes based on genotypic information only (Isik, 2014).

In GWAS, associations are made between the phenotype and the genotype in a structured breeding population (Isik, 2014). It is imperative that the associations are validated in a second population. The method relies on a high density of markers and a large discovery population; however, owing to limited genome information, the former is not always possible. In the case of white pine blister rust, a preliminary analysis using a small set of single-nucleotide polymorphism (SNP) markers was undertaken to identify partial resistance to the disease in *Pinus lambertiana* (Vázquez-Lobo et al., 2017). Significant associations were shown for four of the SNP markers. In Norway spruce, exon-capture sequencing was applied genome wide to identify genes associated with susceptibility to *Heterobasidion parviporum* (Mukrimin et al., 2018). Muchero et al. (2018) found *Populus trichocarpa* loci associated with the response to the fungus, *Sphaerulina musiva,* based on genome resequencing. Three loci were associated with resistance and one, encoding a G-type D-mannose–binding receptor-like kinase, was associated with susceptibility. This example demonstrates that candidate defense or susceptibility genes can be identified in a GWAS. In the case of SNP markers, candidate genes within the genomic loci associated with resistance may be inferred based on genome sequence and further information regarding the relevance of the gene during a biotic stress challenge can be determined using functional genomics, e.g. transcriptomics or proteomics. If genes in these associated regions are differentially expressed under the biotic stress challenge, it is possible that they may be contributing to the resistance phenotype.

The approaches can be even more sophisticated using regional heritability mapping (RHM). RHM uncovers variance not accounted for in GWAS as the approach examines short segments of the genome combining the effects of rare and common SNP variants between individuals to estimate the trait variance explained by such regions (Nagamine et al., 2012). Resende et al. (2017) compared RHM and GWAS approaches in *Eucalyptus* for trait mapping of wood and *Puccinia psidii* disease phenotypes. RHM was considered superior revealing more genomic regions that could be interrogated for underlying candidate defense genes. Additionally, the authors suggest that the genomic architecture revealed by the RHM-QTLs could be used to enhance genomic selection models.

An exciting prospect in forest trees would be to follow the approach of Thoen et al. (2017) who performed a GWAS in *Arabidopsis* to 11 different stresses separately and examined combinations of biotic and abiotic stress responses as a phenotype to explore the genetic architecture that underlies plant immunity. Contrasting and overlapping SNP associations were identified in biotic and abiotic stress combinations and it would be interesting
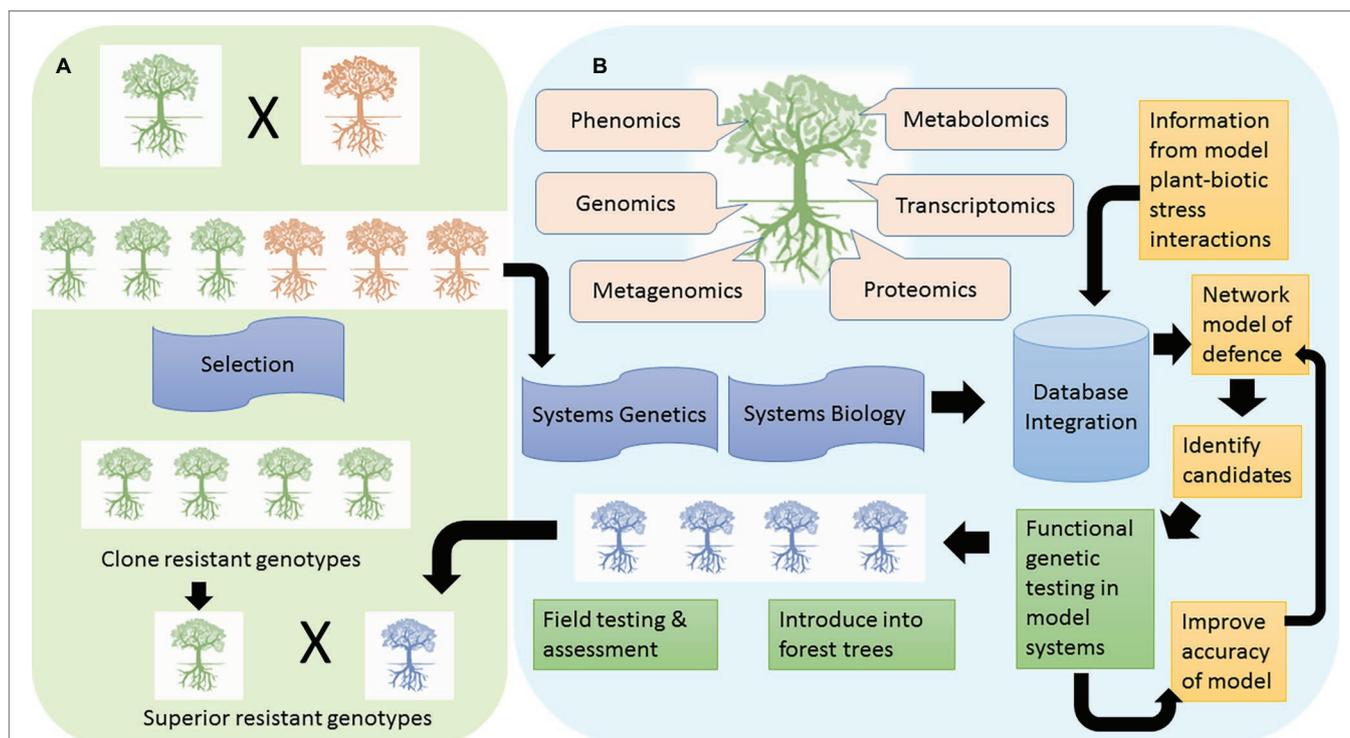
FIGURE 1 | Two avenues to achieving resistance in forest trees are depicted. In tree breeding approaches, **(A)** resistant phenotypes are identified in field and tested under controlled conditions. Such resistant genotypes (green trees) are selected as parents in breeding programs to introduce resistance into a susceptible background (orange trees). Marker-assisted selection has supported plant breeding programs and new genomic tools could enhance selection of resistant trees, e.g. GWAS and GS. In the genetic engineering approach **(B)**, structured breeding programs provide the information about individual resistant or susceptible genotypes that are investigated using a multitude of omic approaches to identify defense mechanisms governing resistance and the power lies in the integration of such studies to a systems level. In addition, summation of such studies into a user-friendly database can assist with comparisons to identify genes and pathways especially important for resistance against specific challenges. A network analysis across species provides support for key genes or pathways to manipulate. The next steps are to test the function of such genes in model systems, and this information is updated in the database. Successful candidates conferring a degree of resistance can be introduced into the forest tree of interest with intense laboratory and field testing (blue trees). The transgenic tree conferring resistance may be bred into various clones selected through breeding programs to pyramid resistance.

to see if similar patterns are apparent in forest trees for which high-density SNP markers are available, e.g. *Eucalyptus*.

Genetic Engineering Tree Resistance

In the absence of natural genetic variation for resistance against a particular biotic stress, genetic engineering may become necessary. The steps toward engineering resistance include gene discovery, candidate gene selection, transformation, and testing (**Figure 1B**). Gene discovery has been aided by omics approaches in forest trees.

Comparative Genomics

A trend observed in the recent genomes of tree species suggests an increase in the number of NLR genes in comparison to herbaceous plants. This may be attributed to their increased exposure to more pathogens during their longer lifetime (Yang et al., 2008; Neale et al., 2017). Specific pathogenesis-related genes were also expanded in different forest tree species such as *Pinus tecnumanii* compared to other sequenced genomes (Visser et al., 2018).

Transcriptomics and Proteomics

Host-pathogen interaction studies have benefitted from the use of transcriptomic approaches and have gained popularity in tree species where comparisons are made between species with disparate tolerance to pathogens (Mangwanda et al., 2015; Meyer et al., 2016). The dual-RNA sequencing approach allows for the intimate study of pathogen and pest as transcripts in both organisms can be captured simultaneously (Naidoo et al., 2017). Host tree responses including defense mechanisms and the corresponding virulence mechanisms in the pathogen can be identified. Some examples are studies on *Notholithocarpus densiflorus* and the oomycete, *Phytophthora ramorum* (Hayden et al., 2014), *Populus* and the fungal pathogen, *Septoria musiva* (Liang et al., 2014), *E. grandis* and *Chrysoporthe austroafricana* as well as *Eucalyptus nitens* and *Phytophthora cinnamomi* (Mangwanda et al., 2015; Meyer et al., 2016). Proteomics offer another level of functional relevance to genes involved in plant defense and iTRAQ (Isobaric Tags for Relative and Absolute Quantitation) was applied in *Eucalyptus* challenged with *C. austroafricana* and *Calonectria pseudoreteaudii* (Chen et al., 2015; Zwart et al., 2017). In the case of the interaction with

C. austroafricana, proteins specific to cell death, salicylic acid signaling and systemic resistance were more apparent compared to the transcriptomic study (Mangwanda et al., 2015). Thus, proteomics may provide novel insight into plant-microbe interactions as there is not a one-to-one relationship between transcript levels and protein abundance.

Metagenomics

Metagenomic studies in forest trees have typically found thousands of taxa associated with individual trees from tropical forests to the boreal region and have linked the importance of these microbiomes to tree health (e.g. Kemler et al., 2013; Kembel and Mueller, 2014; Hacquard and Schadt, 2015; Agler et al., 2016; Jakuschkin et al., 2016; Vivas et al., 2017; Bullington et al., 2018). Above- and below-ground endophytic microbes can contribute to host immunity through multiple direct and indirect (*via* the host) interactions. Indirectly, endophytes can impact pathogens and pests through changes in the host physiology that affect growth or priming of the immune response *via* induced systemic resistance (ISR) and plant volatile production (Hardoim et al., 2015; Brader et al., 2017). Endophytic microbes can also directly impact pathogens and pests directly through competition, exclusion, and antibiosis (Agler et al., 2016; Bamisile et al., 2018). A further mechanism by which endophytic microbes, especially mycorrhizal fungi, may affect plant pathogens and pests is through altered nutritional status of the plant (Gange et al., 2005; Lehr et al., 2007; Pfabel et al., 2012; Kaling et al., 2018). In future, therefore, we need to improve the depth of our understanding on how mutualistic colonization of tree tissues impact pest and pathogen interactions by combining a range of omic approaches to understand the molecular mechanisms at play.

Advances in our knowledge of mechanisms by which beneficial microbes play a role in plant health as well as in our understanding of holistic plant systems are opening up new avenues whereby we can begin to consider actively selecting tree genotypes that specifically foster microbiomes to improve plant health (Mueller and Sachs, 2015). Past research has demonstrated that it is possible to control various aspects of the root and/or foliar microbiome through genetic manipulation of different metabolic pathways within the tree through single-gene alterations (Beckers et al., 2016; Veach et al., 2018). While these examples are a proof of concept that minute genomic changes to a tree's genome can have significant impacts on microbial populations, the question now remains as to whether we can intentionally breed or engineer tree genotypes that specifically target and foster/repress the growth of specific microbes within plant tissues. While much research is needed within this area for tree species of economic importance, advances in this area have been made in agricultural crops (Besserer et al., 2006; Fierer et al., 2007; Carvalhais et al., 2015). Should similar pathways be targeted in tree systems we might be able to similarly improve tree health through naturally present microbial populations.

Metabolomics

Metabolomics in forest tree-biotic stress interactions have been limited to the study of specific secondary metabolites, e.g. terpenes (Oates et al., 2015; Naidoo et al., 2018), using methods such gas or liquid chromatography coupled to mass spectrometry (GC-MS and LC-MS, respectively) and near-infrared reflectance (NIR) technologies. Metabolomics in European common ash, *Fraxinus excelsior*, revealed different profiles associated with resistance and tolerance to the fungus, *Hymenoscyphus fraxineus* (Sambles et al., 2017). Where possible, combining different approaches, e.g. transcriptomics and metabolomics, enhances the selection of pathways implicated in defense. This was applied in pedunculate oak (Kersten et al., 2013). Oaks resistant and susceptible to the leaf roll miner herbivore had specific volatile profiles and susceptible trees produced a volatile blend more attractive for the pest. At the transcriptome level, resistance was attributed to constitutive rather than induced defenses. Plant-plant interactions play an important role in defense signaling against pathogens (for a recent review, please see Subrahmaniam et al., 2018 who describe candidate genes defining plant-plant interactions as cell wall modifications and defense pathways). This phenomenon has been studied in trees to a limited extent as indirect defenses. Indirect defenses involve the release of volatile compounds that attract natural enemies of the pest or signal neighboring plants of the threat (Unsicker et al., 2009). This has been demonstrated in elm (*Ulmus minor*) responses to the elm leaf beetle, *Xanthogaleruca luteola,* where oviposition altered volatile profiles that attracted the specialist egg parasitoid, *Oomyzus gallerucae* (Meiners and Hilker, 2000). *Eucalyptus* also uses these signals to prime neighboring plants, e.g. feeding by *Ctenarytaina eucalypti* induced volatile responses which increased defense-related compounds in unwounded plants (Troncoso et al., 2012). Oviposition or herbivory-induced chemical cues released by plants have been shown to serve as reliable indicators of the presence of prey to predators and parasitoids of the pest (Hilker et al., 2002; Bruinsma et al., 2009).

While these previous applications have proceeded at a rapid rate in forestry research, an area that still lags behind is phenomics. Phenomics involves the high-throughput capture of the phenotypes of an organism, influenced by its genotype and genotype × environment. Phenomics informs the genotype-phenotype map (reviewed in Houle et al., 2010) and new imaging technologies are being pursued to achieve this goal. Ludovisi et al. (2017) describe the use of unmanned aerial vehicle type of remote sensing and imaging to enable high-throughput field phenotyping. The authors applied this to black poplar populations to capture precise phenotypes for drought responses. These techniques are equally promising for biotic stress phenotyping to accelerate the development of superior genotypes.

With the generation of multiple levels of omics data, systems biology views of organisms become possible. This provides a framework for linking molecular interactions with complex traits (Mizrachi et al., 2017). The integration of information from different experiments can be used to generate a network modeling the dynamics and complexity of a biological system. The development of such models is iterative and is used to refine new models. Studies in *Arabidopsis* show the power of this approach to uncover key mechanisms important in plant defense (Mukhtar et al., 2011). A related approach is systems genetics (also known as genetical genomics) which incorporates the genetic variation of organisms with the

systems level phenotyping to dissect complex traits (Jansen and Nap, 2001). Tree breeding programs provide the structured populations for such types of studies (**Figure 1B**). Molecular mechanisms underlying wood properties in *Eucalyptus* were characterized in this manner (Mizrachi et al., 2017), with systems genetics studies involving tree biotic stress interactions not far behind.

SELECTION OF CANDIDATE DEFENSE GENES

The wealth of omics data that is being generated for trees during pest and pathogen challenge requires concomitant development of resources to promote access by the scientific community. User-friendly databases are increasingly becoming available to analyze and visualize omics data in woody perennials, which can assist with identifying key genes and pathways for resistance. The TreeGenes database (Wegrzyn et al., 2008[1]) is a forest tree genomics resource for integration and analysis of genome sequences, transcriptomes, genetic maps, molecular markers, and phenotypic data for 1,749 species. The Plant Genome Integrative Explorer (Sundell et al., 2015[2]) provides access to genomic and transcriptomic data for *Poplar*, *Arabidopsis*, *Eucalyptus*, and conifer. PlantGenIE contains tools for expression analyses and visualization of gene co-expression networks within and across forest species (Netotea et al., 2014). Gramene (Tello-Ruiz et al., 2016[3]) contains the plant reactome database for analysis of plant metabolic and regulatory pathways. MorphDB (Zwaenepoel et al., 2018[4]) assists in identifying missing genes in pathways and regulators such as transcription factors or other signaling genes. Such approaches can place the potential resistance genes in a functional context relative to known resistance genes and thus assist in identifying and prioritizing candidates (Rhee and Mutwil, 2014). Flowing out of such databases should be the development of a predictive network model to understand the potential functional role of genes within a pathway (**Figure 1B**). Mewalal et al. (2014) provides an insightful review on identifying and prioritizing genes using omic approaches to assist new hypothesis-driven experiments. Once candidate genes are prioritized and selected, they are tested in model systems for a disease or pest tolerant phenotype. This has been demonstrated for *P. trichocarpa* whereby overexpression of a salicylic acid-inducible gene, *PtrWRKY73*, in *Arabidopsis* increased resistance to *Pseudomonas syringae* (biotroph) but reduced resistance to *Botrytis cinerea* (necrotroph) (Duan et al., 2015). Another study made use of *Nicotiana tabacum* overexpressing antimicrobial protein Sp-AMP2 (PR-19) from *Pinus sylvestris* L., which enhanced resistance to *B. cinerea* (Jaber et al., 2017). Following successful or unsuccessful demonstration of a candidate gene's function, the model can be refined followed by repeated testing until

a desired phenotype is observed (**Figure 1B**). Then introduction of the gene into the desired forest tree species is pursued *via* genetic transformation. Candidate genes may be overexpressed, knocked down, or knocked out. Jiang et al. (2017) utilized CRISPR/Cas9 knockouts of candidate *WRKY* transcription factors to unravel regulatory mechanisms in *Populus* during interaction with *Melampsora* rust. Once successful transgenic lines are obtained, extensive functional testing under different conditions is made to determine if the candidate gene confers biotic stress resistance. One limitation of producing transgenic trees is that transformation may be optimized for specific clonal backgrounds. This can be circumvented by incorporating transgenic trees as parents in breeding programs to transfer the desired resistant phenotype to other genetic backgrounds (**Figure 1**).

ADVANCES TOWARD ENGINEERING RESISTANCE

Forest biotechnology companies have been developing transgenic forest tree species, *Populus* and *Eucalyptus*, to enhance biomass and resistance to stresses. ArborGen Inc. has developed transgenic freeze-tolerant *Eucalyptus* (Zhang et al., 2012) and SweTree Technologies have developed transgenic hybrid aspen (*Populus tremula* × *Populus tremuloides*) with enhanced growth properties (Eriksson et al., 2006). FuturaGene Ltd. has developed and commercialized the first genetically modified *Eucalyptus* tree with 20% more biomass ("Brazil approves transgenic eucalyptus," 2015) with developments to commercialize disease-resistant trees (Avisar et al., 2013) underway. Additionally, genome editing holds great promise for accelerated breeding in forestry (Bewg et al., 2018). Therefore, once we identify genes or metabolic pathways crucial to disease resistance, we will be able to quickly manipulate these pathways and produce plants that are available to industry in a timelier manner than ever before.

CONCLUSION

With the unprecedented increase in forest pests and pathogens, the two avenues to generate resistant trees are necessary. Breeding and genetic engineering could be combined to accelerate achieving the goal of resistance in forestry. Genomic tools are enhancing our rate of gene discovery in forest trees; however, the vast amounts of data must be more centralized and accessible to identify candidate genes for testing. As is the trend in model systems, tree immunity should also be increasingly viewed from a holistic, systems-based perspective that incorporates the complexity of the biotic interactions outside, within and below the tree. This includes the potential for direct manipulation of host-associated microbes for tree resistance, as well as the recognition that genetic engineering of the host impacts the microbiome with potentially important consequences for tree resistance. A concerted effort has to be made to make engineering resistance a reality in forestry.

---

[1]https://treegenesdb.org/Drupal/

[2]http://plantgenie.org/

[3]http://www.gramene.org/

[4]http://bioinformatics.psb.ugent.be/webtools/morphdb/morphDB/index/

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Agler, M. T., Ruhe, J., Kroll, S., Morhenn, C., Kim, S. T., Weigel, D., et al. (2016). Microbial hub taxa link host and abiotic factors to plant microbiome variation. *PLoS Biol.* 14, 1–31. doi: 10.1371/journal.pbio.1002352

Avisar, D., Stein, H., and Shani, Z. (2013). *Pest-resistant plants containing a combination of a spider toxin and a chitinase.* Google Patents.

Bamisile, B. S., Dash, C. K., Akutse, K. S., Keppanan, R., and Wang, L. (2018). Fungal endophytes: beyond herbivore management. *Front. Microbiol.* 9, 1–11. doi: 10.3389/fmicb.2018.00544

Beckers, B., Op De Beeck, M., Weyens, N., Van Acker, R., Van Montagu, M., Boerjan, W., et al. (2016). Lignin engineering in field-grown poplar trees affects the endosphere bacterial microbiome. *Proc. Natl. Acad. Sci.* 113, 2312–2317. doi: 10.1073/pnas.1523264113

Besserer, A., Puech-Pagès, V., Kiefer, P., Gomez-Roldan, V., Jauneau, A., Roy, S., et al. (2006). Strigolactones stimulate arbuscular mycorrhizal fungi by activating mitochondria. *PLoS Biol.* 4, 1239–1247. doi: 10.1371/journal.pbio.0040226

Bewg, W. P., Ci, D., and Tsai, C.-J. (2018). Genome editing in trees: from multiple repair pathways to long-term stability. *Front. Plant Sci.* 9:1732. doi: 10.3389/fpls.2018.01732

Brader, G., Compant, S., Vescio, K., Mitter, B., Trognitz, F., Ma, L.-J., et al. (2017). Ecology and genomic insights into plant-pathogenic and plant-nonpathogenic endophytes. *Annu. Rev. Phytopathol.* 55, 61–83. doi: 10.1146/annurev-phyto-080516-035641

Bruinsma, M., Posthumus, M. A., Mumm, R., Mueller, M. J., Van Loon, J. J. A., and Dicke, M. (2009). Jasmonic acid-induced volatiles of *Brassica oleracea* attract parasitoids: Effects of time and dose, and comparison with induction by herbivores. *J. Exp. Bot.* 60, 2575–2587. doi: 10.1093/jxb/erp101

Bullington, L. S., Lekberg, Y., Sniezko, R., and Larkin, B. (2018). The influence of genetics, defensive chemistry, and the fungal microbiome on disease outcome in whitebark pine trees. *Mol. Plant Pathol.* 19, 1847–1858. doi: 10.1111/mpp.12663

Carvalhais, L. C., Dennis, P. G., Badri, D. V., Kidd, B. N., Vivanco, J. M., and Schenk, P. M. (2015). Linking jasmonic acid signalling, root exudates and rhizosphere microbiomes. *Mol. Plant-Microbe Interact.* 28, 1049–1058. doi: 10.1094/MPMI-01-15-0016-R

Chen, Q., Guo, W., Feng, L., Ye, X., Xie, W., Huang, X., et al. (2015). Data for transcriptome and proteome analysis of Eucalyptus infected with *Calonectria pseudoreteaudii. Data Br.* 3, 24–28. doi: 10.1016/j.dib.2014.12.008

Das, T., Majumdar, M. H. D., Devi, R. K. T., and Rajesh, T. (2016). Climate change impact on plant diseases. *SAARC J. Agric.* 14, 200–209. doi: 10.3329/sja.v14i2.31259

Dittrich-Schröder, G., Wingfield, M. J., Hurley, B. P., and Slippers, B. (2012). Diversity in *Eucalyptus* susceptibility to the gall-forming wasp *Leptocybe invasa. Agric. For. Entomol.* 14, 419–427. doi: 10.1111/j.1461-9563.2012.00583.x

Duan, Y., Jiang, Y., Ye, S., Karim, A., Ling, Z., He, Y., et al. (2015). PtrWRKY73, a salicylic acid-inducible poplar WRKY transcription factor, is involved in disease resistance in *Arabidopsis thaliana. Plant Cell Rep.* 34, 831–841. doi: 10.1007/s00299-015-1745-5

Eriksson, M., Moritz, T., Israelsson, M., and Olsson, O. (2006). *Transgenic trees exhibiting increased growth, biomass production and xylem fibre length, and methods for their production.* Google Patents.

Fierer, N., Bradford, M. A., and Jackson, R. B. (2007). Toward an ecological classification of soil bacteria. *Ecology* 88, 1354–1364. doi: 10.1007/s00209-018-2115-0

Gange, A. C., Gane, D. R. J., Chen, Y., and Gong, M. (2005). Dual colonization of *Eucalyptus urophylla* S.T. Blake by arbuscular and ectomycorrhizal fungi affects levels of insect herbivore attack. *Agric. For. Entomol.* 7, 253–263. doi: 10.1111/j.1461-9555.2005.00268.x

Gordon, T. R., and Reynolds, G. J. (2017). Plasticity in plant-microbe interactions: a perspective based on the pitch canker pathosystem. *Phytoparasitica* 45, 1–8. doi: 10.1007/s12600-016-0558-6

Granados, G. M., McTaggart, A. R., Barnes, I., Rodas, C. A., Roux, J., and Wingfield, M. J. (2017). The pandemic biotype of *Austropuccinia psidii* discovered in South America. *Australas. Plant Pathol.* 46, 267–275. doi: 10.1007/s13313-017-0488-x

Grattaglia, D., Silva-Junior, O. B., Resende, R. T., Cappa, E. P., Müller, B. S. F., Tan, B., et al. (2018). Quantitative Genetics and Genomics Converge to Accelerate Forest Tree Breeding. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.01693

Hacquard, S., and Schadt, C. W. (2015). Towards a holistic understanding of the beneficial interactions across the Populus microbiome. *New Phytol.* 205, 1424–1430. doi: 10.1111/nph.13133

Hardoim, P. R., van Overbeek, L. S., Berg, G., Pirttilä, A. M., Compant, S., Campisano, A., et al. (2015). The hidden world within plants: ecological and evolutionary considerations for defining functioning of microbial endophytes. *Microbiol. Mol. Biol. Rev.* 79, 293–320. doi: 10.1128/MMBR.00050-14

Hayden, K. J., Garbelotto, M., Knaus, B. J., Cronn, R. C., Rai, H., and Wright, J. W. (2014). Dual RNA-seq of the plant pathogen *Phytophthora ramorum* and its tanoak host. *Tree Genet. Genomes* 10, 489–502. doi: 10.1007/s11295-014-0698-0

Hilker, M., Kobs, C., Varama, M., and Schrank, K. (2002). Insect egg deposition induces *Pinus sylvestris* to attract egg parasitoids. *J. Exp. Biol.* 205, 455–461. doi: 10.1242/jeb.00732

Houle, D., Govindaraju, D. R., and Omholt, S. (2010). Phenomics: the next challenge. *Nat. Rev. Genet.* 11, 855. doi: 10.1038/nrg2897

Hurley, B. P., Garnas, J., Wingfield, M. J., Branco, M., Richardson, D. M., and Slippers, B. (2016). Increasing numbers and intercontinental spread of invasive insects on eucalypts. *Biol. Invasions* 18, 921–933. doi: 10.1007/s10530-016-1081-x

Isik, F. (2014). Genomic selection in forest tree breeding: the concept and an outlook to the future. *New For.* 45, 379–401. doi: 10.1007/s11056-014-9422-z

Jaber, E., Kovalchuk, A., Raffaello, T., Keriö, S., Teeri, T., and Asiegbu, F. O. (2017). A gene encoding scots pine antimicrobial protein Sp-AMP2 (PR-19) confers increased tolerance against *Botrytis cinerea* in transgenic tobacco. *Forests* 9, 1–15. doi: 10.3390/f9010010

Jakuschkin, B., Fievet, V., Schwaller, L., Fort, T., Robin, C., and Vacher, C. (2016). Deciphering the pathobiome: intra- and interkingdom interactions involving the pathogen *Erysiphe alphitoides. Microb. Ecol.* 72, 870–880. doi: 10.1007/s00248-016-0777-x

Jansen, R. C., and Nap, J. P. (2001). Genetical genomics: the added value from segregation. *Trends Genet.* 17, 388–391. doi: 10.1016/S0168-9525(01)02310-1

Jiang, Y., Guo, L., Ma, X., Zhao, X., Jiao, B., Li, C., et al. (2017). The WRKY transcription factors PtrWRKY18 and PtrWRKY35 promote *Melampsora* resistance in *Populus. Tree Physiol.* 37, 665–675. doi: 10.1093/treephys/tpx008

Jones, J. D. G., and Dangl, J. L. (2006). The plant immune system. *Nature* 444, 323–329. doi: 10.1038/nature05286

Junghans, D. T., Alfenas, A. C., Brommonschenkel, S. H., Oda, S., Mello, E. J., and Grattapaglia, D. (2003). Resistance to rust (*Puccinia psidii* Winter) in *Eucalyptus*: mode of inheritance and mapping of a major gene with RAPD markers. *Theor. Appl. Genet.* 108, 175–180. doi: 10.1007/s00122-003-1415-9

Kaling, M., Schmidt, A., Moritz, F., Rosenkranz, M., Witting, M., Kasper, K., et al. (2018). Mycorrhiza-triggered transcriptomic and metabolomic networks impinge on herbivore fitness. *Plant Physiol.* 176, 2639–2656. doi: 10.1104/pp.17.01810

Kembel, S. W., and Mueller, R. C. (2014). Plant traits and taxonomy drive host associations in tropical phyllosphere fungal communities. *Botany* 92, 303–311. doi: 10.1139/cjb-2013-0194

Kemler, M., Garnas, J., Wingfield, M. J., Gryzenhout, M., Pillay, K. A., and Slippers, B. (2013). Ion torrent PGM as tool for fungal community analysis: a case study of endophytes in *Eucalyptus grandis* reveals high taxonomic diversity. *PLoS One* 7:e81718. doi: 10.1371/journal.pone.0081718

Kersten, B., Ghirardo, A., Schnitzler, J. P., Kanawati, B., Schmitt-Kopplin, P., Fladung, M., et al. (2013). Integrated transcriptomics and metabolomics decipher differences in the resistance of pedunculate oak to the herbivore *Tortrix viridana* L. *BMC Genomics* 14:737. doi:10.1186/1471-2164-14-737

Kovalchuk, A., Kerio, S., Oghenekaro, A. O., Jaber, E., Raffaello, T., and Asiegbu, F. O. (2013). Antimicrobial defences and resistance in forest trees: challenges and perspectives in a genomics era. *Annu. Rev. Phytopathol.* 51, 221–244. doi: 10.1146/annurev-phyto-082712-102307

Lehr, N., Schrey, S. D., and Tarkka, M. T. (2007). Root inoculation with a forest soil streptomycete leads to locally and systemically increased resistance against phytopathogens in Norway spruce. *New Phytol.* 177, 965–976. doi: 10.1111/j.1469-8137.2007.02322.x

Liang, H., Staton, M., Xu, Y., Xu, T., and Leboldus, J. (2014). Comparative expression analysis of resistant and susceptible *Populus* clones inoculated with *Septoria musiva*. *Plant Sci.* 223, 69–78. doi: 10.1016/j.plantsci.2014.03.004

Ludovisi, R., Tauro, F., Salvati, R., Khoury, S., Mugnozza Scarascia, G., and Harfouche, A. (2017). UAV-based thermal imaging for high-throughput field phenotyping of black poplar response to drought. *Front. Plant Sci.* 8, 1–18. doi: 10.1109/AMS.2009.139

Mangwanda, R., Myburg, A. A., and Naidoo, S. (2015). Transcriptome and hormone profiling reveals *Eucalyptus grandis* defence responses against *Chrysoporthe austroafricana*. *BMC Genomics* 16, 1–13. doi: 10.1186/s12864-015-1529-x

McTaggart, A. R., Shuey, L. S., Granados, G. M., du Plessis, E., Fraser, S., Barnes, I., et al. (2017). Evidence that *Austropuccinia psidii* may complete its sexual life cycle on Myrtaceae. *Plant Pathol.* 67, 729–734. doi: 10.1111/ppa.12763

Meiners, T., and Hilker, M. (2000). Induction of plant synomones by oviposition of a phytophagus insect. *J. Chem. Ecol.* 26, 221–232. doi: 10.1023/A:1005453830961

Mewalal, R., Mizrachi, E., Mansfield, S. D., and Myburg, A. A. (2014). Cell wall-related proteins of unknown function: missing links in plant cell wall development. *Plant Cell Physiol.* 55, 1031–1043. doi: 10.1093/pcp/pcu050

Meyer, F. E., Shuey, L. S., Naidoo, S., Mamni, T., Berger, D. K., Myburg, A. A., et al. (2016). Dual RNA-sequencing of *Eucalyptus nitens* during *Phytophthora cinnamomi* challenge reveals pathogen and host factors influencing compatibility. *Front. Plant Sci.* 7, 1–15. doi: 10.3389/fpls.2016.00191

Mizrachi, E., Verbeke, L., Christie, N., Fierro, A. C., Mansfield, S. D., Davis, M. F., et al. (2017). Network-based integration of systems genetics data reveals pathways associated with lignocellulosic biomass accumulation and processing. *Proc. Natl. Acad. Sci.* 114, 1195–1200. doi: 10.1073/pnas.1620119114

Muchero, W., Sondreli, K. L., Chen, J.-G., Urbanowicz, B. R., Zhang, J., Singan, V., et al. (2018). Association mapping, transcriptomics, and transient expression identify candidate genes mediating plant–pathogen interactions in a tree. *Proc. Natl. Acad. Sci.* 115, 11573-11578. doi: 10.1073/pnas.1804428115

Mueller, U. G., and Sachs, J. L. (2015). Engineering microbiomes to improve plant and animal health. *Trends Microbiol.* 23, 606–617. doi: 10.1016/j.tim.2015.07.009

Mukhtar, M. S., Carvunis, A.-R., Dreze, M., Epple, P., Steinbrenner, J., Moore, J., et al. (2011). Independently evolved virulence effectors converge onto hubs in a plant immune system network. *Science* 333, 596–601. doi: 10.1126/science.1203659

Mukrimin, M., Kovalchuk, A., Neves, L. G., Jaber, E. H. A., Haapanen, M., Kirst, M., et al. (2018). Genome-wide exon-capture approach identifies genetic variants of norway spruce genes associated with susceptibility to *heterobasidion parviporum* infection. *Front. Plant Sci.* 9, 1–13. doi: 10.3389/fpls.2018.00793

Nagamine, Y., Pong-Wong, R., Navarro, P., Vitart, V., Hayward, C., Rudan, I., et al. (2012). Localising loci underlying complex trait variation using regional genomic relationship mapping. *PLoS One* 7:e46501. doi: 10.1371/journal.pone.0046501

Naidoo, S., Christie, N., Acosta, J. J., Mphahlele, M. M., Payn, K. G., Myburg, A. A., et al. (2018). Terpenes associated with resistance against the gall wasp, Leptocybe invasa, in *Eucalyptus grandis*. *Plant Cell Environ.* 41, 1840–1851. doi: 10.1111/pce.13323

Naidoo, S., Külheim, C., Zwart, L., Mangwanda, R., Oates, C. N., Visser, E. A., et al. (2014). Uncovering the defence responses of eucalyptus to pests and pathogens in the genomics age. *Tree Physiol.* 34, 931–943. doi: 10.1093/treephys/tpu075

Naidoo, S., Visser, E. A., Zwart, L., Toit, Y. d., Bhadauria, V., and Shuey, L. S. (2017). Dual RNA-sequencing to elucidate the plant-pathogen duel. *Curr. Issues Mol. Biol.* 27, 127–141. doi: 10.21775/cimb.027.127

Neale, D. B., Martínez-García, P. J., De La Torre, A. R., Montanari, S., and Wei, X.-X. (2017). Novel insights into tree biology and genome evolution as revealed through genomics. *Annu. Rev. Plant Biol.* 68, 457–483. doi: 10.1146/annurev-arplant-042916-041049

Netotea, S., Sundell, D., Street, N. R., and Hvidsten, T. R. (2014). ComPlEx: conservation and divergence of co-expression networks in *A. thaliana, Populus* and *O. sativa*. *BMC Genomics* 15, 1–17. doi: 10.1186/1471-2164-15-106

Newhouse, A. E., Polin-McGuigan, L. D., Baier, K. A., Valletta, K. E. R., Rottmann, W. H., Tschaplinski, T. J., et al. (2014). Transgenic american chestnuts show enhanced blight resistance and transmit the trait to T1 progeny. *Plant Sci.* 228, 88–97. doi: 10.1016/j.plantsci.2014.04.004

Oates, C. N., Külheim, C., Myburg, A. A., Slippers, B., and Naidoo, S. (2015). The transcriptome and terpene profile of *Eucalyptus grandis* reveals mechanisms of defense against the insect pest, *Leptocybe invasa*. *Plant Cell Physiol.* 56, 1418–1428. doi: 10.1093/pcp/pcv064

Pfabel, C., Eckhardt, K. U., Baum, C., Struck, C., Frey, P., and Weih, M. (2012). Impact of ectomycorrhizal colonization and rust infection on the secondary metabolism of poplar (*Populus trichocarpa × deltoides*). *Tree Physiol.* 32, 1357–1364. doi: 10.1093/treephys/tps093

Resende, R. T., Resende, M. D. V., Silva, F. F., Azevedo, C. F., Takahashi, E. K., Silva-Junior, O. B., et al. (2017). Regional heritability mapping and genome-wide association identify loci for complex growth, wood and disease resistance traits in *Eucalyptus*. *New Phytol.* 213, 1287–1300. doi: 10.1111/nph.14266

Rhee, S. Y., and Mutwil, M. (2014). Towards revealing the functions of all genes in plants. *Trends Plant Sci.* 19, 212–221. doi: 10.1016/j.tplants.2013.10.006

Roux, J., Greyling, I., Coutinho, T. A., Verleur, M., and Wingfield, M. J. (2013). The Myrtle rust pathogen, *Puccinia psidii*, discovered in Africa. *IMA Fungus* 4, 155–159. doi: 10.5598/imafungus.2013.04.01.14

Sambles, C. M., Salmon, D. L., Florance, H., Howard, T. P., Smirnoff, N., Nielsen, L. R., et al. (2017). Data descriptor: ash leaf metabolomes reveal differences between trees tolerant and susceptible to ash dieback disease. *Sci. Data* 4, 1–13. doi: 10.1038/sdata.2017.190

Santini, A., Ghelardini, L., De Pace, C., Desprez-Loustau, M. L., Capretti, P., Chandelier, A., et al. (2013). Biogeographical patterns and determinants of invasion by forest pathogens in Europe. *New Phytol.* 197, 238–250. doi: 10.1111/j.1469-8137.2012.04364.x

Six, D. L., Vergobbi, C., and Cutter, M. (2018). Are survivors different? Genetic-based selection of trees by mountain pine beetle during a climate change-driven outbreak in a high-elevation pine forest. *Front. Plant Sci.* 9, 1–11. doi: 10.3389/fpls.2018.00993

Sniezko, R. A., and Koch, J. (2017). Breeding trees resistant to insects and diseases: putting theory into application. *Biol. Invasions* 19, 3377–3400. doi: 10.1007/s10530-017-1482-5

Sturrock, R. N., Frankel, S. J., Brown, A. V., Hennon, P. E., Kliejunas, J. T., Lewis, K. J., et al. (2011). Climate change and forest diseases. *Plant Pathol.* 60, 133–149. doi: 10.1111/j.1365-3059.2010.02406.x

Subrahmaniam, H. J., Libourel, C., Journet, E. P., Morel, J. B., Muños, S., Niebel, A., et al. (2018). The genetics underlying natural variation of plant–plant interactions, a beloved but forgotten member of the family of biotic interactions. *Plant J.* 93, 747–770. doi: 10.1111/tpj.13799

Sundell, D., Mannapperuma, C., Netotea, S., Delhomme, N., Lin, Y. C., Sjödin, A., et al. (2015). The plant genome integrative explorer resource: PlantGenIE.org. *New Phytol.* 208, 1149–1156. doi: 10.1111/nph.13557

Tello-Ruiz, M. K., Stein, J., Wei, S., Preece, J., Olson, A., Naithani, S., et al. (2016). Gramene 2016: comparative plant genomics and pathway resources. *Nucleic Acids Res.* 44, D1133–D1140. doi: 10.1093/nar/gkv1179

Thoen, M. P. M., Davila Olivas, N. H., Kloth, K. J., Coolen, S., Huang, P. P., Aarts, M. G. M., et al. (2017). Genetic architecture of plant stress resistance: multi-trait genome-wide association mapping. *New Phytol.* 213, 1346–1362. doi: 10.1111/nph.14220

Troncoso, C., Becerra, J., Perez, C., Hernandez, V., Martin, A., Sanchez-olate, M., et al. (2012). Induction of defensive responses in *Eucalyptus globulus* (Labill) plants, against *Ctenarytaina eucalypti* (Maskell) (Hemiptera: Psyllidae). *Am. J. Plant Sci.* 2012, 589–595. doi: 10.4236/ajps.2012.35071

Unsicker, S. B., Kunert, G., and Gershenzon, J. (2009). Protective perfumes: the role of vegetative volatiles in plant defense against herbivores. *Curr. Opin. Plant Biol.* 12, 479–485. doi: 10.1016/j.pbi.2009.04.001

Vázquez-Lobo, A., De La Torre, A. R., Martínez-García, P. J., Vangestel, C., Wegzryn, J. L., Ćalić, I., et al. (2017). Finding loci associated to partial resistance to white pine blister rust in sugar pine (*Pinus lambertiana* Dougl.). *Tree Genet. Genomes* 13:108. doi: 10.1007/s11295-017-1190-4

Veach, A. M., Yip, D., Engle, N. L., Yang, Z. K., Bible, A., Morrell-Falvey, J., et al. (2018). Modification of plant cell wall chemistry impacts metabolome and microbiome composition in *Populus PdKOR1* RNAi plants. *Plant Soil* 429, 349–361. doi: 10.1007/s11104-018-3692-8

Visser, E. A., Wegzryn, J. L., Myburg, A. A., and Naidoo, S. (2018). Defence transcriptome assembly and pathogenesis related gene family analysis in *Pinus tecunumanii* (low elevation). *BMC Genomics* 19, 1–13. doi: 10.1186/s12864-018-5015-0

Vivas, M., Kemler, M., Mphahlele, M. M., Wingfield, M. J., and Slippers, B. (2017). Maternal effects on phenotype, resistance and the structuring of fungal communities in *Eucalyptus grandis*. *Environ. Exp. Bot.* 140, 120–127. doi: 10.1016/j.envexpbot.2017.06.002

Wegzryn, J. L., Lee, J. M., Tearse, B. R., and Neale, D. B. (2008). TreeGenes: a forest tree genome database. *Int. J. Plant Genomics* 2008, 1–7. doi: 10.1155/2008/412875

Wingfield, M. J. (2003). Daniel McAlpine Memorial Lecture: increasing threat of diseases to plantation forests in the Southern Hemisphere: lessons from *Cryphonectria* canker. *Australas. Plant Pathol.* 32, 133–139. doi: 10.1071/AP03024

Wingfield, M. J., Brockerhoff, E. G., Wingfield, B. D., and Slippers, B. (2015). Planted forest health: the need for a global strategy. *Science* 349, 832–836. doi: 10.1126/science.aac6674

Yang, S., Zhang, X., Yue, J. X., Tian, D., and Chen, J. Q. (2008). Recent duplications dominate NBS-encoding gene expansion in two woody species. *Mol. Gen. Genomics.* 280, 187–198. doi: 10.1007/s00438-008-0355-0

Zhang, C., Kwan, B., Chang, S., and Raymond, P. (2012). *Freeze tolerant eucalyptus*. Google Patents.

Zwaenepoel, A., Diels, T., Amar, D., Van Parys, T., Shamir, R., Van de Peer, Y., et al. (2018). MorphDB: prioritizing genes for specialized metabolism pathways and gene ontology categories in plants. *Front. Plant Sci.* 9, 1–13. doi: 10.3389/fpls.2018.00352

Zwart, L., Berger, D. K., Moleleki, L. N., Van Der Merwe, N. A., Myburg, A. A., and Naidoo, S. (2017). Evidence for salicylic acid signalling and histological changes in the defence response of *Eucalyptus grandis* to *Chrysoporthe austroafricana*. *Sci. Rep.* 7, 1–12. doi: 10.1038/srep45402

Cold Hardiness in Trees: A Mini-Review

Michael Wisniewski[1]*, Annette Nassuth[2]* and Rajeev Arora[3]*

[1] United States Department of Agriculture – Agricultural Research Service, Kearneysville, WV, United States, [2] Department of Molecular and Cellular Biology, University of Guelph, Ontario, ON, Canada, [3] Department of Horticulture, Iowa State University, Ames, IA, United States

Significant advances have been made in our understanding of the regulation of cold hardiness. The existence of numerous biophysical and biochemical adaptive mechanisms in perennial woody plants and the complexity their regulation has made the development of methods for managing and improving cold hardiness in perennial woody plants has been very difficult. This may be partially attributed to viewing cold hardiness as a single dimensional response, rather than as a complex phenomenon, involving different mechanisms (avoidance and tolerance), different stages (mid-winter vs. late winter), and having an intimate overlap with the genetic regulation of dormancy. In particular separating the molecular regulation of cold hardiness from growth processes has been challenging. ICE and C-repeat binding factor (CBF), transcription factors (Inducer of CBF expression and CRT-binding factor) have been shown to be an important aspect in the regulation of cold-induced gene expression. Evidence has emerged, however, that they are also intimately involved in the regulation of growth, flowering, dormancy, and stomatal development. This evidence includes the presence of CBF binding motifs in genes regulating these processes, or through cross-talk between the pathways that regulate them. Recent changes in climate that have resulted in erratic episodes of unseasonal warming followed by more seasonal patterns of low temperatures has also highlighted the need to better understand the genetic and molecular regulation of deacclimation, a topic of research that is only more recently being addressed. Environmentally-induced epigenetic regulation of stress responses and seasonal processes such as cold acclimation, deacclimation, and dormancy have been documented but are still poorly understood. Advances in the ability to efficiently generate large DNA and RNA datasets and genetic transformation technologies have greatly increased our ability to explore the regulation of gene expression and explore genetic diversity. Greater knowledge of the interplay between epigenetic and genetic regulation of cold hardiness, along with the application of advanced genetic analyses, such as genome-wide-association-studies (GWAS), are needed to develop strategies for addressing the complex processes associated with cold hardiness in woody plants. A cautionary note is also indicated regarding the time-scale needed to examine and interpret plant response to freezing temperatures if progress is to be made in developing effective approaches for manipulating and improving cold hardiness.

Keywords: freezing tolerance, ice nucleation, cold acclimation, deacclimation, dormancy, C-repeat binding factor (CBF), *DAM* genes, antifreeze protein (AFP)

INTRODUCTION

Ever since the first microscopic observations of the freezing response of cells were made in the latter part of the 19th century and early 20th century (Molisch, 1897; Wiegand, 1906), and it was discovered that plant cells undergo cytorrhysis rather than plasmolysis in response to freezing, an elusive search has been conducted to develop a complete and integrated understanding of cold hardiness and freezing tolerance in plants (Wisniewski et al., 2003; Gusta and Wisniewski, 2013; Arora, 2018). Despite thousands of reports and countless reviews, reliable approaches to improving freezing tolerance, without affecting other aspects of plant development, have yet to be developed, either at the molecular/genetic level or the physiological level. New technologies have allowed us to understand plant response to low temperatures in greater and greater detail, but the picture has greatly increased in complexity.

The lack of progress may be partially attributable to two factors. One factor is interpreting cold hardiness as a singular on/off response rather than a combination of many diverse mechanisms that involve significant structural, biochemical, and genetic adjustments, as well as the complexity of manipulating cold hardiness without having a negative impact on other plant developmental processes. The characteristics of these components are species-specific (often genotype-specific), potentially under separate genetic control. Therefore, it is essential when investigating plant cold hardiness to be cognizant of what aspect of the process is being studied and its potential impact on the aspect of cold hardiness that is deemed to be critical for survival. The second factor, is related to the difficulty of studying the biology of organisms at low temperatures, where the kinetics of reactions, and the time required for processes to reach an equilibrium can be problematic when conducting experiments. As noted in Gusta and Wisniewski (2013), the admonition made by Felix Franks in his book on the biophysics of water at low temperature (Franks, 1985), is very relevant. "Too frequently experimental observations on highly complex systems are based on measurements performed under non-equilibrium conditions and rationalized in terms of elementary textbook science. The degree of undercooling (mostly presented using the incorrect terminology, supercooling), the mechanism of ice nucleation, the growth and type of crystals, their size and distribution, the flow properties of the unfrozen matrix, and long-term effects of aging, all need to be taken into account." The book published by Franks, 1985 still serves as an invaluable primer on the low temperature biology.

Cold hardiness adaptations in plants have been divided into two general categories, tolerance and avoidance. The former involves transcriptomic reprogramming and a host of subsequent biochemical changes that allow plants to tolerate freezing temperatures and the presence of ice in their tissues, while the latter involves mechanisms that allows pockets of water to remain undercooled (deep supercooling) to very low, sub-zero temperatures (-20 to $-40°C$), so that the supercooled cells are not exposed to the dehydrative effects associated with a freeze tolerance response (often referred to as extracellular freezing).

Deep supercooling is characteristic of the dormant buds of many woody perennials and the xylem parenchyma cells of many temperate tree species. The terms freeze tolerance and avoidance, however, are somewhat inaccurate, though widely used, as in both cases cells are avoiding freezing. In the case of freezing tolerance, this is accomplished by the loss of cellular water to extracellular ice, which then decreases the freezing point of the cytoplasm. In the second case, water is not relocated to sites of extracellular ice, even though extracellular ice is present, but instead remains in a metastable condition, and prone to "flash" intracellular freezing (Fujikawa et al., 2009; Wisniewski et al., 2014b). Processes relevant to these strategies are ice nucleation and propagation (Wisniewski et al., 2009, 2014b), the ability to specifically determine where ice crystals are initiated in plant tissues and what shape they form as they grow (McCully et al., 2004), and the formation of cryoprotective and antifreeze compounds (Duman and Wisniewski, 2014).

Despite the complexity of plant cold hardiness, considerable progress has been made in understanding the various components that comprise cold hardiness (Gusta and Wisniewski, 2013). This mini-review highlights one area where considerable progress has been made in understanding the genetic regulation of cold acclimation, and another topic, deacclimation, that is deserving of considerable more focus due to the erratic patterns of warming and cooling temperatures that have developed in the context of climate change. These highly variable weather patterns have had a major impact on dormancy, cold acclimation, and chilling requirements.

THE MOLECULAR REGULATION OF PLANT COLD HARDINESS

Plants cannot move but rather must adapt to a stressful environment. Genes encoding transcription factors in the model plant *Arabidopsis* constitute 6–10% of their genome, compared to 5% in humans, and it is therefore not surprising that adaptation to stresses in plants includes a dramatic change in transcriptional cascades (Pireyre and Burow, 2015). The C-repeat binding factor (CBF), transcription factor pathway has been demonstrated to play an exceptionally important role in plant cold acclimation, a process in which low temperatures lead to biochemical and physiological changes that confer freezing tolerance. These changes are largely associated with the expression of so-called Cold Responsive (COR) genes. In *Arabidopsis*, two or three CBFs co-regulate, often with other transcription factors, more than two-thirds of *COR* genes (Shi et al., 2017). The increase in frost tolerance under ambient conditions that has been demonstrated to occur in many plants as a result of *CBF* overexpression, and the decrease in frost tolerance in *CBF* triple mutants (Jia et al., 2016; Zhao et al., 2016), further underscores the importance of *CBF* genes. *CBF* overexpression in herbaceous and tree species, however, can also reduce growth and induce dormancy (Wisniewski et al., 2011, 2014a), thus it is not surprising that CBF activity is tightly regulated and exhibits only short periods of elevated presence in an active form. This regulation occurs at various levels, including

transcriptional (transcript quantity and variant), translational (protein quantity), and post-translational (protein activity), and can have an immediate effect, because many changes are made to pre-existing molecules. Detailed insights into the regulation of the CBF pathway in *Arabidopsis* has only recently been emerging (**Figure 1**), while limited information for other plant species suggests they have similar but unique regulatory processes of their own. An overview of the information on *Arabidopsis,* and woody plants when available, is presented in the current mini-review.

Low temperature-induced chromatin modification provides physical access to certain genes and allows their transcription. This apparently includes access to and activation of *COR* genes by CBFs (Park et al., 2018). *CBF* gene expression itself is regulated by a large number of transcription factors (Shi et al., 2018). Transcriptional activators include ICE1 and 2 (Inducer of CBF expression 1 and 2), CAMTA 3 and 5 (Calmodulin binding transcription activator 3 and 5), CESTA, BZR1 (Brassinazole-resistant 1), and CCA1/LHY1 (circadian clock-associated 1/late elongated hypocotyl). In contrast, CBF transcription repressors include MYB15, EIN3 (Ethylene insensitive 3), PIF3 (phytochrome-interacting factor 3), and PIF4/7. The activity of these transcription factors is modulated by low temperature, light, and/or a circadian clock in such a way that it results in *CBF* expression only at specific times during constant low-temperature conditions (**Figure 1**). Each transduction most likely also involves one or more hormones (Eremina et al., 2016; Barrero-Gill and Salinas, 2017; Li et al., 2017b; Zhou et al., 2017), whereby CBFs also affect hormone levels (Li et al., 2017c), but details are currently relatively sparse. The fact that a gradual or rapid decrease in temperature has slightly different effects make the CBF pathway even more complicated (Kidokoro et al., 2017).

In addition to transcript levels, the type of transcripts can also be altered in response to a cold period. Approximately 60% of intron-containing genes in *Arabidopsis* were reported to undergo alternative splicing (Marquez et al., 2012), especially under stress conditions. Recent RNAseq analysis for *Arabidopsis* identified the often rapid cold induction of alternative splicing (AS) of over 2,400 genes, with over 1,600 regulated only at the AS level and therefore not detected in most previous analyses (Calixto et al., 2018). *CBF* genes do not have introns, thus AS does not directly affect them, however, alternatively spliced transcripts have been detected for PIF7, PHYB, and CAMTA3 (Calixto et al., 2018), which may alter *CBF* expression.

Once produced, the stability of CBF1/3 proteins is downregulated by their interaction with cold-induced CRPK1-phosphorylated 14-3-3 protein (Liu et al., 2017), and upregulated by their interaction with cold-induced OST1-phosphorylated basic transcription factor 3/BTF3-like protein (BTF3/BTF3L; Ding et al., 2018). While some reports have suggested that *Arabidopsis* CBF1-3 are equally important, others suggest that *Arabidopsis* CBF2 and CBF3 play a more important role in directing the cold response (Jia et al., 2016; Zhao et al., 2016; Shi et al., 2017), and adaptation to low temperature in natural populations (Gehan et al., 2015). This apparently occurs through the employment of different regulons. Whereas the main genes

upregulated in CBF2 overexpressing plants were related to lipid localization, starch metabolic process, light stimulus response, and regulation of transcription, the genes regulated by CBF3 were mainly related to oxidative stress response (Li et al., 2017c).

The presence of a similar pathway in perennial woody plants, such as poplar, apple, grape, *Prunus* sp., and eucalyptus, is suggested by the identification of usually a larger number of *ICE*- and *CBF*-like genes, increasing the possible further delineation of functions (Wisniewski et al., 2014a). Investigations into their regulation found that *ICE* RNA levels are not much affected by treatments, suggesting an emphasis on regulation of *ICE* by post-translational modifications (Wisniewski et al., 2014a). *CBF* expression is often induced by low temperatures and/or drought or high salt (Wisniewski et al., 2014a), can be affected by the circadian clock (Artlip et al., 2013), and, for some *CBFs*, is induced by a continuous cold treatment later and/or for a longer time period than reported for the *Arabidopsis CBFs* (Xiao et al., 2006, 2008; Artlip et al., 2013; Leyva-Pérez et al., 2015; Li et al., 2017d). AS was determined to be a prevalent occurence in the transcription of many genes in apple, orange, and grape, with genes in grape plants showing the most AS events (Sablok et al., 2017). Not much is currently known about AS events, however, in *ICE*, or *COR* genes in low temperature conditions, except for *ICE* transcripts in grape (Rahman et al., 2014). AS, however, has been suggested to regulate responses to environmental stresses in many plants, including Western poplar (*Populus trichocarpa*) (Filichkin et al., 2018). Amino acid motifs thought to be involved in post-translational modifications have been identified in predicted sequences for ICE and CBF proteins, and functional studies suggest that they are important (Feng et al., 2012; Nguyen et al., 2016; Carlow et al., 2017), but much more studies are needed to determine when and how they regulate ICE and CBF activity. Together the collective studies suggest that woody perennial plants have a CBF-like pathway similar to *Arabidopsis* (Benedict et al., 2006), even including a trade-off between growth and cold stress tolerance (Ibanez et al., 2010; Tillett et al., 2011; Nguyen et al., 2016) but details of *CBF* regulation in woody plants is still very limited.

PROSPECTS FOR THE GENERATION OF PLANTS WITH ENHANCED FREEZING TOLERANCE

The tight regulation of CBF activity is largely lost in transgenic plants using a *CBF* construct driven by a constitutive 35S promoter, thus the use of natural promoters is preferred. Recent reports suggest that plants with constitutive brassinosteroid (BR) response display higher *CBF* expression but no signs of growth retardation (Li et al., 2017b). Therefore, it may be possible to avoid the reduced growth associated with *CBF* overexpression (Artlip et al., 2014), if the "correct" CBF or CBF regulators are utilized. Recently, Dong et al. (2017) conducted a meta-analysis of the effect of *CBF* overexpression on temperature stress tolerance and related responses. In that study data from 75 published articles were analyzed to determine the impact of a host of factors such as origin of the *CBF* gene, promoter
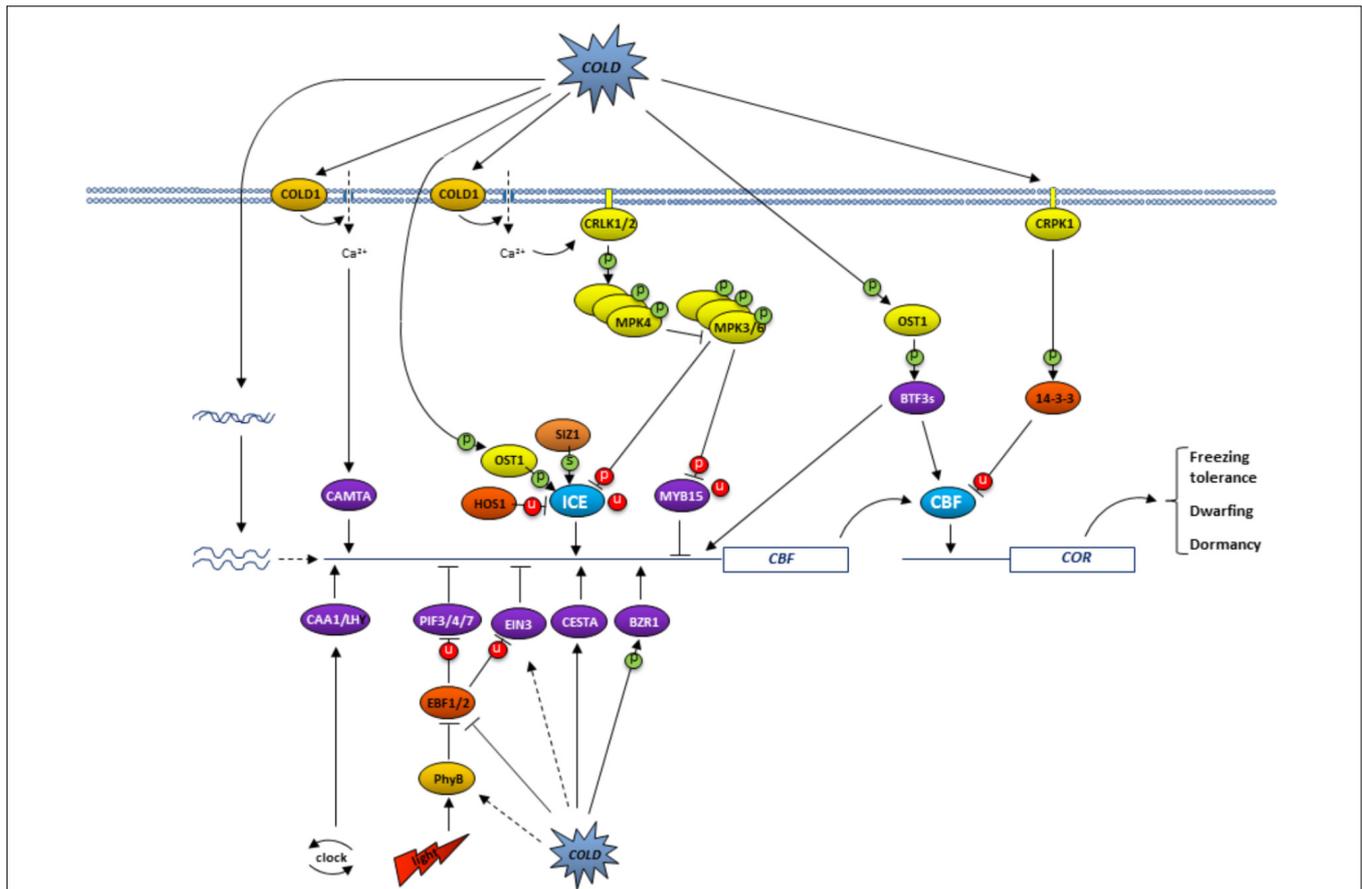
FIGURE 1 | Overview of the regulation of the CBF pathway in *Arabidopsis*. Low temperatures trigger plasma membrane rigidification which leads, presumably via COLD1-like protein, to the opening of Ca^{2+} channels. The resulting higher calcium levels activate CRLK1/2 (calcium/calmodulin-regulated receptor-like kinase; Yang et al., 2010a,b). In turn, this CRKL1/2 triggers the MEKK1-MMK2-MPK4 cascade to ultimately increase ICE activity, because it inhibits the phosphorylation of ICE1 by MPK3/6 and subsequent ICE1 ubiquitination and degradation (Li et al., 2017a; Zhao et al., 2017). ICE activity is further regulated by low temperatures via OST1 (open stomata 1), induced phosphorylation (Ding et al., 2015), and SIZ1-induced sumoylation (Miura et al., 2007), both of which interfere with HOS1 (high expression of osmotically responsive protein 1), directed ubiquitination and subsequent degradation of ICE (Dong et al., 2006). The resulting active ICE directs *CBF* expression. Low temperature activated phosphorylation of 14-3-3 proteins by CRPK1 cause the degradation of CBF proteins (Liu et al., 2017). In contrast, cold-induced OST1-directed phosphorylation of BTF3s promotes its binding to CBFs and thereby prevents CBF degradation (Ding et al., 2018). Photoperiod regulates *CBF* expression via red light perception by PhyB and subsequent degradation of PIF3 (phytochrome-interacting factor 3), thereby relieving its inhibition of *CBF* expression (Jiang et al., 2017), whereas the circadian clock regulates CCA1 and LHY activity (Dong et al., 2011). Interestingly, PIF3 stability is increased by low temperatures, presumably at a later time to downregulate *CBF* expression (Jiang et al., 2017). PIF4/7 and EIN3 (Ethylene insensitive 3) downregulate whereas BZR and CESTA upregulate *CBF* expression, but how this is triggered is not yet known (Shi et al., 2018). Phosphorylation, sumoylation and ubiquitination events are indicated by P, S and U, respectively, with activating modifications in green and inhibiting modifications in red.

used to drive expression, the method of stress evaluation, etc., on temperature response and associated indicators, such as electrolyte leakage, growth, chlorophyll fluorescence, sugar and proline levels, etc., Results indicated that 7 of 8 measured variables were significantly modulated in *CBF* (*DREB*)-transgenic plants, while two of the eight parameters were only modulated in non-stressed plants. The measured parameters were modulated by 32% or more by various experimental variables. The modulating variables included, acclimated vs. non-acclimated, type of promoter, duration of stress and its severity, source of the donor gene, and whether the donor and recipient were the same genus. *CBF* overexpression had a consistent negative impact on plant height, a reduction in electrolyte leakage, and positive impact on survival. The impact was evident in both

acclimated and non-acclimated plants, although the greatest impact was observed in acclimated plants. Such analyses may provide a more comprehensive understanding of how to best utilize *CBF* genes with modified promoters to improve freezing stress tolerance.

An alternative, enterprising approach is the modification of endogenous *CBF* genes into variants that lead to a higher frost tolerance using a clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated protein 9 nuclease (Cas9)-like system which has recently been optimized for use in vegetatively propagated perennial plants (Chen et al., 2018). For example, existing sequences could be modified based on variants present in more frost tolerant cultivars or species (Carlow et al., 2017; Li et al., 2017d), and thereby change their

regulation. Because of the precise change, the resulting plants may not be considered genetically modified by government regulatory agencies and may be more acceptable to the general public.

DEACCLIMATION (DA) RESPONSE, A CRITICAL FACTOR FOR WINTER-SURVIVAL

The ability to increase freeze-tolerance in temperate- and boreal-zone woody perennials (fruit and forest tree species) via autumnal cold acclimation is undoubtedly the first line of defense against harsh and long winters. Seasonally induced freeze-tolerance is lost under relatively warmer conditions in a process called 'deacclimation', a process that typically occurs in response to spring-warming. The maintenance of a sufficient level of cold-induced freeze-tolerance until the danger of killing frosts is passed, however, is an imperative to avoid frost-damage. For example, erratic temperature fluctuations, i.e., sudden winter-warming or premature spring-like conditions followed by more "normal" freezing temperatures, could render partially or fully deacclimated tissues vulnerable to freeze-damage. Indeed, the frequency of such fluctuations has been increasing (Jentsch et al., 2007; IPCC, 2014), and some of the most devastating killer-frosts across North America have been attributed to such events, e.g., Easter freeze of 2007 (Gu et al., 2008), Mother's Day freeze of 2010, killer frost of 2012, and the polar vortex of 2014. Field simulations of winter-warming events have also confirmed their damaging effects on overwintering perennials (Taulavuori et al., 1997; Bokhorst et al., 2009, 2010).

DORMANCY STATUS AND SPRING PHENOLOGY (BUDBREAK) IN RELATION TO DA

Temperate trees have evolved the ability to tolerate harsh winters by undergoing a period of endodormancy (rest), during which cold-acclimated meristems are less prone to DA when trees are exposed to unseasonal episodes of warming (Kalberer et al., 2006, 2007a,b). Buds of native temperate and boreal trees must satisfy a genetically defined chilling requirement to exit endodormancy (Richardson et al., 1974). Post-endodormancy, buds enter an ecodormant state where they must be exposed to a genetically defined threshold of warming ('heat units') (Charrier et al., 2011), for the resumption of meristematic activity and growth to occur (budbreak or spring phenology). Ecodormant buds are substantially more sensitive to warmer temperatures and DA than endodormant buds (Kalberer et al., 2007a). This sensitivity increases progressively as the period of ecodormancy increases, finally culminating in complete DA and spring budbreak (Kalberer et al., 2006; Arora and Taulavuori, 2016). Any shift in this annual cycle of spring phenology could potentially increase the risk of trees encountering frost injury (Vitasse et al., 2014 and references therein). Whether

an increased risk occurs depends on a variety of internal factors, including species, chilling requirement, the genetic ability to resist deacclimation in response to transient, unseasonal episodes of warm temperatures, and the capacity to reacclimate. External (environmental) mitigating factors include, temperature fluctuations (intensity and timing), and the region/site (latitude and altitude) where the trees are located (Pagter and Arora, 2013; Arora and Taulavuori, 2016; Vitasse et al., 2018b). The sensitivity of ecodormant buds to deacclimating temperatures has also been reported to increase with increasing photoperiod in spring in species such as European beech (*Fagus sylvatica*) (Vitasse and Basler, 2013). There is ample evidence that warming trends in recent decades have advanced spring budbreak and leaf development in many plant species growing in cold regions (Penuelas and Filella, 2001; Menzel et al., 2006, and references therein). Some studies have indicated, however, that the degree of advancement in spring phenology appears to be declining in the recent years (Yu et al., 2010; Fu et al., 2015).

Advances in spring phenology due to climate change can occur under two scenarios. In the first scenario, a faster than normal accumulation of heat-units by ecodormant buds occurs due to earlier and warmer spring-like temperatures (Cleland et al., 2007). This could render prematurely deacclimated buds vulnerable to subsequent spring frosts. The second, somewhat ignored and paradoxical scenario, involves a more rapid fulfillment of chilling requirement due to warmer winter temperatures than has occurred in more typical, historical winters in certain regions. For example, tree species in northern latitudes or high elevations could experience a greater level of 'dormancy-breaking chill units' since the warmer winter temperatures could expose trees to temperatures $>0°C$ that are more effective in breaking endodormancy and reduce exposure to sub-freezing temperatures that do not contribute to chill unit accumulation (Hanninen, 2006). Based on this premise, spring phenology could be expected to advance more rapidly in historically colder areas under warming climate conditions. This would result in premature deacclimation and a greater risk of spring freeze-damage. Indeed, Vitasse et al. (2018a) reported that spring phenology in fruit (apple, cherry), and forest (Norway spruce and European beech) trees has advanced at a faster rate during 1975–2016 at fifty high elevation locations in Switzerland than in other temperate locations. The authors argued that even if the frequency and severity of late spring frosts remains unchanged in the future or changes less than the spring phenology of plants, deacclimated organs may be more exposed to a greater number of freeze-damage events (Vitasse et al., 2018a).

On the other hand, proponents of a possible decline in the advancement or delay in spring phenology by warming climate make their case as follows. They suggest that elevated winter temperature may result in a chilling-deficit, i.e., reduced duration and/or sum of cold. And since heat unit requirement for spring phenology is believed to be inversely correlated with the chill accumulation during dormancy (Harrington et al., 2010; Laube et al., 2014), any reduction in accumulated chilling would

result in higher heat-unit requirement, thus slowing down the advance in, or delaying, the spring leaf-unfolding (Yu et al., 2010; Fu et al., 2015). One of these studies (Fu et al., 2015) noted that while spring phenology for seven deciduous forest tree species had advanced by ∼4 d during 1980–1994 in Europe, this response has decreased by ∼40% (to 2.3 d) during 1999–2013 (Pan European Phenology Network). A caveat that must be added for such an observation to be practically and widely applicable, however, is that the temperatures during warmer winters have to be high enough to cause real chilling deficit, i.e., either negate accumulated chilling or be ineffective to meet the chilling requirement.

Although the focus of this mini-review is not on bud dormancy, it is relevant to note that several studies have associated the genetic regulation of chilling requirement and dormancy with *Dormancy Associated MADs-box* (*DAM*) genes in peach (Bielenberg et al., 2008; Li et al., 2009) apple (Wu et al., 2017a), pear (Saito et al., 2013), apricot (Sasaki et al., 2011), leafy spurge (*Euphorbia esula*) (Horvath et al., 2010), kiwifruit (Wu et al., 2011, 2017b), and tea plant (*Camellia sinensis*) (Hao et al., 2017). Bud-dormancy associated candidate genes have also been identified in blackcurrant (*Ribes nigrum*) by Hedley et al. (2010) and Yordanov et al. (2014) identified an *early budbreak* (*EBB*) gene in poplar that was an APETALA2/Ethylene responsive transcription factor responsible for early bud flush.

C-repeat binding factor-binding motifs have been identified in promoters of several *DAM* genes and an *EBB* homolog in apple (Wisniewski et al., 2015a), as well as other plant species. A comprehensive analysis of *DAM* genes in the ornamental woody plant, *P. mume* (Chinese plum, Japanese apricot) demonstrated an interaction between *CBF*s and *DAM* genes, especially *PMCBF1 - PMDAM1* (Zhao et al., 2018) and *CBF* expression lowered whereas MADS-box gene (1 and 3) expression increased in almond flower buds after bud break (Barros et al., 2012). Wisniewski et al. (2015a) noted that overexpression of a peach *CBF* gene (*PpCBF1*) in apple altered the expression of *DAM*, *EBB*, and *RGL* (DELLA) genes and that some members of each of these gene families contained C-repeat regions in their promoter regions that are the target sites for CBF. They provided a model linking *CBF* expression with the regulation of dormancy, bud-break, freezing tolerance, and growth. Interestingly, a subsequent study indicated that the impact of the apple transgenic rootstock overexpressing *CBF* was not graft-transmissible and thus did not affect the cold hardiness of dormancy of the scion cultivar grafted to the transgenic 'M.26' rootstock, although growth and flowering were significantly impacted (Artlip et al., 2016).

Current research has highlighted the impact of dormancy status and spring phenology on the propensity of trees to deacclimate. Spring-phenology is an outcome influenced by both the chilling and heat unit requirements of overwintering tree species. The fact that chill- and heat-units can be satisfied by the same temperatures for certain species (Cooke et al., 2012 and references therein) makes their combined effect on spring phenology even more complex. It is therefore critical to include dormancy-status and the interactions between chilling and heat requirements as key parameters in models designed to predict the relationship between deacclimation response and freezing-tolerance.

FUTURE DIRECTIONS

The past 50 years of research has provided a wealth of information on the genetic and molecular regulation of plant cold hardiness, as well as the regulation of dormancy. These advances have been fueled by new technologies associated with high-throughput sequencing, genetic mapping, and transformation technologies. In particular the regulatory role of *CBF* genes in freezing tolerance, and of *DAM* genes in the regulation of chilling requirement stand out as major advances. The discovery that CBF activity is regulated continually and at various levels, helps explain, at least in part, why slightly different treatment of plants with respect to light, duration and rate of low temperature treatment for example, lead to different outcomes with respect to their frost tolerance. Notable advances have also been made with the use of high-resolution infrared thermography in our understanding of ice nucleation and propagation (Wisniewski et al., 2014b, 2015b), and the properties of antifreeze proteins (Duman and Wisniewski, 2014). Despite these advances, significant improvements in plant cold hardiness have been elusive and problematic due to the complexity of this trait and its intimate connection to other plant developmental processes, especially growth and flowering. In addition, the relatively new field of epigenetics has demonstrated the key role that the environment can play on imprinting plant response to abiotic stress (Kumar, 2018).

Future studies will need to better understand the cross-talk that occurs between different plant developmental processes and how it can be manipulated in a prescribed manner. A key question will be whether processes that determine cold hardiness can be separated from processes that restrict growth. Can *CBF* genes be regulated in a manner that removes their negative impact on growth and reproductive output? What are the genetic mechanisms that can be used to tease these processes apart? Which *CBF* gene or gene variant present in more frost tolerant species is best targeted for manipulation and how can epigenetic modifications affecting their activity best be harnessed?

Although, not as glamorous, a comprehensive understanding of the underlying biophysical mechanisms responsible for freeze avoidance, especially in woody plants, is still lacking. Deep supercooling of xylem parenchyma (Wisniewski, 1995; Fujikawa et al., 2009) and floral buds (Kuprian et al., 2016, 2017) is an integral aspect of the cold hardiness of many temperate tree species, especially fruit trees, however, few advances have been made on this topic over the past 30 years. What new technologies that can be applied to better understand how, when, and where ice is initiated in plants, how it is propagated, and how the size and shape of ice crystals are regulated. Genetic studies of the inheritance of avoidance traits, such as supercooling, have yet to be conducted, but would provide very useful information.

An integrated approach that takes into account the complexity of traits that contribute to plant cold hardiness will be needed to achieve advances that can be translated into practical solutions that address the challenges of a rapidly changing climate.

AUTHOR CONTRIBUTIONS

MW was responsible for the general overview of the opinions stated in the manuscript and any faults or shortcoming in logic directly fall on him. AN contributed the overview of CBF regulation, and RA the information on the importance of deacclimation in a changing climate. All authors reviewed and agreed with the final version of the submitted manuscript.

REFERENCES

Arora, R., and Taulavuori, K. (2016). Increased risk of freeze damage in woody perennials vis-à-vis climate change: importance of deacclimation and dormancy response. *Front. Environ. Sci.* 4:44. doi: 10.3389/fenvs.2016.00044

Arora, R. (2018). Mechanism of freeze-thaw injury and recovery: a cool retrospective and warming up to new ideas. *Plant Sci.* 270, 301–313. doi: 10.1016/j.plantsci.2018.03.002

Artlip, T. A., Wisniewski, M., Bassett, C. L., and Norelli, J. L. (2013). CBF gene expression in peach leaf and bark tissues is gated by a circadian clock. *Tree Physiol.* 33, 866–877. doi: 10.1093/treephys/tpt056

Artlip, T., Wisniewski, M., Arora, R., and Norelli, J. (2016). An apple rootstock overexpressing a peach CBF gene alters growth and flowering in the scion but does not impact cold hardiness or dormancy. *Hort. Res.* 3:16006. doi: 10.1038/hortres.2016.6

Artlip, T., Wisniewski, M., and Norelli, J. (2014). Field evaluations of apple overexpressing a peach CBF gene confirms its effect on cold hardiness, dormancy, and growth. *Environ. Exp. Bot.* 106, 79–86. doi: 10.1038/hortres.2016.6

Barrero-Gill, J., and Salinas, J. (2017). CBFs at the crossroads of plant hormone signaling in cold stress response. *Mol. Plant.* 10, 542–544. doi: 10.1016/j.molp.2017.03.004

Barros, P. M., Gonçalves, N., Saibo, N. J. M., and Oliveira, M. M. (2012). Cold acclimation and floral development in almond bud break: insights into the regulatory pathways. *J. Exp. Biol.* 63, 4585–4596. doi: 10.1093/jxb/ers144

Benedict, C., Skinner, J. S., Meng, R., Chang, Y., Bhalerao, R., Huner, N. P. A., et al. (2006). The CBF1-dependent low temperature signaling pathway, regulon and increase in freeze tolerance are conserved in *Populus* spp. *Plant Cell Environ.* 29, 1259–1272. doi: 10.1111/j.1365-3040.2006.01505

Bielenberg, D. G., Wang, Y., Li, Z., Zhebentyayeva, T., Fan, S., Reighard, G., et al. (2008). Sequencing and annotation of the evergrowing locus in peach (*Prunus persica* L. Batsch) reveals a cluster of six MADS-box transcription factors as candidate genes for regulation of terminal bud formation. *Tree Genet. Genomes* 4, 495–507. doi: 10.1007/s11295-007-0126-9

Bokhorst, S., Bjerke, J. W., Davey, M. P., Taulavuori, K., Taulavuori, E., Laine, K., et al. (2010). Impacts of extreme winter warming events on plant physiology in a sub-Arctic heath community. *Physiol. Plant.* 140, 128–140. doi: 10.1111/j.1399-3054.2010.01386.x

Bokhorst, S., Bjerke, J. W., Tommervik, H., Callaghan, T. V., and Phoenix, G. K. (2009). Winter warming events damage sub-Arctic vegetation: consistent evidence from an experimental manipulation and natural event. *J. Ecol.* 97, 1408–1415. doi: 10.1111/j.1365-2745.2009.01554.x

Calixto, C. P. G., Guo, W., James, A. B., Tzioutziou, N. A., Entizne, J. C., Panter, P. E., et al. (2018). Rapid and dynamic alternative splicing impacts the Arabidopsis cold response transcriptome. *Plant Cell* 30, 1424–1444. doi: 10.1105/tpc.18.00177

Carlow, C. E., Faultless, J. T., Lee, C., Siddiqua, M., Edge, A., and Nassuth, A. (2017). Nuclear localization and transactivation by Vitis CBF transcription factors are regulated by combinations of conserved amino acid domains. *Plant Phys. Biochem.* 118, 306–319. doi: 10.1016/j.plaphy.2017.06.027

Charrier, G., Bonhomme, M., Lacointe, A., and Ameglio, T. (2011). Are budburst dates, dormancy and cold acclimation in walnut trees (*Juglans regia* L.) under mainly genotypic or environmental control? *Int. J. Biometeorol.* 55, 763–774. doi: 10.1007/s00484-011-0470-1

Chen, L., Li, W., Katin-Grazzini, L., Ding, J., Gu, X., Li, Y., et al. (2018). A method for the production and expedient screening of CRISPR/Cas9-mediated non-transgenic mutant plants. *Hort. Res.* 5:13. doi: 10.1038/s41438-018-0023-4

Cleland, E. E., Chuine, I., Menzel, A., Mooney, H. A., and Schwartz, M. D. (2007). Shifting plant phenology in response to global change. *Trends Ecol. Evol.* 22, 357–365. doi: 10.1016/j.tree.2007.04.003

Cooke, J. E., Eriksson, M. E., and Juntilla, O. (2012). The dynamic nature of bud dormancy in trees: environmental control and molecular mechanisms. *Plant Cell Environ.* 35, 1707–1728. doi: 10.1111/j.1365-3040.2012.02552.x

Ding, Y., Jia, Y., Shi, Y., Zhang, X., Song, C., Gong, Z., et al. (2018). OST1-mediated BTF3L phosphorylation positively regulates CBFs during plant cold responses. *EMBO J.* 37:e982288. doi: 10.15252/embj.201798228

Ding, Y., Li, H., Zhang, X., Xie, Q., Gong, Z., and Yang, S. (2015). OST1 kinase modulates freezing tolerance by enhancing ICE1 stability in Arabidopsis. *Dev. Cell* 32, 278–289. doi: 10.1016/j.devcel.2014.12.023

Dong, C., Ma, Y., Wisniewski, M., and Cheng, Z.-M. (2017). Meta-analysis of the effect of overexpression of CBF/DREB family genes on drought stress response. *Environ. Exp. Bot.* 142, 1–14. doi: 10.1016/j.envexpbot.2017.07.014

Dong, C. H., Agarwal, M., Zhang, Y., Xie, Q., and Zhu, J. K. (2006). The negative regulator of plant cold responses, HOS1, is a RING E3 ligase that mediates the ubiquitination and degradation of ICE1. *Proc. Natl. Acad. Sci. U.S.A.* 103, 8281–8286. doi: 10.1073/pnas.0602874103

Dong, M. A., Farre, E. M., and Thomashow, M. F. (2011). Circadian clock-associated 1 and late elongated hypocotyl regulate expression of the C-repeat binding factor (CBF) pathway in Arabidopsis. *Proc. Natl. Acad. Sci. U.S.A.* 108, 7241–7246. doi: 10.1073/pnas.1103741108

Duman, J. G., and Wisniewski, M. (2014). The use of antifreeze proteins for frost protection in sensitive crop plants. *Environ. Exp. Bot.* 106, 60–69. doi: 10.1016/j.envexpbot.2014.01.001

Eremina, M., Rozhon, W., and Poppenberger, B. (2016). Hormonal control of cold stress responses in plants. *Cell. Mol. Life Sci.* 73, 797–810. doi: 10.1007/s00018-015-2089-6

Feng, X.-M., Zhao, Q., Zhao, L.-L., Qiao, Y., Xie, X.-B., Li, H.-F., et al. (2012). The cold-induced basic helix-loop-helix transcription factor gene MdCIbHLH1encodes an ICE-like protein in apple. *BMC Plant Biol.* 12:22. doi: 10.1186/1471-2229-12-22

Filichkin, S. A., Hamilton, M., Dharmawardhana, P. D., Singh, S. K., Sullivan, C., Ben-Hur, A., et al. (2018). Abiotic stresses modulate landscape of poplar transcriptome via alternative splicing, differential intron retention, and isoform ratio switching. *Front. Plant Sci.* 9:5. doi: 10.3389/fpls.2018.00005

Franks, F. (1985). *Biophysics and Biochemistry at Low Temperatures*. Cambridge: Cambridge University Press.

Fu, Y. H., Zhao, H., Piao, S., Peaucelle, M., Peng, S., Zhou, G., et al. (2015). Declining global warming effects on the phenology of spring leaf unfolding. *Nature* 526, 104–107. doi: 10.1038/nature15402

Fujikawa, S., Kasuga, J., Takata, N., and Arakawa, K. (2009). "Factors related to changes of deep supercooling capability in xylem parenchyma cells of trees," in *Plant Cold Hardiness: From the Laboratory to the Field*, eds L. V. Gusta, M. E. Wisniewski, and K. Tanino (Oxfordshire: CABI), 29–42.

Gehan, M. A., Park, S., Gilmour, S. J., An, C., Lee, C.-M., and Thomashow, M. F. (2015). Natural variation in the C-repeat binding factor cold response pathway correlates with local adaptation of Arabidopsis ecotypes. *Plant J.* 84, 682–693. doi: 10.1111/tpj.13027

Gu, I., Hanson, P. J., Post, W. M., Kaiser, D. P., Yang, B., Nemani, R., et al. (2008). The 2007 eastern US spring freeze: increased cold damage in a warming world? *Bioscience* 58, 253–262. doi: 10.1641/B580311

Gusta, L. V., and Wisniewski, M. (2013). Understanding plant cold hardiness: an opinion. *Physiol. Plant.* 147, 4–14. doi: 10.1111/j.1399-3054.2012.01611.x

Hanninen, H. (2006). Climate warming and the risk of frost damage to boreal forest trees: identification of critical ecophysiological traits. *Tree Physiol.* 26, 889–898. doi: 10.1093/treephys/26.7.889

Hao, X., Yang, Y., Yue, C., Wang, L., Horvath, D. P., and Wang, X. (2017). Comprehensive transcriptome analyses reveal differential gene expression profiles of *Camellia sinensis* axillary buds at para-, endo-, ecodormancy, and bud flush stages. *Front. Plant Sci.* 8:553. doi: 10.3389/fpls.2017.00553

Harrington, C. A., Gould, P. J., and St. Clair, J. B. (2010). Modeling the effects of winter environment on dormancy release of Douglas-fir. *For. Ecol. Manag.* 259, 798–808. doi: 10.1016/j.foreco.2009.06.018

Hedley, P. E., Russell, J. R., Jorgensen, L., Gordon, S., Morris, J. A., Hackett, C. A., et al. (2010). Candidate genes associated with bud dormancy release in blackcurrant (*Ribes nigrum* L.). *BMC Plant Biol.* 10:202. doi: 10.1186/1471-2229-10-202

Horvath, D. P., Sung, S., Kim, D., Chao, W., and Anderson, J. (2010). Characterization, expression and function of DORMANCY ASSOCIATED MADS-BOX genes from leafy spurge. *Plant Mol. Biol.* 73, 169–179. doi: 10.1007/s11103-009-9596-5

Ibanez, C., Kozarewa, I., Johansson, M., Ogren, E., Rohde, A., and Eriksson, M. E. (2010). Circadian clock components regulate entry and affect exit of seasonal dormancy as well as winter hardiness in populus trees. *Plant Physiol.* 153, 1823–1833. doi: 10.1104/pp.110.158220

IPCC (2014). *Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II of the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC)*, eds C. B. Field, et al. (Cambridge: Cambridge University Press).

Jentsch, A., Kreyling, J., and Beierkuhnlein, C. (2007). A new generation of climate-change experiments: events, not trends. *Front. Ecol. Environ.* 5:7. doi: 10.1890/1540

Jia, Y., Ding, Y., Shi, Y., Zhang, X., Gong, Z., and Yang, S. (2016). The cbfs triple mutants reveal the essential functions of CBFs in cold acclimation and allow the definition of CBF regulons in *Arabidopsis*. *New Phytol.* 212, 345–353. doi: 10.1111/nph.14088

Jiang, B., Shi, Y., Zhang, X., Xin, X., Qi, L., Guo, H., et al. (2017). PIF3 is a negative regulator of the CBF pathway and freezing tolerance in Arabidopsis. *Proc. Natl. Acad. Sci. U.S.A.* 114, E6695–E6702. doi: 10.1073/pnas.1706226114

Kalberer, S. R., Leyva-Estrada, N., Krebs, S. L., and Arora, R. (2007a). Cold hardiness of floral buds of deciduous azaleas: dehardening, rehardening, and endodormancy in late winter. *J. Amer. Soc. Hort. Sci.* 132, 73–79.

Kalberer, S. R., Leyva-Estrada, N., Krebs, S. L., and Arora, R. (2007b). Frost dehardening and rehardening of floral buds of deciduous azaleas depend on genotypic biogeography. *Environ. Exp. Bot.* 59, 264–275. doi: 10.1016/j.envexpbot.2006.02.001

Kalberer, S. R., Wisniewski, M., and Arora, R. (2006). Deacclimation and reacclimation of cold-hardy plants: current understanding and emerging concepts. *Plant Sci.* 171, 3–16. doi: 10.1016/j.plantsci.2006.02.013

Kidokoro, S., Yoneda, K., Takasaki, H., Takahashi, F., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2017). Different cold-signaling pathways function in the responses to rapid and gradual decreases in temperature. *Plant Cell* 29, 760–774. doi: 10.1105/tpc.16.00669

Kumar, S. (2018). Epigenetic memory of stress responses in plants. *J. Phytochem. Biochem.* 2:1.

Kuprian, E., Koch, S., Munkler, S., Resnyak, A., Buchner, O., Oberhammer, M., et al. (2017). Does winter desiccation account for seasonal increases in supercooling capacity of Norway spruce bud primordia. *Tree Physiol.* 38, 591–601. doi: 10.1093treephys/tpx142

Kuprian, E., Tuong, T. D., Pfaller, K., Wagner, J., Livingston, D. P. III, and Neuner, G. (2016). Persistent supercooling of reproductive shoots in enabled by structural ice barriers being activie despite and intact xylem connection. *PLoS One* 11:e0163160. doi: 10.1371/journal.pone.0163160

Laube, J., Sparks, T. H., Estrella, N., Hofler, J. F., Ankerst, D. P., and Menzel, A. (2014). Chilling outweighs photoperiod in preventing precocious spring development. *Glob. Change Biol.* 20, 170–182. doi: 10.1111/gcb.12360

Leyva-Pérez, M. L., Valverde-Corredor, A., Valderrama, R., Jiménez-Ruiz, J., Muñoz-Merida, A., Trelles, O., et al. (2015). Early and delayed long-term transcriptional changes and short-term transient responses during cold acclimation in olive leaves. *DNA Res.* 22, 1–11. doi: 10.1093/dnares/dsu033

Li, A., Zhou, M., Wei, D., Chen, H., You, C., and Lin, J. (2017c). Transcriptome profiling reveals the negative regulation of multiple plant hormone signaling pathways elicited by overexpression of C-repeat binding factors. *Front. Plant Sci.* 8:1647. doi: 10.3389/fpls.2017.01647

Li, H., Ding, Y., Shi, Y., Zhang, X., Zhang, S., Gong, Z., et al. (2017a). MPK3- and MPK6-mediated ICE1 phosphorylation negatively regulates ICE1 stability and freezing tolerance in *Arabidopsis. Dev. Cell* 43, 630–642. doi: 10.17632/jyxknn3gmr.1

Li, H., Ye, K., Shi, Y., Cheng, J., Zhang, X., and Yang, S. (2017b). BZR1 positively regulates freezing tolerance via CBF-dependent and CBF independent pathways in *Arabidopsis. Mol. Plant.* 10, 545–559. doi: 10.1016/j.molp.2017.01.004

Li, Y., Song, Y., Xu, B., Xie, J., and Zhang, D. (2017d). Poplar CBF1 functions specifically in an integrated cold regulatory network. *Tree Physiol.* 37, 98–115. doi: 10.1093/treephys/tpw079

Li, Z., Reighard, G., Abbott, A., and Bielenberg, D. (2009). Dormancy-associated MADS genes from the EVG locus of peach (*Prunus perica* L. Batsch.) have distinct seasonal and photoperiodic expression patterns. *J. Exp. Bot.* 60, 3521–3530. doi: 10.1093/jxb/erp195

Liu, Z., Jia, Y., Ding, Y., Shi, Y., Li, Z., Guo, Y., et al. (2017). Plasma membrane CRPK1-mediated phosphorylation of 14-3-3 proteins induces their nuclear import to fine-tune CBF signaling during cold response. *Mol. Cell* 66, 117.e5–128.e5. doi: 10.1016/j.molcel.2017.02.016

Marquez, Y., Brown, J. W., Simpson, C., Barta, A., and Kalyna, M. (2012). Transcriptome survey reveals increased complexity of the alternative splicing landscape in *Arabidopsis. Genome Res.* 22, 1184–1195. doi: 10.1101/gr.134106.111

McCully, M. E., Canny, M. J., and Huang, C. X. (2004). The management of extracellular ice by petioles of frost-resistant herbaceous plants. *Ann. Bot.* 94, 665–674. doi: 10.1093/aob/mch191

Menzel, A., Sparks, T. H., Estrella, N., Koch, E., Aasa, A., Ahas, R., et al. (2006). European phenological response to climate change matches the warming pattern. *Glob. Change Biol.* 12, 1969–1976. doi: 10.1111/J.1365-2486.2006.01193.X

Miura, K., Jin, J. B., Lee, J., Yoo, C. Y., Stirm, V., Miura, T., et al. (2007). SIZ1-mediated sumoylation of ICE1 controls CBF3/DREB1A expression and freezing tolerance in *Arabidopsis. Plant Cell* 19, 1403–1414. doi: 10.1105/tpc.106.048397

Molisch, H. (1897). Untersuchungen über das erfrieren der Pflanzen (reprinted in English in 1982). *Cryo Lett.* 3, 332–390.

Nguyen, H. C., Cao, P. B., San Clemente, H., Ployet, R., Mounet, F., Ladouce, N., et al. (2016). Special trends in CBF and DREB2 groups in *Eucalyptus gunnii* vs *Eucalyptus grandis* suggest that CBF are master players in the trade-off between growth and stress resistance. *Physiol. Plant.* 4, 445–467. doi: 10.1111/ppl.12529

Pagter, M., and Arora, R. (2013). Winter survival and deacclimation of perennials under warming climate: physiological perspectives. *Physiol. Plant.* 147, 75–87. doi: 10.1111/j.1399-3054.2012.01650.x

Park, J., Lim, C. J., Shen, M., Park, H. J., Cha, J.-Y., Iniesto, E., et al. (2018). Epigenetic switch from repressive to permissive chromatin in response to cold stress. *Proc. Natl. Acad. Sci. U.S.A.* 115, E5400–E5409. doi: 10.1073/pnas.1721241115

Penuelas, J., and Filella, I. (2001). Responses to a warming world. *Science* 294, 793–795. doi: 10.1126/science.1066860

Pireyre, M., and Burow, M. (2015). Regulation of MYB and bHLH transcription Factors: a glance at the protein level. *Mol. Plant.* 8, 378–388. doi: 10.1016/j.molp.2014.11.022

Rahman, M. A., Moody, M. A., and Nassuth, A. (2014). Grape contains 4 ICE genes whose expression includes alternative polyadenylation, leading to transcripts encoding at least 7 different ICE proteins. *Environ. Exp. Bot.* 106, 70–78. doi: 10.1016/j.envexpbot.2014.01.003

Richardson, E. A., Seeley, S. D., and Walker, D. R. (1974). A model for estimating the completion of rest for 'Redhaven' and 'Elberta' peach trees. *Hortscience* 9, 331–332.

Sablok, G., Powell, B., Braessler, J., Yu, F., and Min, X. J. (2017). Comparative landscape of alternative splicing in fruit plants. *Curr. Plant Biol.* 9-10, 29–36. doi: 10.1016/j.cpb.2017.06.001

Saito, T., Bai, S., Ito, A., Sakamoto, D., Saito, T., EwaUbi, B., et al. (2013). Expression and genomic structure of the dormancy-associated MADS box genes MADS13 in Japanese pears (*Pyrus pyrifolia* Nakia) that differ in their chilling requirement for endodormancy release. *Tree Physiol.* 33, 654–667. doi: 10.1093/treephys/tpt037

Sasaki, R., Yamane, H., Ooka, T., Jotatsu, H., Kitamura, Y., Akagi, T., et al. (2011). Functional and expressional analyses of PmDAM genes associated with endodormancy in Japanese apricot. *Plant Physiol.* 157, 485–497. doi: 10.1104/pp.111.181982

Shi, Y., Ding, Y., and Yang, D. (2018). Molecular regulation of CBF signaling in cold acclimation. *Trends Plant Sci.* 23, 623–637. doi: 10.1016/j.tplants.2018.04.002

Shi, Y., Huang, J., Sun, T., Wang, X., Zhu, C., Ai, Y., et al. (2017). The precise regulation of different COR genes by individual CBF transcription factors in *Arabidopsis thaliana*. *J. Integr. Plant Biol.* 59, 118–133. doi: 10.1111/jipb.12515

Taulavuori, K., Laine, K., Taulavuori, E., Pakonen, T., and Saari, E. (1997). Accelerated dehardening in the bilberry (*Vaccinium myrtillus* L.) induced by a small elevation in air temperature. *Environ. Pollut.* 98, 91–95. doi: 10.1016/s0269-7491(97)00115-2

Tillett, R. L., Wheatley, M. D., Tattersall, E. A., Schlauch, K. A., Cramer, G. R., and Cushman, J. C. (2011). The Vitis vinifera C-repeat binding protein 4 (VvCBF4) transcriptional factor enhances freezing tolerance in wine grape. *Plant Biotechnol. J.* 10, 105–124. doi: 10.1111/j.1467-7652.2011.00648

Vitasse, Y., and Basler, D. (2013). What role for photoperiod in the bud burst phenology of European beech. *J. For. Res.* 132, 1–8. doi: 10.1007/s10342-012-0661-2

Vitasse, Y., Lenz, A., and Korner, C. (2014). The interaction between freezing tolerance and phenology in temperate deciduous trees. *Front. Plant Sci.* 5:541. doi: 10.3389/fpls.2014.00541

Vitasse, Y., Schneider, L., Rixen, C., Christen, D., and Rebetez, M. (2018a). Increase in the risk of exposure of forest and fruit trees to spring frosts at higher elevations in Switzerland over the last four decades. *Agric. For. Meteorol.* 248, 60–69. doi: 10.1016/j.agrformet.2017.09.005

Vitasse, Y., Signarbieux, C., and Fu, Y. H. (2018b). Global warming leads to more uniform spring phenology across elevations. *Proc. Natl. Acad. Sci. U.S.A.* 115, 1004–1008. doi: 10.1073/pnas.1717342115

Wiegand, K. M. (1906). The occurrence of ice in plant tissue. *Plant World* 9, 25–39.

Wisniewski, M. (1995). "Deep supercooling in woody plants and the role of cell wall structure," in *Biological Ice Nucleation and Its Applications*, eds R. E. Lee, G. J. Warren, and L. V. Gusta (Minneapolis, MN: APS Press), 163–181.

Wisniewski, M., Basset, C. L., and Gusta, L. V. (2003). An overview of cold hardiness in woody plants: seeing the forest through the trees. *Hortscience* 38, 952–959.

Wisniewski, M., Gusta, L. V., Fuller, M. P., and Karlson, D. (2009). "Ice nucleation, propagation, and deep supercooling: the lost tribes of freezing studies," in *Plant Cold Hardiness: From the Laboratory to the Field*, eds L. V. Gusta, M. E. Wisniewski, and K. Tanino (Cambridge: CAB International), 1–11.

Wisniewski, M., Gusta, L. V., and Neuner, G. (2014a). Adaptive mechanisms of freeze avoidance in plants: a brief update. *Environ. Exp. Bot.* 99, 133–140. doi: 10.1016/j.envexpbot.2013.11.011

Wisniewski, M., Nassuth, A., Teulières, C., Marque, C., Rowland, J., Cao, P. B., et al. (2014b). Genomics of cold hardiness in woody plants. *Crit. Rev. Plant Sci.* 33, 92–124. doi: 10.1080/07352689.2014.870408

Wisniewski, M., Neuner, G., and Gusta, L. V. (2015a). The use of high-resolution infrared thermography (HRIT) for the study of ice nucleation and propagation in plants. *J. Vis. Exp.* 99:52703. doi: 10.3791/52703

Wisniewski, M., Norelli, J., and Artlip, T. (2015b). Overexpression of a peach CBF gene in apple: a model for understanding the integration of growth, dormancy, and cold hardiness in woody plants. *Front. Plant Sci.* 6:85. doi: 10.3389/fpls.2015.00085

Wisniewski, M., Norelli, J., Bassett, C., Artlip, T., and Macarisn, D. (2011). Ectopic expression of a novel peach (*Prunus persica*) CBF transcription actor in apple (*Malus x domestica*) results in short-day induced dormancy and increased cold hardiness. *Planta* 233, 971–983. doi: 10.1007/s00425-011-1358-3

Wu, R., Tomes, S., Karunairetnam, S., Tustin, S. D., Hellens, R., Allen, A., et al. (2017a). SVP-like MADS box genes control dormancy and budbreak in apple. *Front. Plant Sci.* 8:477. doi: 10.3389/fpls.2017.00477

Wu, R., Want, T., Warren, B., Allan, A., Macknight, R. C., and Varkonyi-Gasic, E. (2017b). Kiwifruit SVP2 gene prevents premature budbreak during dormancy. *J. Exp. Bot.* 68, 1071–1082. doi: 10.1093/jxb/erx014

Wu, R.-M., Walton, E., Richardsons, A. G., Wood, M., Hellens, R. P., and Varkonyi-Gasic, E. (2011). Conservation and divergence of four kiwifruit SVP-like MADS-box genes suggest distinct roles in kiwifruit bud dormancy and flowering. *J. Exp. Bot.* 63, 797–807. doi: 10.1093/jxb/err304

Xiao, H., Siddiqua, M., Braybrook, S., and Nassuth, A. (2006). Three grape CBF/DREB1 genes are regulated by low temperature, drought and abscisic acid. *Plant Cell Environ.* 29, 1410–1421. doi: 10.1111/j.1365-3040.2006.01524

Xiao, H., Tattersall, E., Siddiqua, M., Cramer, G. R., and Nassuth, A. (2008). CBF4 is a unique member of the CBF transcription factor family of Vitis vinifera and Vitis riparia. *Plant Cell Environ.* 31, 1–10. doi: 10.1111/j.1365-3040.2007.01741

Yang, T., Chaudhuri, S., Yang, L., Du, L., and Poovaiah, B. W. (2010a). A calcium/calmodulin-regulated member of the receptor-like kinase family confers cold tolerance in plants. *J. Biol. Chem.* 285, 7119–7126. doi: 10.1074/jbc.M109.035659

Yang, T., Shad Ali, G., Yang, L., Du, L., Reddy, A. S., and Poovaiah, B. W. (2010b). Calcium/calmodulin-regulated receptor-like kinase CRLK1 interacts with MEKK1 in plants. *Plant Signal. Behav.* 5, 991–994. doi: 10.4161/psb.5.8.12225

Yordanov, Y. S., Ma, C., Strauss, S. H., and Buscov, V. B. (2014). EARLY BUD-BREAK 1 (EBB1) is a regulator of release from seasonal dormancy in poplar trees. *Proc. Natl. Acad. Sci. U.S.A.* 111, 10001–10006. doi: 10.1073/pnas.1405621111

Yu, H., Luedeling, E., and Xu, J. (2010). Winter and spring warming result in delayed spring phenology on the Tibetan Plateau. *Proc. Natl. Acad. Sci. U.S.A.* 107, 22151–22156. doi: 10.1073/pnas.1012490107

Zhao, C., Wang, P., Si, T., Hsu, C.-C., Wang, L., Zayed, O., et al. (2017). MAP kinase cascades regulate the cold response by modulating ICE1 protein stability. *Dev. Cell* 43, 618–629. doi: 10.1016/j.devcel.2017.09.024

Zhao, C., Zhang, Z., Xie, S., Si, T., Li, Y., and Zhu, J. K. (2016). Mutational evidence for the critical role of CBF transcription factors in cold acclimation in Arabidopsis. *Plant Physiol.* 171, 2744–2759. doi: 10.1104/pp.16.00533

Zhao, K., Zhou, Y., Li, Y., Zhuo, X., Ahmad, S., Han, Y., et al. (2018). Crosstalk of PmCBFs and PmDAMs based on the changes of phytohormones under seasonal cold stress in the stem of *Prunus mume*. *Int. J. Mol. Sci.* 19:15. doi: 10.3390/ijms19020015

Zhou, M., Chen, H., Wei, D., Ma, H., and Lin, J. (2017). Arabidopsis CBF3 and DELLAs positively regulate each other in response to low temperature. *Sci. Rep.* 7:39819. doi: 10.1038/srep39819

Check for
updates

Nitrogen Metabolism and Biomass Production in Forest Trees

Francisco M. Cánovas*, Rafael A. Cañas, Fernando N. de la Torre, María Belén Pascual, Vanessa Castro-Rodríguez and Concepción Avila

Grupo de Biología Molecular y Biotecnología de Plantas, Departamento de Biología Molecular y Bioquímica, Universidad de Málaga, Málaga, Spain

Low nitrogen (N) availability is a major limiting factor for tree growth and development. N uptake, assimilation, storage and remobilization are key processes in the economy of this essential nutrient, and its efficient metabolic use largely determines vascular development, tree productivity and biomass production. Recently, advances have been made that improve our knowledge about the molecular regulation of acquisition, assimilation and internal recycling of N in forest trees. In poplar, a model tree widely used for molecular and functional studies, the biosynthesis of glutamine plays a central role in N metabolism, influencing multiple pathways both in primary and secondary metabolism. Moreover, the molecular regulation of glutamine biosynthesis is particularly relevant for accumulation of N reserves during dormancy and in N remobilization that takes place at the onset of the next growing season. The characterization of transgenic poplars overexpressing structural and regulatory genes involved in glutamine biosynthesis has provided insights into how glutamine metabolism may influence the N economy and biomass production in forest trees. Here, a general overview of this research topic is outlined, recent progress are analyzed and challenges for future research are discussed.

Keywords: Populus, nitrogen acquisition, nitrogen recycling, glutamine biosynthesis, transgenic trees

INTRODUCTION

Forest trees include a large group of gymnosperm and angiosperm species that play a crucial role in the overall balance of ecosystems. Forest species also have great economic importance in the production of wood, paper pulp, biofuels and a variety of resins and secondary metabolites. The biorefinery and nanotechnology of forest products are emerging areas of industrial interest in Europe within the so-called bioeconomy of the forestry sector (Horizons – Vision 2030 for the European Forest-based Sector[1]). Despite the relevance of forest species from the environmental, economic and social point of view, our knowledge of the mechanisms underlying forest growth, development, and productivity is still limited when compared to crop plants. However, recent developments in genomics and biotechnology are providing new tools to unravel key regulatory processes in fundamental tree biology (Tuskan et al., 2006; De la Torre et al., 2014; Plomion et al., 2016; Tsai et al., 2018).

A sustainable management of forest resources is needed to satisfy the increasing demand of forest-derived products and to preserve natural forest stands. For example, highly productive

---

[1]http://www.forestplatform.org/

plantations with increased levels of tree biomass production are necessary to meet the demands of second-generation bioenergy and other forest resources (Hinchee et al., 2009; Allwright and Taylor, 2016). These new forests will require a sustainable use of fertilizers with N as one of the most relevant components. N use efficiency (NUE) is defined in general terms as the amount of plant product per unit of N fertilizer supplied (Good et al., 2004), and trees with improved NUE will be required to enhance the yield of future plantations. Nitrogen acquisition and metabolism are therefore important targets to improve forest biomass production, and key genes involved in N acquisition from soil and assimilation into amino acids have been studied (Cánovas et al., 2007; de la Torre et al., 2014; Castro-Rodríguez et al., 2016a, 2017). In addition, processes of N storage and recycling are particularly relevant in forest species with long life cycles exhibiting seasonal periods of growth and development (Cantón et al., 2005; Minocha et al., 2015). In deciduous trees, most of the leaf N that is present in the stromal and thylakoidal proteins and chlorophylls of green plastids is allocated and stored in the stem during seasonal dormancy; these N reserves are rapidly mobilized to sustain metabolic activities in the next growing season (Babst and Coleman, 2018).

The genus *Populus* (poplars) include a variety of tree species with fast growth in temperate habitats. In fact, poplars are widely used for biomass production and considered to be one of the most important bioenergy crops (Ye et al., 2011; Allwright and Taylor, 2016). In addition, *Populus* has become a model tree due to its favorable characteristics for experimental analyses and the advances made during the last 15 years in its structural and functional genomics (Jansson and Douglas, 2007; Douglas, 2017). In this article, the functional genomics of N metabolism in *Populus* is reviewed. The relevance of these studies to enhance biomass production is highlighted. Finally, potential avenues for future research on this topic are discussed.

N ACQUISITION AND METABOLISM IN POPLAR

Poplars are able to acquire inorganic N forms from soil, such as NH_4^+ and NO_3^-, and its relative preference will depend on the soil pH (Rennenberg et al., 2010). The genome of *Populus trichocarpa* contains an abundant repertoire of genes encoding low- and high-affinity transporters involved in N uptake and allocation. NO_3^- uptake is mediated by a large family of transporters consisting of 68 *PtNPF* genes encoding nitrate and peptide transporters and a smaller family of 11 *PtNRT2/NRT3* genes also encoding nitrate transport systems (Bai et al., 2013). NH_4^+ transporters (AMT) are encoded by 22 genes distributed in two separate subfamilies (*AMT1* and *AMT2*) (Couturier et al., 2007; Wu et al., 2015; Calabrese et al., 2017). Interestingly, the number of AMT2 members in poplar is much higher than in *Arabidopsis,* suggesting differences in the way that these plants incorporate and transport NH_4^+ ions (Castro-Rodríguez et al., 2017; Gojon, 2017). The specific expression patterns of these genes strongly suggest that they play non-overlapping roles in N

acquisition and intercellular transport (Castro-Rodríguez et al., 2017).

Regardless of the source of inorganic N that is taken up by the roots, NH_4^+ is the ultimate N form to be assimilated into amino acids. The main step for N entry in plant metabolism is catalyzed by the enzyme glutamine synthetase (GS; EC 6.3.1.2) and it involves the ATP-dependent condensation of NH_4^+ and glutamate for the biosynthesis of glutamine. Unlike that found in other plant species, GS is encoded by a duplicated gene family in *Populus* consisting of 4 groups of genes, 3 of which code for GS isoforms of cytosolic localization (GS1.1, GS1.2, and GS1.3) and one group that codes for a plastid-located isoform (GS2). Duplicated *GS* genes display similar structures with well conserved regulatory regions and intron-exon boundaries. However, they are dispersed in the poplar genome and distributed on separate chromosomes (Castro-Rodríguez et al., 2011). Functional analyses of recombinant duplicates revealed that they exhibit similar molecular and kinetic properties and therefore are functionally equivalent enzymes (Castro-Rodríguez et al., 2015). The specific spatial and seasonal patterns of expression of GS genes support non-overlapping roles in poplar N metabolism (Castro-Rodríguez et al., 2011).

In situ hybridization and laser capture microdissection of the whole GS family disclosed that the expression of duplicates is confined to specific cell-types, confirming and extending previous findings (Castro-Rodríguez et al., 2015). Expression studies supported a relevant role of the GS1.1 isoforms in N metabolism of photosynthetic cells in coordination with the role of the chloroplast-located GS2, for example, in the reassimilation of NH_4^+ released during photorespiration (Betti et al., 2006). In contrast, the role of the GS1.2 enzymes could be related to N mobilization during seasonal N recycling at the onset of dormancy and with senescence associated with pathogen attack. Finally, the expression patterns of the GS1.3 isoforms are consistent with playing an essential role in the biosynthesis of glutamine for N transport (Cánovas et al., 2007) and reassimilation of NH_4^+ released in the metabolism of phenylalanine during wood formation (Craven-Bartle et al., 2013). Interestingly, genes encoding glutamate synthases (GOGAT, EC 1.4.7.1, and EC 1.4.1.14) and cytosolic isocitrate dehydrogenase (ICDH, EC 1.1.1.42) are also duplicated in poplar. The main players in N uptake and metabolism are shown in **Figure 1**. Gene IDs are listed in **Supplementary Table S1**.

Enzyme redundancy in glutamine biosynthesis occurs in particular cell types of poplar. Paralogous genes may have been retained in the poplar genome to increase the amount of enzyme because their expression is restricted to specific cell-types, and the accumulation of a GS isoform could contribute to maintaining the homeostasis of the N metabolism in a particular cell-type. Glutamine biosynthesis is at the crossroad of many metabolic pathways, and according to the above hypothesis, functions associated with glutamine-derived metabolic products would be enhanced in specialized tissues such as meristems, photosynthetic parenchyma, xylem, and phloem of vascular bundles. The availability of enhanced levels of organic N in the form of glutamine could have boosted the growth and vigor of

FIGURE 1 | Nitrogen transporters and enzymes involved in N acquisition and metabolism in poplar. **(A)** leaves, **(B)** stem, and **(C)** roots. AMT1, ammonium transporter 1; AMT2, ammonium transporter 2; NRT1, nitrate transporter 1 (NPF); NRT2, nitrate transporter 2; NRT3, nitrate transporter 3; GS1.1, cytosolic glutamine synthetase, isoform 1; GS1.2, cytosolic glutamine synthetase, isoform 2; GS1.3, cytosolic glutamine synthetase, isoform 3; GS2, glutamine synthetase, chloroplastic isoform; Fd-GOGAT, ferredoxin-dependent glutamate synthase; NADH-GOGAT, NADH-dependent glutamate synthase; ICDH, cytosolic isocitrate dehydrogenase; NiR, nitrite reductase; NR, nitrate reductase.

GENE FUNCTIONAL ANALYSIS IN TRANSGENIC TREES

Classical breeding has been widely used for tree improvement but new developments in genomics and biotechnology can accelerate the process. In regard to N nutrition, the above results point to GS as a key enzyme in N metabolism; however, it is important to elucidate which GS isoform contributes to a major extent to poplar growth and biomass production. Transgenic hybrid poplars (*P. tremula x P. alba*) overexpressing a cytosolic GS of pine exhibited enhanced growth and increased levels of proteins and chlorophylls (Fu et al., 2003; **Figure 2A**). Furthermore, the observed phenotype as a consequence of transgene expression was related to the correct assembly of GS1 subunits in the cytosol of photosynthetic cells. Further characterization showed that GS transgenics had a better NUE (Man et al., 2005; Castro-Rodríguez et al., 2017) and enhanced tolerance to abiotic stress (El-Khatib et al., 2004; Pascual et al., 2008; Shestibratov et al., 2010; Molina-Rueda and Kirby, 2015). A similar approach, overexpressing *GS1*, has also been used to improve NUE in birch species (Lebedev et al., 2017).

Enhanced growth of GS transgenics was associated with increases in the transcript and protein levels of anthranilate synthase, the enzyme catalyzing the biosynthesis of tryptophan, a precursor of auxin biosynthesis (Man et al., 2011). These findings highlight the paramount importance of GS1 in poplar growth and biomass production, which has also been demonstrated in herbaceous plant models such as maize (Hirel et al., 2007) and rice (Tabuchi et al., 2007). In contrast, chloroplastic GS plays a well-established role in reassimilation of NH_4^+ released in the photorespiratory pathway as determined by characterization of photorespiratory mutants lacking GS2 (Blackwell et al., 1987; Betti et al., 2006).

The ability of a regulatory gene to influence growth and biomass production has also been tested in poplar. Dof factors are regulators of N metabolism and potential targets to enhance N assimilation and plant growth (Rueda-López et al., 2008; Tsujimoto-Inui et al., 2009; Wang Y. et al., 2013). The transcription factor Dof5 that regulates GS1 isoforms in maritime pine (Rueda-López et al., 2008) was overexpressed in hybrid poplars (**Figure 2B**). In comparison to untransformed controls, young transgenic plants exhibited enhanced growth, an increased capacity for inorganic N uptake, and accumulated significantly more carbohydrates and lignin (Rueda-López et al., 2017).

The assimilation of NH_4^+ into amino acids by the GS/GOGAT pathway also requires the provision of carbon skeletons in the form of 2-oxoglutarate (2-OG) (Hodges, 2002). The role of ICDH, a key enzyme in the provision of 2-OG, has been investigated in hybrid poplar and overexpression of *ICDH* causes an alteration in vascular development (Pascual et al., 2018). Transgenic trees with higher levels of ICDH also displayed increased expression of *GS1.3* and other genes associated with vascular differentiation. Phenotypic characterization of the transgenic plants showed increased growth in height, longer internodes and enhanced development in young leaves and the apical region of the stem (**Figures 2C,D**). *ICDH* overexpression altered the contents of organic acids including citrate, malate and 2-OG, and the levels of glutamate and γ-aminobutyric acid. These results show that the provision of carbon skeletons for NH_4^+ assimilation and glutamine biosynthesis is a key metabolic process for growth and vascular development in poplar.

Field trials of genetically modified trees are extremely important to assess transgene behavior under natural conditions and potential risks of transgenes spreading before commercialization. A field trial of independently transformed lines expressing GS1 was established, and the performance of these transgenic lines was studied in natural conditions over 3 years (Jing et al., 2004; **Figure 2E**). The transgene was stably expressed in the field resulting in enhanced vegetative growth of transgenic poplars reaching average heights that were 41% greater than non-transformed controls (Jing et al., 2004; Cánovas et al., 2006). These results likely reflect a higher capacity of transgenic trees for N remobilization and N recycling, resulting in a better exploitation of nutrient resources. Interestingly, analysis of wood samples from these 3-year-old trees revealed alterations in cell wall characteristics resulting in improved attributes for pulp and paper production (Coleman et al., 2012). GS1 transgenics also tolerate high levels of NO_3^- supply, exhibiting greater NUE and accumulating increased biomass, particularly enriched in cellulose in the above-ground part of the plant (Castro-Rodríguez et al., 2016b). These results are consistent with an efficient N allocation and metabolism in the transgenics. Transcriptomic analysis revealed that transgenic trees are able to reprogramme the transcriptome in response to N excess by differential expression of a greater number of genes than untransformed plants. The above findings strongly support the potential use of these genetically modified trees for phytoremediation of NO_3^- pollution with enhanced production of biomass and cellulose for bioenergy applications.

The performance of transgenic trees overexpressing Dof5 was also studied in a field trial during two growing seasons (**Figure 2D**). Interestingly, these transgenic lines showed attenuated growth and no modification of carbon or N metabolism when growing under natural conditions. As the expression of the transgene was stable during the period of study the observed differences in the performance of transgenic trees were attributed to the low levels of N nutrients available in the soil (Rueda-López et al., 2017). These findings reinforce the importance of field studies and indicate that the manipulation of structural rather than regulatory genes has been more effective for increasing biomass production and forest productivity. **Figure 2** illustrates the phenotypes of genetically modified poplars growing under controlled and natural conditions.

PERSPECTIVES AND FUTURE DEVELOPMENTS

Recent advances made in model and crop plants highlight the importance of NO_3^- and NH_4^+ transporters as key
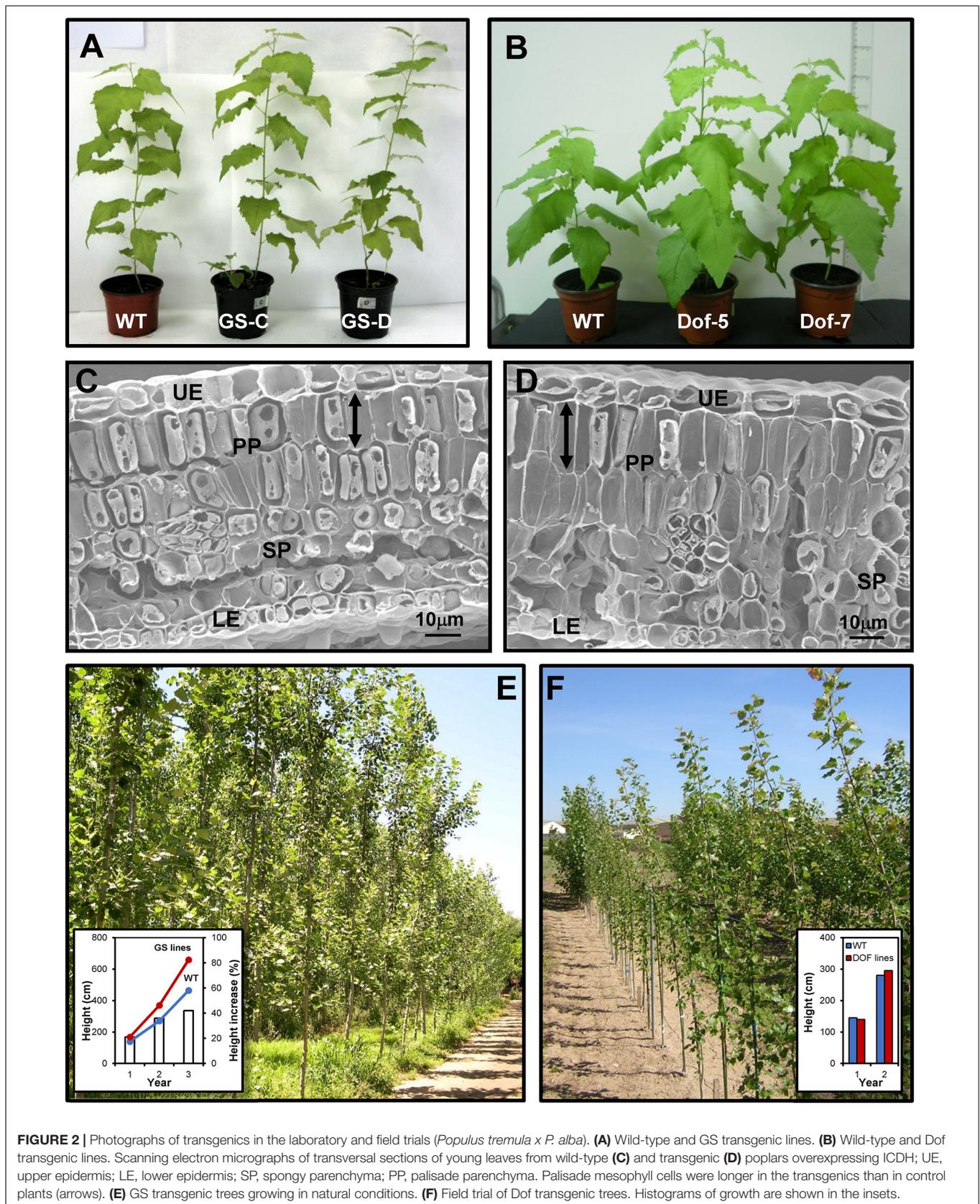
FIGURE 2 | Photographs of transgenics in the laboratory and field trials (*Populus tremula x P. alba*). **(A)** Wild-type and GS transgenic lines. **(B)** Wild-type and Dof transgenic lines. Scanning electron micrographs of transversal sections of young leaves from wild-type **(C)** and transgenic **(D)** poplars overexpressing ICDH; UE, upper epidermis; LE, lower epidermis; SP, spongy parenchyma; PP, palisade parenchyma. Palisade mesophyll cells were longer in the transgenics than in control plants (arrows). **(E)** GS transgenic trees growing in natural conditions. **(F)** Field trial of Dof transgenic trees. Histograms of growth are shown in the insets.

components of NUE. Consequently, manipulation of N acquisition and intercellular transport should be addressed to explore potential benefits in growth and productivity. In rice plants, overexpression of NO_3^- transporters such as OsNRT1.1B (Hu et al., 2015) and OsNRT2.3a (Fan et al., 2017) led to an accumulation of more biomass and increased yield. The identification of potential orthologs of these genes in poplar and subsequent functional analysis will deserve special attention to increasing tree productivity. Additional efforts are also needed to characterize amino acid transporters, particularly those involved in mechanisms of N allocation and recycling (Babst and Coleman, 2018). A strict coordination between N transporters and GS isoforms needs to exist to sustain the glutamine flux that is necessary for the biosynthesis of all nitrogenous compounds during poplar growth and development. A comparative analysis of gene expression in poplar showed co-expression profiles for several *AMT1, AMT2,* and *GS1* genes in young leaves, mature leaves and stems (Castro-Rodríguez et al., 2017). The coordinated function of N transporters and GS1 in different poplar tissues need to be investigated in future studies. Interestingly, Zhang et al. (2018) reported that under low nitrogen, the excess of carbon is redirected to the biosynthesis of aromatic amino acids and lignin, resulting in improved NUE that could be of practical value in terms of biomass production. Deciphering of regulatory networks involved in the response of roots to nitrogen availability is also deserving special attention (Wei et al., 2013; Dash et al., 2015; Luo et al., 2015).

As previously discussed, the manipulation of glutamine biosynthesis is a reasonable strategy to improve NUE and biomass production in poplar. According to the currently available data, the overproduction of GS1 isoforms in particular cell-types is largely beneficial; however, further studies are necessary to fully understand how the increase in glutamine biosynthesis influences tree growth and biomass production. For example, the specific contribution of the GS1 duplicates, GS1.1, GS1.2, and GS1.3, should be explored by performing functional studies using classic transformation strategies, or alternatively, by using the powerful CRISPR-Cas9 technology (Zhou et al., 2015). A recent study reported that overexpression of poplar *GS1.2* in tobacco altered secondary cell wall and fiber characteristics and accelerated auxin biosynthesis (Lu et al., 2018), but unfortunately, we still do not know what the impact of *GS1.2* manipulation in poplar may be. Of particular interest will be to specifically elucidate how the GS1.1 duplicates are associated with photosynthetic primary and/or secondary NH_4^+ assimilation in mature leaves. The impact of GS1 overproduction can be further studied with refined approaches using tissue-specific and inducible promoters (Filichkin et al., 2006; Wang L. et al., 2013).

Nevertheless, the existence of intrinsic regulatory mechanisms *in planta* cannot be ruled out. GS1 expression driven by the constitutive 35S promoter is potentially modulated in photosynthetic tissues through the interaction with a microRNA leading to improved biomass production (Fu et al., 2012).

Furthermore, the co-transformation of GS1 transgenics with cellulase genes driven by inducible promoters could facilitate processing of feedstocks for bioenergy applications. Another challenge for the future is the identification and molecular dissection of QTLs potentially associated with GS1/glutamine biosynthesis. QTLs for biomass production under nitrogen limitation and excess have been mapped in poplar but candidate genes of N metabolism were not identified (Novaes et al., 2009). In contrast, interactions have been shown between genes involved in glutamine and glutamate metabolism and QTLs associated to yield traits in maize (Hirel et al., 2007) and rice (Yamaya et al., 2002).

It is worth mentioning that the observed effects of GS1 transgene expression are explained by the altered expression of other genes involved in primary and secondary metabolism. Significant changes in the leaf transcriptome were observed when growing trees at high NO_3^- levels with a high number of genes differentially expressed, including those involved in photosynthesis, cell wall formation and phenylpropanoid biosynthesis (Castro-Rodríguez et al., 2016b). In turn, the upregulation of transcription factors strongly suggests that chromatin organization differs in transgenics and wild-type plants, particularly in the response of trees to N availability. Genome-wide identification of regions containing targeted genes involved in the nutritional responses can be achieved by conducting comparative ChiP-Seq analysis in poplar (Liu et al., 2015). A recent genomic resource will facilitate this task, the whole genome-assembly of the *P. tremula* x *P. alba* clone INRA 717-1B4 that is used as a tree model in transgenic experiments (Mader et al., 2016). New knowledge derived from these studies and/or those derived from the molecular dissection of QTLs will facilitate the identification of genes linking GS1/glutamine biosynthesis and biomass production. Using gene capture approaches (Seoane-Zonjic et al., 2016), structural variability in these genes can be analyzed in poplar genotypes with a contrasted ability for biomass production. In fact, a recent study has confirmed substantial structural variation in the poplar pan-genome (Pinosio et al., 2016).

CONCLUSION

A combination of functional genomic approaches, including transgenic and gene editing technology, chromatin analysis, and systematic identification of genome regions involved in NUE, will facilitate the exploration of the molecular basis of how N metabolism influences biomass production in forest trees.

AUTHOR CONTRIBUTIONS

FC and CA conceived and wrote the manuscript. RC, FdlT, MP, and VC-R made additional contributions and edited the manuscript. RC and VC-R composed the figures.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2018.01449/full#supplementary-material

TABLE S1 | Gene IDs and accession numbers for N transporters and enzymes involved in N acquisition and metabolism in poplar.

REFERENCES

Allwright, M. R., and Taylor, G. (2016). Molecular breeding for improved second generation bioenergy crops. *Trends Plant Sci.* 21, 43–54. doi: 10.1016/j.tplants.2015.10.002

Babst, B. A., and Coleman, G. D. (2018). Seasonal nitrogen cycling in temperate trees: transport and regulatory mechanisms are key missing links. *Plant Sci.* 270, 268–277. doi: 10.1016/j.plantsci.2018.02.021

Bai, H., Euring, D., Volmer, K., Janz, D., and Polle, A. (2013). The nitrate transporter (NRT) gene family in poplar. *PLoS One* 8:e72126. doi: 10.1371/journal.pone.0072126

Betti, M., Arcondéguy, T., and Márquez, A. J. (2006). Molecular analysis of two mutants from *Lotus japonicus* deficient in plastidic glutamine synthetase: functional properties of purified GLN2 enzymes. *Planta* 224, 1068–1079. doi: 10.1007/s00425-006-0279-z

Blackwell, R. D., Murray, A. J. S., and Lea, P. J. (1987). Inhibition of photosynthesis in barley with decreased levels of chloroplast glutamine synthetase activity. *J. Exp. Bot.* 38, 1799–1809. doi: 10.1093/jxb/38.11.1799

Calabrese, S., Kohler, A., Niehl, A., Veneault-Fourrey, C., Boller, T., and Courty, P. E. (2017). Transcriptome analysis of the *Populus trichocarpa-Rhizophagus irregularis* mycorrhizal symbiosis: regulation of plant and fungal transportomes under nitrogen starvation. *Plant Cell Physiol.* 58, 1003–1017. doi: 10.1093/pcp/pcx044

Cánovas, F. M., Ávila, C., Cantón, F. R., Cañas, R., and de la Torre, F. (2007). Ammonium assimilation and amino acid metabolism in conifers. *J. Exp. Bot.* 58, 2307–2318. doi: 10.1093/jxb/erm051

Cánovas, F. M., Gallardo, F., Jing, Z. P., and Pascual, M. B. (2006). "Transgenic approaches to engineer nitrogen metabolism," in *Tree Transgenesis*, eds M. Fladung, and D. Ewald (Berlin: Springer), 157–178. doi: 10.1007/3-540-32199-3_8

Cantón, F. R., Suárez, M. F., and Cánovas, F. M. (2005). Molecular aspects of nitrogen mobilisation and recycling in trees. *Photosynth. Res.* 83, 265–278. doi: 10.1007/s11120-004-9366-9

Castro-Rodríguez, V., Assaf-Casals, I., López-Tienda, J., Ávila, C., Miller, A. J., and Cánovas, F. M. (2016a). Deciphering the molecular basis of ammonium uptake and transport in maritime pine. *Plant Cell Environ.* 39, 1669–1682. doi: 10.1111/pce.12692

Castro-Rodríguez, V., Cañas, R. A., de la Torre, F., Pascual, M. B., Ávila, C., and Cánovas, F. M. (2017). Molecular fundamentals of nitrogen uptake and transport in trees. *J. Exp. Bot.* 68, 2489–2500. doi: 10.1093/jxb/erx037

Castro-Rodríguez, V., García-Gutiérrez, A., Canales, J., Ávila, C., Kirby, E. G., and Cánovas, F. M. (2011). The glutamine synthetase gene family in *Populus*. *BMC Plant Biol.* 11:119. doi: 10.1186/1471-2229-11-119

Castro-Rodríguez, V., García-Gutiérrez, A., Canales, J., Cañas, R. A., Kirby, E. G., Ávila, C., et al. (2016b). Poplar trees for phytoremediation of high levels of nitrate and applications in bioenergy. *Plant Biotech. J.* 14, 299–312. doi: 10.1111/pbi.12384

Castro-Rodríguez, V., García-Gutiérrez, A., Cañas, R. A., Pascual, M. B., Ávila, C., and Cánovas, F. M. (2015). Redundancy and metabolic regulation of the glutamine synthetase gene family in poplar. *BMC Plant Biol.* 15:20. doi: 10.1186/s12870-014-0365-5

Coleman, H. D., Mansfield, S. D., Cánovas, F. M., Huimin Man, H., and Kirby, E. G. (2012). Enhanced expression of glutamine synthetase (GS1a) confers altered fiber and wood chemistry in field grown hybrid poplar (*Populus tremula x P. alba*) (717-1B4). *Plant Biotech. J.* 10, 883–889. doi: 10.1111/j.1467-7652.2012.00714.x

Couturier, J., Montanini, B., Martin, F., Brun, A., Blaudez, D., and Chalot, M. (2007). The expanded family of ammonium transporters in the perennial poplar plant. *New Phytol.* 174, 137–150. doi: 10.1111/j.1469-8137.2007.01992.x

Craven-Bartle, B., Pascual, M. B., Cánovas, F. M., and Ávila, C. (2013). A Myb transcription factor regulates genes of the phenylalanine pathway in maritime pine. *Plant J.* 74, 755–766. doi: 10.1111/tpj.12158

Dash, M., Yordanov, Y. S., Georgieva, T., Kumari, S., Wei, H., and Busov, V. (2015). A systems biology approach identifies new regulators of poplar root development under low nitrogen. *Plant J.* 84, 335–346. doi: 10.1111/tpj.13002

De la Torre, A. R., Birol, I., Bousquet, J., Ingvarsson, P. K., Jansson, S., Jones, S. J. M., et al. (2014). Insights into conifer giga-genomes. *Plant Physiol.* 166, 1724–1732. doi: 10.1104/pp.114.248708

de la Torre, F., Cañas, R. A., Pascual, M. B., Ávila, C., and Cánovas, F. M. (2014). Plastidic aspartate aminotransferases and the biosynthesis of essential amino acids. *J. Exp. Bot.* 65, 5527–5534. doi: 10.1093/jxb/eru240

Douglas, C. J. (2017). "*Populus* as a model tree," in *Comparative and Evolutionary Genomics of Angiosperm Trees. Plant Genetics and Genomics: Crops and Models*, eds A. T. Groover and Q. C. B. Cronk (New York, NY: Springer International Publishing), 61–84. doi: 10.1007/7397_2017_3

El-Khatib, R. T., Hamerlynck, E. P., Gallardo, F., and Kirby, E. G. (2004). Transgenic poplar characterized by ectopic expression of a pine cytosolic glutamine synthetase gene exhibits enhanced tolerance to water stress. *Tree Physiol.* 24, 729–736. doi: 10.1093/treephys/24.7.729

Fan, X., Naz, M., Fan, X., Xuan, W., Miller, A., and Xu, G. (2017). Plant nitrate transporters: from gene function to application. *J. Exp. Bot.* 68, 2463–2475. doi: 10.1093/jxb/erx011

Filichkin, S. A., Meilan, R., Busov, V. B., Ma, C., Brunner, A. M., and Strauss, A. M. (2006). Alcohol-inducible gene expression in transgenic *Populus*. *Plant Cell Rep.* 25, 660–667. doi: 10.1007/s00299-005-0112-3

Fu, C., Sunkar, R., Zhou, C., Shen, H., Zhang, J. Y., Matts, J., et al. (2012). Overexpression of miR156 in switchgrass (*Panicum virgatum* L.) results in various morphological alterations and leads to improved biomass production. *Plant Biotech. J.* 10, 443–452. doi: 10.1111/j.1467-7652.2011.00677.x

Fu, J., Sampalo, R., Gallardo, F., Cánovas, F. M., and Kirby, E. G. (2003). Assembly of a cytosolic pine glutamine synthetase holoenzyme in the leaf of transgenic poplar leads to enhanced vegetative growth in young plants. *Plant Cell Environ.* 26, 411–418. doi: 10.1046/j.1365-3040.2003.00972.x

Gojon, A. (2017). Nitrogen nutrition in plants: rapid progress and new challenges. *J. Exp. Bot.* 68, 2457–2462. doi: 10.1093/jxb/erx171

Good, A. G., Shrwat, A. K., and Muench, D. G. (2004). Can less yield more? Is reducing nutrient input into the environment compatible with maintaining

crop production. *Trends Plant Sci.* 9, 597–605. doi: 10.1016/j.tplants.2004. 10.008

Hinchee, M., Rottmann, W., Mullinax, L., Zhang, C., Chang, S., Cunningham, M., et al. (2009). Short-rotation woody crops for bioenergy and biofuels applications. *In Vitro Cell. Dev. Biol. Plant* 45, 619–629. doi: 10.1007/s11627-009-9235-5

Hirel, B., Le Gouis, J., Ney, B., and Gallais, A. (2007). The challenge of improving nitrogen use efficiency in crop plants: towards a more central role for genetic variability and quantitative genetics within integrated approaches. *J. Exp. Bot.* 58, 2369–2387. doi: 10.1093/jxb/erm097

Hodges, M. (2002). Enzyme redundancy and the importance of 2-oxoglutarate in plant ammonium assimilation. *J. Exp. Bot.* 53, 905–916. doi: 10.1093/jexbot/53. 370.905

Hu, B., Wang, W., Ou, S., Tang, J., Li, H., Che, R., et al. (2015). Variation in NRT1.1B contributes to nitrate-use divergence between rice subspecies. *Nat. Genet.* 47, 834–838. doi: 10.1038/ng.3337

Jansson, S., and Douglas, C. J. (2007). *Populus*: a model system for plant biology. *Annu. Rev. Plant Biol.* 58, 435–458. doi: 10.1146/annurev.arplant.58.032806. 103956

Jing, Z. P., Gallardo, F., Pascual, M. B., Sampalo, R., Romero, J., Torres de Navarra, A., et al. (2004). Improved growth in a field trial of transgenic hybrid poplar overexpressing glutamine synthetase. *New Phytol.* 164, 137–145. doi: 10.1111/j.1469-8137.2004.01173.x

Lebedev, V. G., Kovalenko, N. P., and Shestibratov, K. A. (2017). Influence of nitrogen availability on growth of two transgenic birch species carrying the pine *GS1a* gene. *Plants* 6:E4. doi: 10.3390/plants6010004

Liu, L., Ramsay, T., Zinkgraf, M., Sundell, D., Street, N. B., Fikov, V., et al. (2015). A resource for characterizing genome-wide binding and putative target genes of transcription factors expressed during secondary growth and wood formation in *Populus*. *Plant J.* 82, 887–898. doi: 10.1111/tpj. 12850

Lu, T., Liu, L., Wei, M., Liu, Y., Qu, Z., Yang, C., et al. (2018). The effect of poplar *PsnGS1.2* overexpression on growth, secondary cell wall, and fiber characteristics in tobacco. *Front. Plant Sci.* 9:9. doi: 10.3389/fpls.2018. 00009

Luo, J., Zhou, J., Li, H., Shi, W., Polle, A., Lu, M., et al. (2015). Global poplar root and leaf transcriptomes reveal links between growth and stress responses under nitrogen starvation and excess. *Tree Physiol.* 35, 1283–1302. doi: 10.1093/ treephys/tpv091

Mader, M., Le Paslier, M. C., Bounon, R., Bérard, A., Rampant, P. F., Fladung, M., et al. (2016). Whole-genome draft assembly of *Populus tremula* x *P. alba* clone INRA 717-1B4. *Silvae Genet.* 65, 74–79. doi: 10.1515/sg-2016-0019

Man, H., Pollmann, S., Weiler, E. W., and Kirby, E. G. (2011). Increased glutamine in leaves of poplar transgenic with pine GS1a caused greater anthranilate synthase α-subunit (ASA1) transcript and protein abundances: an auxin-related mechanism for enhanced growth in GS transgenics? *J. Exp. Bot.* 62, 4423–4431. doi: 10.1093/jxb/err026

Man, H. M., Boriel, R., El-Khatib, R., and Kirby, E. G. (2005). Characterization of transgenic poplar with ectopic expression of pine cytosolic glutamine synthetase under conditions of varying nitrogen availability. *New Phytol.* 167, 31–39. doi: 10.1111/j.1469-8137.2005. 01461.x

Minocha, R., Turlapati, S. A., Long, S., McDowell, W. H., and Minocha, S. (2015). Long-term trends of changes in pine and oak foliar nitrogen metabolism in response to chronic nitrogen amendments at Harvard Forest, MA. *Tree Physiol.* 35, 894–909. doi: 10.1093/treephys/tpv044

Molina-Rueda, J. J., and Kirby, E. G. (2015). Transgenic poplar expressing the pine GS1a show alterations in nitrogen homeostasis during drought. *Plant Physiol. Biochem.* 94, 181–190. doi: 10.1016/j.plaphy.2015. 06.009

Novaes, E., Osorio, L., Drost, D. R., Mile, B. L., Boaventura-Novaes, C. R. D., Benedict, C., et al. (2009). Quantitative genetic analysis of biomass and wood chemistry of *Populus* under different nitrogen levels. *New Phytol.* 182, 878–890. doi: 10.1111/j.1469-8137.2009.02785.x

Pascual, M. B., Jing, Z. P., Kirby, E. G., Cánovas, F. M., and Gallardo, F. (2008). Response of transgenic poplar overexpressing cytosolic glutamine synthetase to

phosphinothricin. *Phytochemistry* 69, 382–389. doi: 10.1016/j.phytochem.2007. 07.031

Pascual, M. B., Molina-Rueda, J. J., Cánovas, F. M., and Gallardo, F. (2018). Overexpression of a cytosolic NADP + -isocitrate dehydrogenase causes alterations in the vascular development of hybrid poplar. *Tree Physiol.* 38, 992–1005. doi: 10.1093/treephys/tpy044

Pinosio, S., Giacomello, S., Faivre-Rampant, P., Taylor, G., Jorge, V., Le Paslier, M. C., et al. (2016). Characterization of the poplar pan-genome by genom-wide identification of structural variation. *Mol. Biol. Evol.* 33, 2706–2719. doi: 10.1093/molbev/msw161

Plomion, C., Bastien, C., Bogeat-Triboulot, M. B., Bouffier, L., Déjardin, A., Duplessis, S., et al. (2016). Forest tree genomics: 10 achievements from the past 10 years and future prospects. *Ann. For. Sci.* 73, 77–103. doi: 10.1007/s13595-015-0488-3

Rennenberg, H., Wildhagen, H., and Ehlting, B. (2010). Nitrogen nutrition of poplar trees. *Plant Biol.* 12, 275–291. doi: 10.1111/j.1438-8677.2009. 00309.x

Rueda-López, M., Crespillo, R., Cánovas, F. M., and Ávila, C. (2008). Differential regulation of two glutamine synthetase genes by a single Dof transcription factor. *Plant J.* 56, 73–85. doi: 10.1111/j.1365-313X.2008. 03573.x

Rueda-López, M., Pascual, M. B., Pallero, M., Lasa, B., Jauregui, I., Aparicio-Tejo, P., et al. (2017). Overexpression of a pine Dof transcription factor in hybrid poplars: a comparative study in trees growing under controlled and natural conditions. *PLoS One* 12:e0174748. doi: 10.1371/journal.pone. 0174748

Seoane-Zonjic, P., Cañas, R., Bautista, R., Gómez-Maldonado, J., Arrillaga, I., Fernández-Pozo, N., et al. (2016). Establishing gene models from the *Pinus pinaster* genome using gene capture and BAC sequencing. *BMC Genomics* 17:148. doi: 10.1186/s12864-016-2490-z

Shestibratov, K. A., Bulatova, I. V., and Novikov, P. S. (2010). In vitro response of transgenic aspen containing glutamine synthetase gene GS1 to the sublethal dose of phosphinothricin. *Appl. Biochem. Microbiol.* 46, 763–768. doi: 10.1134/ S0003683810080053

Tabuchi, M., Abiko, T., and Yamaya, T. (2007). Assimilation of ammonium ions and reutilization of nitrogen in rice (*Oryza sativa* L.). *J. Exp. Bot.* 58, 2319–2327. doi: 10.1093/jxb/erm016

Tsai, C. J., Harding, S. A., and Cooke, J. E. K. (2018). Branching out: a new era of investigating physiological processes in forest trees using genomic tools. *Tree Physiol.* 38, 303–310. doi: 10.1093/treephys/ tpy026

Tsujimoto-Inui, Y., Naito, Y., Sakurai, N., Suzuki, H., Sasaki, R., Takahashi, H., et al. (2009). Functional genomics of the Dof transcription factorfamily genes in suspension-cultured cells of *Arabidopsis thaliana*. *Plant Biotechnol.* 26, 15–28. doi: 10.5511/plantbiotechnology.26.15

Tuskan, G. A., DiFazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313, 1596–1604. doi: 10.1126/science.112 8691

Wang, L., Li, L., Xu, L., Zhou, J., Zhuang, H., Gong, X., et al. (2013). Isolation and functional analaysis of the poplar *RbcS* gene promoter. *Plant Mol. Biol. Rep.* 31, 120–127. doi: 10.1007/s11105-012-0482-y

Wang, Y., Fu, B., Pan, L., Chen, L., Fu, X., and Li, K. (2013). Overexpression of *Arabidopsis* Dof1, GS1 and GS2 enhanced nitrogen assimilation in transgenic tobacco grown under low-nitrogen conditions. *Plant Mol. Biol. Rep.* 31, 886–900. doi: 10.1007/s11105-013-0561-8

Wei, H., Yordanov, Y. S., Kumari, S., Georgieva, T., and Busov, V. (2013). Genetic networks involved in poplar root response to low nitrogen. *Plant Signal. Behav.* 8, e27211. doi: 10.4161/psb.27211

Wu, X., Yang, H., Qu, C., Xu, Z., Li, W., Hao, B., et al. (2015). Sequence and expression analysis of the AMT gene family in poplar. *Front. Plant Sci.* 6:337. doi: 10.3389/fpls.2015.00337

Yamaya, T., Obara, M., Nakajima, H., Sasaki, S., Hayakawa, T., and Sato, T. (2002). Genetic manipulation and quantitative-trait loci mapping for nitrogen recycling in rice. *J. Exp. Bot.* 53, 917–925. doi: 10.1093/jexbot/53. 370.917

Ye, X., Busov, V., Zhao, N., Meilan, R., McDonnell, L. M., Coleman, H. D., et al. (2011). Transgenic *Populus* trees for forest products, bioenergy, and functional genomics. *Crit. Rev. Plant Sci.* 30, 415–434. doi: 10.1080/07352689.2011.605737

Zhang, X., Misra, A., Nargund, S., Coleman, G. D., and Sriram, G. (2018). Concurrent isotope-assisted metabolic flux analysis and transcriptome profiling reveal responses of poplar cells to altered nitrogen and carbon supply. *Plant J.* 3, 472–488. doi: 10.1111/tpj.13792

Zhou, X., Jacobs, T. B., Xue, L. J., Harding, S. A., and Tsai, C. J. (2015). Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial *Populus* reveals 4-coumarate:CoA ligase specificity and redundancy. *New Phytol.* 208, 298–301. doi: 10.1111/nph.13470

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Engineering Drought Resistance in Forest Trees

Andrea Polle [1,2,3*], Shao Liang Chen [1], Christian Eckert [2] and Antoine Harfouche [4]

[1] Beijing Advanced Innovation Center for Tree Breeding by Molecular Design, College of Biological Sciences and Technology, Beijing Forestry University, Beijing, China, [2] Forest Botany and Tree Physiology, University of Goettingen, Göttingen, Germany, [3] Centre of Biodiversity and Sustainable Land Use, University of Goettingen, Göttingen, Germany, [4] Department for Innovation in Biological, Agrofood and Forest systems, University of Tuscia, Viterbo, Italy

Climatic stresses limit plant growth and productivity. In the past decade, tree improvement programs were mainly focused on yield but it is obvious that enhanced stress resistance is also required. In this review we highlight important drought avoidance and tolerance mechanisms in forest trees. Genomes of economically important trees species with divergent resistance mechanisms can now be exploited to uncover the mechanistic basis of long-term drought adaptation at the whole plant level. Molecular tree physiology indicates that osmotic adjustment, antioxidative defense and increased water use efficiency are important targets for enhanced drought tolerance at the cellular and tissue level. Recent biotechnological approaches focused on overexpression of genes involved in stress sensing and signaling, such as the abscisic acid core pathway, and down-stream transcription factors. By this strategy, a suite of defense systems was recruited, generally enhancing drought and salt stress tolerance under laboratory conditions. However, field studies are still scarce. Under field conditions trees are exposed to combinations of stresses that vary in duration and magnitude. Variable stresses may overrule the positive effect achieved by engineering an individual defense pathway. To assess the usability of distinct modifications, large-scale experimental field studies in different environments are necessary. To optimize the balance between growth and defense, the use of stress-inducible promoters may be useful. Future improvement programs for drought resistance will benefit from a better understanding of the intricate networks that ameliorate molecular and ecological traits of forest trees.

Keywords: water limitation, antioxidative systems, genetic engineering, forest tree species, isohydric, anisohydric, avoidance, tolerance

INTRODUCTION

Forests cover about 30% of the terrestrial land (FAO, 2016). They have strong effects on the local climate (Li et al., 2015), by interacting with biogeochemical water cycles (Ellison et al., 2017). When forest trees die or forests are cleared across large-scale landscapes, the negative consequences of drought are aggravated (Allen et al., 2015; Reyer et al., 2015), as shown for many areas world-wide (Laurance, 1998, 2004; van der Werf et al., 2008; Malone, 2017). Over-utilization of forests as a feedstock for energy, construction materials, or the generation of value-added products for the chemical industries, intensifies the problem.

The negative consequences of drought become even more urgent in current times of climate change because projections suggest that such events will occur more frequently and be more

extreme (Allen et al., 2010; Reyer et al., 2015). In the past decades, global warming has resulted in a drastic reduction of ice-covered northern polar areas during summer (NSIDC, 2018). Over smaller polar areas air temperatures are cooling down less, thus, resulting in lower differences between boreal, temperate and tropical areas. A possible climate implication of this atmospheric situation is an effect on jet-stream oscillation, which in turn may extend stable meteorological high- and low-pressure (anticyclone/cyclone) conditions; the consequences of such conditions are manifested in periods of precipitation on the one hand, and periods of drought on the other hand (Schaller et al., 2018). During lows, flooding events are frequent, whereas the long-lasting highs lead to scarcity of water in many regions world-wide (FAO, 2016). The dry spells promote salt accumulation in upper soil layers, soil degradation, and erosion (Polle and Chen, 2015). Salt and drought are, thus, often co-occurring stresses with which plants have to cope although their physiological implications vary to some extent (Chen and Polle, 2010; Polle and Chen, 2015). The current review is focused on tree responses and improvement by genetic engineering in response to drought. Since most studies in which trees were ameliorated for improved stress resistance included both drought and salt, salinity cannot be completely ignored.

In this review we highlight the molecular physiology of drought stressed forest trees and present an overview on recent biotechnological approaches to improve the drought tolerance of trees with a focus on yield and enhanced stress resistance. Drought effects on woody plants and measures for tree improvements have regularly been reviewed (Wang et al., 2003; Polle et al., 2006; Rennenberg et al., 2006; Fischer and Polle, 2010; Harfouche et al., 2014). Therefore, this review briefly recapitulates the molecular physiology of drought and salt tolerance mechanisms. We summarize novel studies, published in the past 5 years on the performance of trees engineered for better osmotic resistance. We also pinpoint research gaps that need to be addressed for future improvement of drought resistance in trees.

Concepts and Strategies

Growth and reproduction of plants requires access to water. Water is the solvent for nutrients in soil, the transport medium for nutrients in the plants and the solvent for cellular solutes. Because water is essentially the "stuff of life", the plant water status is tightly controlled by a multitude of general and specific measures such as stomatal control on water loss (Buckley, 2005; Daszkowska-Golec and Szarejko, 2013), osmotic adjustment (Harfouche et al., 2014), anatomical adjustment of the water conducting system (Sperry and Love, 2015; Leuschner and Meier, 2018), deposition of cuticular waxes (Hadley and Smith, 1990) and morphological adjustments such as leaf shedding to avoid uncontrolled desiccation (Munné-Bosch and Alegre, 2004; Fischer and Polle, 2010). Periods of severe and long-lasting drought threaten the existence of plants when overruling their acclimation capacities. These broad examples show that drought responses act at different scales, i.e., inside the plant body and at the level of populations; the responses occur at different time scales, and thus, invoke short- and long-term adjustments that

can be flexible or reflect evolutionary adaptation (**Figure 1**). As a result, drought resistance can be achieved by avoidance (homeostasis of tissue water status) or by tolerance mechanisms (acclimation that enable metabolism at a low water potential) (Levitt, 1980; Jones, 1993). These distinctions are important when considering strategies for engineering drought resistance in tree species.

Avoidance mechanisms generally act at the scale of organs or at the whole-plant and the species level (**Figure 1**). Drought adaptation is characterized by ecological traits such as leaf thickness, root morphology, leaf shedding, etc. It is obvious that these traits are the result of a distinct developmental repertoire in a given species. Due to their presumed complexity they have rarely been incorporated in molecular breeding programs. However, the availability of increasing number of tree reference genomes may open new avenues to better understanding and exploiting their ecological traits. For example, the genomes of European beech (*Fagus sylvatica*) and pedunculate oak (*Quercus robur*) have recently been published (Mishra et al., 2018; Plomion et al., 2018). These two species are closely related members of the *Fagaceae* but exhibit strongly divergent ecological behavior (Aranda et al., 2015; Roman et al., 2015). Beech has a shallow fine root system, while oak has a deep root system (Leuschner and Meier, 2018). Deeper tree roots make a valuable contribution to drought resistance and thus, root morphology is one of the traits targeted for improving water use by capturing subsoil water. At the whole-plant level drought stress avoidance is dependent on the capability of the tree to minimize loss and maximize uptake of water (Chaves et al., 2003) through stomatal control and extensive, deep root systems (Nguyen and Lamant, 1989; Brodribb et al., 2010). We envisage that exploiting genomic information, for instance by comparing the molecular differences in root development of important forest species such as beech and oak, novel approaches that could direct breeding for drought avoidance may become available.

FIGURE 1 | Schematic overview of the different scales of drought impact from the cellular to the population level. Blue arrows: full line (scale of molecular and physiological traits), broken line (scale of ecological traits).
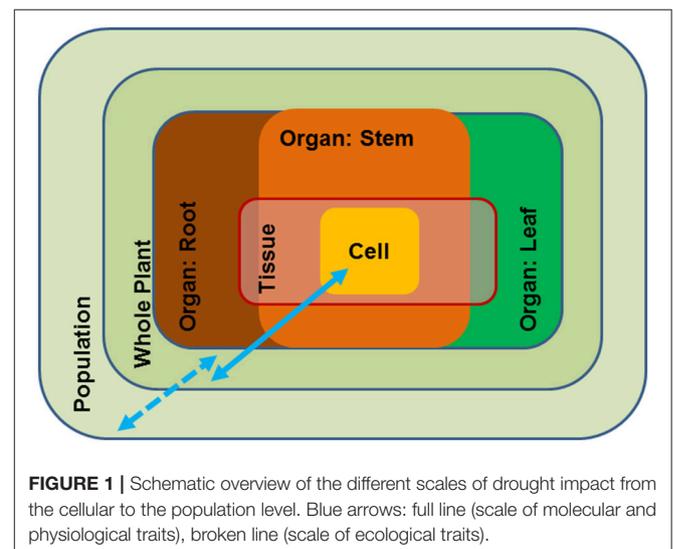
Another interesting example for drought avoidance is leaf shedding, a common phenomenon in tropical dry forests (Wolfe et al., 2016). Leaf shedding is controlled by an intricate interplay of phytohormones, including ethylene, abscisic acid (ABA), and auxin (González-Carranza et al., 1998; Chen et al., 2002a,b; Jin et al., 2015; Paul et al., 2018), which could be harnessed to improve tree drought resistance. In polyploid poplars regenerated from protoplast fusion accelerated drought-induced leaf shedding was observed that resulted in increased tree survival under extreme drought (Hennig et al., 2015). The exact genetic basis of this phenotype is not known but it is apparently associated with partial genome duplication (Hennig, 2016).

Trees must exist in their environment over decades and centuries and therefore require not only drought adaptedness but also metabolic flexibility to adjust their metabolism to changing conditions. Drought tolerance is usually achieved by biochemical modification of the cellular metabolism (**Figure 1**). Acclimation to drought by an individual plant invokes changes in membrane composition, protection of protein folding, osmotic adjustment, scavenging of reactive oxygen species (ROS), etc. (Harfouche et al., 2014) and, thus, acts at the level of cells to organs (**Figure 1**). An important feature of plant drought tolerance is the increase in osmotic pressure as a countermeasure to maintain water flux under declining soil water potentials. The production of osmolytes is costly in terms of carbohydrates because it diverts carbohydrates from growth to defense. A striking example of how woody species from arid, saline deserts can economize their carbon budget is the succulent xerophyte *Zygophyllum xanthoxylum* (Janz and Polle, 2012). This species exploits sodium as a "cheap" osmolyte, thereby, improving photosynthesis and growth under harsh environmental conditions (Ma et al., 2012). The discovery of such amelioration mechanisms constitutes an important basis to improve drought tolerance in trees (Bao et al., 2015, more details are found below) and underpins our understanding of the physiological consequences of novel features, which is crucial to harness the critical molecular mechanisms for drought acclimation and adaptation.

MOLECULAR PHYSIOLOGY OF OSMOTIC STRESSES

Roots

Roots are the first organ to sense and signal soil water deficits (Hamanishi and Campbell, 2011; Brunner et al., 2015). Since enhanced salinity decreases water availability to roots by increasing the osmotic potential in soil solution, the consequences for water uptake are partly similar to those of drought. Both drought and salt result in a decline in root-to-shoot water flow in poplars (Chen et al., 1997, 2002b; Shi et al., 2010), but the consequences are generally less severe in salt tolerant than in sensitive species (Chen et al., 2002b,c, 2003).

At the biochemical level, increased ABA concentrations are a hallmark of osmotic stress across all organs (Wasilewska et al., 2008; Kuromori et al., 2018) (**Figure 2**). The stress signal ABA interacts with pyrabactin resistance 1 (PYR1)/PYR1-like (PYL)/regulatory components of ABA receptors (RCAR)

proteins, which then can then form a complex with PP2Cs (Type 2C phosphatases). Thereby, phosphorylation of a SnRK (Kinase) is enabled, which subsequently activates down-stream transcription factors and target genes (Fujita et al., 2011; de Zelicourt et al., 2016). Transcriptomic analyses of pine and poplar roots under drought revealed upregulation of genes for ABA biosynthesis [9-cis-epoxycarotenoid dioxygenase (*NCED*)], signaling and response factors such as *DREB1*, *bZIP*, *AP2/ERF*, *MYB*, *NAC*, and *WRKY* (Wilkins et al., 2009; Cohen et al., 2010; Lorenz et al., 2011; Perdiguero et al., 2012). Salinity and drought share similar response patterns in poplar roots, which are likely mediated by ABA (Chen et al., 1997, 2001, 2002b; Luo et al., 2009).

ABA in roots further promotes increased biosynthesis of proline (Davies and Bacon, 2003). High concentrations of proline may act as osmolyte and contribute to osmotic adjustment; a further possible function of proline is the protection of appropriate protein folding (Claeys and Inzé, 2013). However, only a few studies have demonstrated increases in proline concentrations in the roots of trees in response to drought stress (Cocozza et al., 2010; Naser et al., 2010). This casts doubts on a prominent role of proline as an osmolyte in roots. Another function of ABA biosynthesis and transport in roots may be related to stimulate endodermal suberization (Tan et al., 2003; Koiwai et al., 2004; Kuromori et al., 2010; Kanno et al., 2012; Zhang et al., 2014). A recent study demonstrated that endodermis is reversibly impregnated with hydrophobic compounds such as suberin in response to abiotic stresses, which might restrict apoplastic movement of water (Barberon et al., 2016).

Evidence is accumulating that ABA plays a role in regulating hydraulic conductance in roots possibly via aquaporin activity (Parent et al., 2009; Almeida-Rodriguez et al., 2011). At the molecular level, aquaporins are important for the control of water

FIGURE 2 | Molecular and physiological key responses of trees to drought stress.
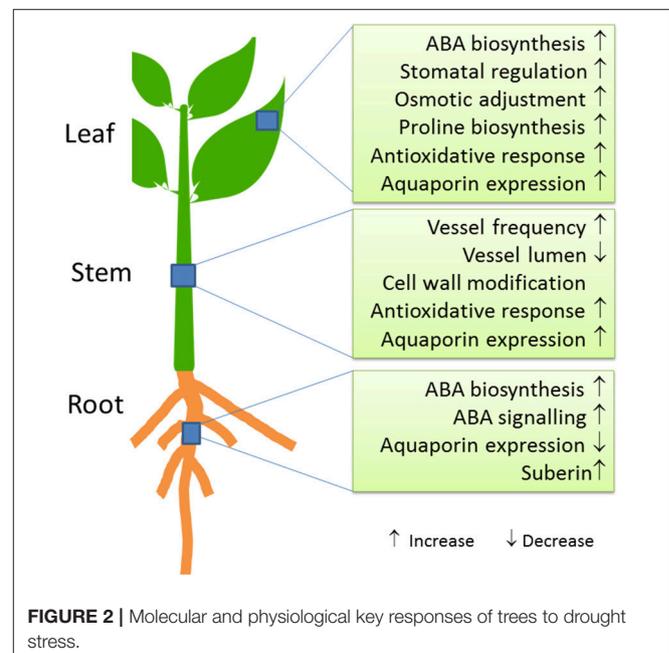
uptake (Fox et al., 2017). Since contrasting responses to different osmotic stress factors have been observed, their regulation is still unclear. For example, changes in the evaporative demand resulted in aquaporin upregulation (Parent et al., 2009; Almeida-Rodriguez et al., 2011), while drought or salt stress caused declines in their expression (Bogeat-Triboulot et al., 2007; Wang et al., 2017a). A decrease in aquaporins in roots is supposed to increase cellular water conservation due to reduced membrane water permeability during periods of dehydration stress (Smart et al., 2001; Bogeat-Triboulot et al., 2007) and would fit with the observed restriction of apoplastic water loss by enhanced suberization (Barberon et al., 2016).

Mycorrhizal fungi also increase tree stress tolerance by regulation of aquaporins and stress metabolites (Luo et al., 2009; Dietz et al., 2011; Xu et al., 2015; Peter et al., 2016), but an in-depth treatment of this aspect is beyond the scope of this review. The reader is advised to consult Brunner et al. (2015).

Stem

Physiological responses of trees to drought stress lead to hydraulic and carbon cycle adjustments (Parker, 1956; Bréda et al., 2006). The hydraulic architecture of the stem is important to keep up water transport under drought and re-establish water flux after re-irrigation. Hydraulic acclimation can be achieved by increased vessel frequencies and decreased vessel lumina (Hacke et al., 2006). Drought-resistant trees decrease the ratio of vessel lumen to cell wall thickness to enhance wall strength under water stress (Hacke et al., 2001) (**Figure 2**). While the anatomy and biophysics of xylem adjustment to water-limited conditions has often been studied (Tyree and Ewers, 1991; Anderegg, 2015; Sperry and Love, 2015), our knowledge on the molecular processes underlying these changes is limited. Aquaporins (PIP1 family) are important for refilling of embolized vessels and thereby helping the tree to recover after drought (Secchi and Zwieniecki, 2011; Laur and Hacke, 2014).

Anatomical and transcriptomic analyses of the developing xylem of poplars revealed that drought imposed changes that are similar to those found after salt exposure (Junghans et al., 2006; Bogeat-Triboulot et al., 2007; Janz et al., 2012; Wildhagen et al., 2018). For example, salt stressed poplar trees reinforce cell walls by increasing wall thickness relative to lumen area and avoid a strong loss in hydraulic conductivity by enhancing vessel frequency (Janz et al., 2012). Transcript abundances of of genes encoding fasciclin-like arabinogalactans (FLA), COBRA-like proteins, xyloglucan-endo-transglycolyases, pectin methylesterases were jointly repressed in developing xylem, while those of genes activating stress and defense responses increased (Janz et al., 2012). Similarly, in water-stressed poplars, transcript abundances of several cellulose synthases, arabinogalactan (AGP) and fasciclin-like proteins decreased (Berta et al., 2010). Wildhagen et al. (2018) also reported massive changes in the regulation of genes required for cell wall forming enzymes. Unexpectedly, drought decreased lignin and increased the saccharification potential of the wood (Wildhagen et al., 2018), indicating positive changes with regard to the biotechnological usage of wood. It would therefore be worthwhile

to test whether these changes can be achieved without the typical drought-induced growth-defense trade-off.

Drought further activates antioxidant defenses in the cambium of different poplar clones [Dvina (*P. deltoides*), I-214 (*P.* × *canadensis*), Pallara et al., 2012]. A distinct isoform of CATALASE 3 (CAT3) was strongly enhanced under water deficit, unpinning an essential role for this enzyme in ROS control under drought stress (Pallara et al., 2012). Furthermore, increases in the concentrations of osmotically active solutes in the cambial region of *P. alba* accompanied reductions in predawn leaf water potentials and stem dehydration (Pallara et al., 2012).

Leaves

Stomatal regulation is one of the most important mechanisms to adjust water consumption to fluctuations in water availability (Tardieu and Simonneau, 1998; McDowell et al., 2008; Skelton et al., 2015). Regulation of stomatal aperture reduces water loss by leaf transpiration (Stålfelt, 1955; Barrs, 1971; Brodribb and Holbrook, 2003; Araújo et al., 2011) but there is a trade-off between transpirational water loss and CO_2 assimilation (Jarvis and Jarvis, 1963; Cowan, 1978). Pioneering studies involving poplar species and hybrids shed light on two drought stress response strategies, anisohydric and isohydric behavior with divergent consequences for water flux and biomass production (Marron et al., 2003; Monclus et al., 2006; Giovannelli et al., 2007). Anisohydric plants keep their stomata relatively widely open and prevent dehydration by increasing the osmotic pressure in leaves (Gebre et al., 1994; Marron et al., 2002; Hanin et al., 2011; Barchet et al., 2014; Martorell et al., 2015); thereby, they are capable to support growth and biomass production (Passioura, 2002). Isohydric plants limit water loss through sensitive stomatal regulation and closure and/or by leaf abscission (Couso and Fernández, 2012). Poplars are isohydric species but exhibit a suite of adaptive measures (Brignolas et al., 2000) such as variation in stomatal sensitivity (Hamanishi et al., 2012), leaf shedding (Marron et al., 2002) and growth decline (Giovannelli et al., 2007). In a population of tree species a continuum of responses to water deficit between isohydryic and anisohydric behavior can be found (Klein, 2014). For instance, beech generally exhibits isohydric behavior but progenies from dry habitats showed stronger anisohydric behavior than those from wet habitats (Nguyen et al., 2017). In poplar, heritability of stomatal responsiveness to water deficit is generally high, indicating that this trait is a useful target for genetic engineering (Orlovic et al., 1998; Al Afas et al., 2006; Monclus et al., 2006).

Endogenous ABA is rapidly produced upon water deficit, initiating a signaling cascade that results in down-stream responses such as stomatal closure (Fujii et al., 2009). Besides roots, ABA biosynthesis takes place in leaves (Kuromori et al., 2018). Stomatal responsiveness to ABA shows large variation among different species and is evolutionary determined (Sussmilch et al., 2017). In angiosperms, ABA induced stomatal closure is usually rapid and can occur within seconds or minutes (Geiger et al., 2011), thus, not requiring *de novo* transcription. Over-expression of the ABA biosynthesis protein 9-cis-epoxycarotenoid dioxygenase 3 (NCED3) is beneficial for water-use efficiency (WUE) and results in enhanced drought

resistance in several plant species (Iuchi et al., 2001; Tung et al., 2008). Drought induced changes in stomatal development involve regulation of transcript abundance of the poplar orthologs of STOMAGEN, ERECTA, and STOMATA DENSITY AND DISTRIBUTION 1 (SDD1) (Harfouche et al., 2014). Interestingly, WUE was increased by the over-expression of a poplar ortholog of ERECTA in A. thaliana (Xing et al., 2011). ERECTA controls stomatal density but the sequence of events leading to this effect is still unknown (Xing et al., 2011). Genes such as ERECTA, SDD1, or NCED3 should be the focus of future research programs aimed at developing transgenic or gene-edited trees with resistance to naturally occurring field drought conditions.

At the cellular level, biochemical protection measures are activated in response to drought to avoid negative consequences of stress-induced ROS and to endure water deficit (Wang et al., 2003, **Figure 2**). Moderate water stress results in increased concentrations of soluble carbohydrates and polyols, which potentially promote maintenance of cell turgor in P. euphratica leaves through increased osmotic pressure (Bogeat-Triboulot et al., 2007). However, after salt acclimation bulk soluble carbohydrates (including glucose and fructose, sucrose), sugar alcohols, organic acids, mostly decrease or remain almost unaffected (Ottow et al., 2005; Dluzniewska et al., 2007; Ehlting et al., 2007; Brinker et al., 2010), suggesting that moderate salt accumulation in leaves may compensate for osmotic adjustment as observed in some halophytic species (Ma et al., 2012). It is noteworthy that amino acids, in particular proline, increase drastically in both water- and salt-stressed leaves (Brosché et al., 2005; Ottow et al., 2005; Dluzniewska et al., 2007; Ehlting et al., 2007; Pallara et al., 2012). In accordance, the mRNA levels of genes encoding enzymes that catalyze rate-limiting steps of proline synthesis and degradation [delta-1-pyrroline-5-carboxylate synthase (PcP5CS) and proline dehydrogenase] accumulate under osmotic stress (Dluzniewska et al., 2007). However, the bulk rise of proline to μM levels is insufficient to explain the observed change in osmotic pressure in salt exposed trees required to maintain water uptake (Ottow et al., 2005; Brinker et al., 2010). Therefore, increased proline may act as a protectant of protein integrity rather than function as an osmolyte in leaves.

Antioxidative systems also play an important role in the defense against negative consequences of drought stress (Hasanuzzaman et al., 2013). Therefore, one would expect that populations of wild tree species (beech) from dry conditions contain higher ROS protection than those from mesic conditions. Unexpectedly, the opposite was observed: unstressed beech from a mesic habitat showed a higher antioxidative capacity than those from a dry habitat and moreover, those from mesic conditions showed a stronger antioxidative response to drought than those from the dry habitat (Carsjens et al., 2014). These observations suggest that trees exposed infrequently to stress respond more flexibly, whereas long-term stress adapted trees are protected by resistance measures, which are already in place before the onset of acute stress. This view is also supported by constitutively enhanced salt tolerance of P. euphratica compared to salt sensitive poplars (Janz et al., 2010). The enhanced tolerance

of P. euphratica is, for example, based on the expansion of the sodium:proton antiporter family in the genome of this species (Ma et al., 2013). These few and selective examples highlight that divergent strategies may be required for improving drought resistance in short- or long-term water limited environments.

GENETIC APPROACHES FOR INCREASED STRESS TOLERANCE

Because drought and other osmotic stresses result in multiple tolerance or avoidance mechanisms, simple strategies for improving the performance of trees in water-limited environments do not exist. In order to target a suite of genes that can enhance drought tolerance, recent attempts to improve plant performance have often focused on signal perception and transduction (**Table 1**), whereas overexpression of structural genes found fewer applications (**Table 2**). Strategies for the selection of candidate genes relied mainly on the inducibility of genes under stress or on the gene origin in a highly stress tolerant species (**Tables 1**, **2**).

In most cases, candidate genes for stress tolerance were expressed under the 35S promoter, leading to high constitutive production in the transgenic plant (**Table 1**). A drawback of this approach is that more drought resistant plants often show biomass yield trade-off (e.g., the dwarfed eui mutant, a mutant in the GA-regulating CYP714 A3 gene, Wang C. et al., 2016). The utilization of stress-inducible promoters may be promising to achieve a balance between growth under non-stress conditions and enhanced defense activation under drought conditions. For example, a novel zinc finger protein from the succulent, xerophytic species Z. xanthoxylum rendered transgenic plants more tolerant to osmotic stress (Chu et al., 2016; **Table 1**). Similarly, overexpression of DREB (dehydration responsive protein binding element) under the RD29 promoter activated osmolytes (sugars) and enhanced the performance of transgenic plants under drought stress (Zhou et al., 2012; **Table 1**). Other studies showed successful activation of antioxidants, reduction of membrane leakage and increased photosynthesis, when YUCCA6 (a flavin mono-oxygenase-like from Arabidopsis thaliana) or choline oxidase (from bacteria) were overexpressed under an oxidative stress-inducible promoter (Ke et al., 2015, 2016; **Table 2**). However, overall utilization of stress-inducible promoters is still rare.

Plant model species, in which drought responses have often been studied and for which genomic information is available for a decade or longer, such as A. thaliana and Populus spp. were often used as the source species of inducible genes. In recent years, the gene pool of drought and salt tolerant woody species has increasingly been tapped. Among these species are: the succulent, xerophyte Z. xanthoxylum, the salt-tolerant and facultative succulent poplar, P. euphratica, the salt- and drought-tolerant species, Tamarix hispida and the salt-tolerant Fraxinus velutina (**Tables 1**, **2**). Other crops and woody species that have also been increasingly used as gene source are: Diospyros kaki (a widely cultivated fruit tree in China), Phyllostachys edulis (bamboo), Morus spp. (mulberry, feed for silkworms),

TABLE 1 | Functional characterization of drought- and salt-inducible protein kinase and transcription factors originating from or expressed in trees species.

Genes	Sources	Abiotic and chemical elicitors	Functional characterization				
			Promoter	Transgenic species	Phenotypes of stress tolerance conditions	Functional traits of transgenics under stress conditions	References
SIGNAL PERCEPTION AND TRANSDUCTION							
PYR/PYL/RCAR PYRL1 PYRL5	Populus trichocarpa	drought ABA	2x CaMV 35S	Populus davidiana x P. bolleana	Drought osmotic stress cold	↑ stomatal closure ↑ drought tolerance ↑ cold tolerance ↑ osmotic stress resistance ↓ water loss	Yu J. et al., 2017
PYR/PYL/RCAR PYRL1 PYRL5	Populus trichocarpa	drought ABA	2x CaMV 35S	A. thaliana	drought	↑ interaction with PP2C ↑ ABA sensitivity ↑ recovery after re-watering ↓ seed germination	Yu J. et al., 2016
PP2C HAB1	Populus euphratica	drought ABA	CaMV 35S	A. thaliana	drought	↑ ABA response ↑ drought sensitivity ↑ interaction with PYL4 in yeast two-hybrid assays	Chen et al., 2015
MAPK MAPK1	Morus L	Drought Cold Heat NaCl	CaMV 35S	A. thaliana	NaCl H_2O_2	↑ NaCl tolerance ↑ H_2O_2 tolerance ↓ heat and drought tolerance	Liu et al., 2017
MYB MYB1	Fraxinus velutina	NaCl drought	CaMV 35S	Nicotiana tabacum	NaCl	↑ biomass production ↑ SOD and POD activity ↑ proline content	Li et al., 2016
NAC NAC034 NAC035 NAC045	Populus euphratica	Drought NaCl	CaMV 35S	A. thaliana P. tomentosa	drought NaCl	↑ root development ↑ stem elongation ↑ drought sensitivity ↑ NaCl sensitivity	Lu et al., 2017
NAC NAC3	Cicer arietinum	drought ABA NaCl	CaMV 35S	Populus deltoides x P. euramericana	drought NaCl	↑ improved growth under drought, comparable to WT ↑ proline content ↑ photosynthetic pigments ↑ antioxidant enzymes	Movahedi et al., 2015
NAC NAC13	Tamarix hispida	NaCl drought	CaMV 35S	Tamarix hispida A. thaliana	NaCl osmotic stress	↑ POD and SOD activity ↑ proline content ↑ chlorophyll content ↓ ROS ↓ MDA	Wang L. et al., 2017

(Continued)

TABLE 1 | Continued

Genes	Sources	Abiotic and chemical elicitors	Promoter	Transgenic species	Phenotypes of stress tolerance conditions	Functional characterization — Functional traits of transgenics under stress conditions	References
NAC SNAC1	Phyllostachys edulis	auxin	CaMV 35S	A. thaliana	drought PEG	↑ survival rate under drought ↑ overall growth	Wang L. et al., 2016
ErbB-3 binding protein1 EBP1	Hevea brasiliensis	cold drought ABA	CaMV 35S	A. thaliana	drought	↑ enhanced expression of RD29A, RD22 and CYCD3;1 ↑ organ size	Cheng et al., 2016
ZFP (zink finger protein) ZxZF	Zygophyllum xanthoxylum	drought	RD29A	thaliana Populus x euramericana	Mannitol ABA	↑ PSII activity ↑ root growth ↑ antioxidant enzymes ↓ H$_2$O$_2$	Chu et al., 2016
EPF (epidermal patterning factor) PdEPF2	Populus nigra x (P. deltoids x P. nigra)	drought	CaMV 35S	A. thaliana	drought	↑ germination rate ↑ primary root length ↑ proline content ↑ chlorophyll content ↑ photosynthetic activity	Liu et al., 2016
CBF (C-repeat binding factor) CBF4	Populus euphratica	cold oxidative stress drought ABA NaCl	CaMV 35S	Populus tomentosa	drought ABA NaCl	↑ iWUE ↑ photosynthesis rate ↑ antioxidative enzymes ↑ proline ↓ overall growth (dwarf)	Tian et al., 2017
ERF ERF76	P. simonii x P. nigra	NaCl	CaMV 35S	P. simonii x P. nigra	NaCl	↑ ABA ↑ GA ↑ upregulation of 16 TFs and 45 stress related genes	Yao et al., 2016
START transcription factor EDT1/HDG11	A. thaliana	drought NaCl	CaMV 35S TUB2	Populus tomentosa, Gossypium hirsutum	drought NaCl	↑ proline content ↑ soluble sugar content ↑ antioxidant enzymes ↑ NaCl tolerance ↑ drought tolerance	Yu L. H. et al., 2016
CDPK (Calcium -dependent protein kinase) PeCPK10	Populus euphratica	Drought Salt Cold	CaMV 35S	A. thaliana	Drought Freezing	↑ expression of stress/ABA- responsive genes	Chen et al., 2013
bZIP (Basic leucine zipper protein) ThbZIP1	Tamarix hispida	NaCl PEG 6000 NaHCO$_3$ CdCl$_2$	CaMV 35S	Nicotiana tabacum	Salt	↑ activity of POD and SOD ↑ soluble sugars ↑ soluble proteins	Wang et al., 2010

(Continued)

TABLE 1 | Continued

Genes	Sources	Abiotic and chemical elicitors	Promoter	Transgenic species	Phenotypes of stress tolerance	Functional characterization — Functional traits of transgenics under stress conditions	References
Nucleoside diphosphate kinases NDPK2	Populus trichocarpa	ROS	CaMV 35S	Populus deltoides x P. euramericana	NaCl drought	↑ auxin related genes ↑ drought tolerance ↑ NaCl tolerance	Zhang et al., 2017
Nuclear factor Y NF-YB3	Picea wilsonii	NaCl heat PEG not ABA not cold	CaMV 35S	A. thaliana	drought NaCl ABA	↑ NaCl tolerance ↑ SOS3 transcript levels ↑ drought tolerance ↑ CDPK1 transcript levels ↑ CBF marker genes	Zhang et al., 2015
DREB (Dehydration-responsive element binding protein) PeDREB2a	Populus euphratica	Drought NaCl Low temperature NAA 6-BA GA$_3$ Not by ABA	RD29A	A. thaliana Lotus corniculatus	Salt Drought	↑ root length and plant height ↑ soluble sugars levels ↓ MDA levels	Zhou et al., 2012
DREB PeDREB2	Populus euphratica	Drought High salinity Cold Not by ABA	CaMV 35S	Tobacco	High-salt	↑ seed germination ↑ chlorophyll levels ↑ chlorophyll fluorescence ↓ growth retardation	Chen et al., 2009
DREB PeDREB2L	Populus euphratica	dehydration salt ABA	CaMV 35S	A. thaliana	Drought Freezing	↑ DRE/CRT-containing stress-responsive genes, RD29A and RD29B	Chen et al., 2011
DREB ThDREB	Tamarix hispida	NaCl PEG NaHCO$_3$ CdCl$_2$	CAMV 35S	Tobacco Tamarix hispida (transient transgenic)	NaCl Mannitol	↑ antioxidase activity (SOD,POD) ↓ ROS (H$_2$O$_2$) ↓ MDA content ↓ electrolyte leakage (EL)	Yang et al., 2017
DREB DREB4A	Morus alba	cold drought	CaMV 35S	A. thaliana N. tabacum	Drought NaCl heat	↑ regulation of ABI1 and ABI2 ↑ proline content ↓ MDA ↑ NaCl, heat and drought tolerance in tobacco	Liu et al., 2015
DREB DREB2A	Fraxinus pennsylvanica	drought	2x CaMV 35S	Robinia pseudoacacia	drought	↑ stress-inducible genes ↑ ABA levels ↑ auxin levels ↓ GA levels ↓ zeatin riboside	Xiu et al., 2016

(Continued)

TABLE 1 | Continued

Genes	Sources	Abiotic and chemical elicitors	Promoter	Transgenic species	Functional characterization		
					Phenotypes of stress tolerance	Functional traits of transgenics under stress conditions	References
CBL *PeCBL6 PeCBL10*	*Populus euphratica*	Drought High salinity Cold Not by ABA	CAMV35S	*A. thaliana*	Salinity Drought Low temperature	↑ fresh mass, survival rate ↑ chlorophyll content ↑ chlorophyll fluorescence (Fv/Fm)	Li et al., 2013
NAC [No apical meristem (NAM), (CUC)] superfamily *PeNAC1*	*Populus euphratica*	Drought Salt stress ABA (slightly induced)	CaMV 35S	*A. thaliana*	Salt	↑ survival rates, fresh weights ↑ capacity for K^+ uptake and transport ↓ AtHKT1 gene expression ↓ Na^+ content	Wang et al., 2013
NAC *PeNAC036*	*Populus euphratica*	Drought Salt ABA	CaMV 35S	*A. thaliana* (WT) and mutant *anac072*	Salt Drought	↑ plant height, primary root length (1/2 MS agar medium) ↑ survival rate (Soil culture) ↑ expression levels of COR47, RD29B, ERD11, RD22 and DREB2A	Lu et al., 2017
ERF (Ethylene) response factor *ERF76*	*Populus simonii* x *Populus nigra*	Drought Salinity ABA .	CaMV 35S	*P. simonii* x *P. nigra*	Salt	↑ Plant height, root length, and fresh weight ↑ gene expression of ABA and GA signal pathways ↑ defense-related genes, such as LEA, GST and HRGP genes ↑ signal transfer-related Genes (incl). P-tyrosine phosphatases, MAPKKK and PR5K	Wang et al., 2014 Yao et al., 2016
WRKY *ThWRKY4*	*Tamarix hispida*	Drought Salt ABA	CaMV 35S	*A. thaliana*	ABA Salt Oxidative stress	↑ SOD and POD activity ↓ ROS level ↓ cell death	Zheng et al., 2013
ZFP (Zinc finger protein) *ThZFP1*	*Tamarix hispida*	NaCl Mannitol ABA	CaMV 35S	*A. thaliana Tamarix hispida* (transient overexpression or knockdown)	ABA NaCl Mannitol	↑ SOD and POD activity and encoding genes ↑ proline level, and *ThP5CS1&2* genes ↑ chlorophyll content ↓ O_2^- and H_2O_2 ↓ MDA content, membrane lipid peroxidation, electrolytic leakage	Zang et al., 2015

Polyethyleneglycol (PEG), 1-naphthaleneacetic acid (NAA), 6-benzyl aminopurine (6-BA), gibberellic acid (GA), abscisic acid (ABA), superoxide dismutase (SOD), peroxidase (POD), malondialdehyde (MDA), Cauliflower mosaic virus (CaMV), Electrolyte leakage (EL), instantaneous water use efficieny (iWUE), ↑, Increase; ↓, Decrease.
Data were searched between 2013 and 2018 with the key words: Tree and overexpression and drought in Web of Science.

TABLE 2 | Functional characterization of drought- and salt-inducible structural genes originating from or expressed in trees species.

Genes	Sources	Abiotic and chemical elicitors	Promoter	Transgenic species	Phenotypes of stress tolerance	Functional traits of transgenics under stress conditions	References
DEFENSE AND OTHER STRUCTURAL GENES							
Cysteine protease *SmCP*	*Salix matsudana*	NaCl	CaMV 35S	*A. thaliana*	NaCl	↑ germination rate ↑ SOD activity ↓ MDA content ↓ electrolytic leakage	Zheng et al., 2018
Aquaporin *PeAQUA1*	*Populus x euramericana*	drought NaCl wounding	CaMV 35S	*Populus alba*	zinc	↑ growth rate ↑ intrinsic Transpiration Efficiency	Ariani et al., 2016
NXH (proton-sodium antiporter) *ZxNHX* **H$^+$-PPase** *ZxVP1-1*	*Zygophyllum xanthoxylum*	NaCl drought auxin	CaMV 35S	*Medicago sativa* (Alfalfa)	drought	↑ biomass production ↑ Na$^+$, K$^+$ and Ca^{2+} accumulation in leaves and roots ↑ leaf relative water content ↑ greater photosynthesis capacity	Bao et al., 2015
Dehydrin *HbDHN1* *HbDHN2*	*Hevea brasiliensis*	ABA ET JA NaCl drought heat wounding	CaMV 35S	*A. thaliana*	NaCl drought water stress osmotic stress	↑ APX & SOD ↑ drought tolerance ↑ NaCl tolerance ↑ tolerance to osmotic stress ↓ H$_2$O$_2$	Cao Y. et al., 2017
Ascorbate Peroxidase *APX*	*Populus tomentosa*	H$_2$O$_2$	CaMV 35S	*Nicotiana tabacum*	Drought NaCl Oxidative stress	↑ chlorophyll content ↑ NADP to NADPH ratio ↓ H$_2$O$_2$ ↓ MDA	Cao S. et al., 2017
XTH *DkXTH1*	*Diospyros kaki*	darkness cold/heat-shock ET ABA GA IAA	CaMV 35S	*A. thaliana, Solanum lycopersicum*	NaCl drought	↑ NaCl tolerance in A. thaliana and tomato ↑ drought tolerance in A. thaliana and tomato	Han et al., 2017
Choline oxidase *codA*	bacteria	cold NaCl drought	SWPA2 (oxidative stress inducible)	*Populus alba x glandulosa*	NaCl drought	↑ higher glycin betain levels ↑ auxin responsive genes ↑ PSII activity ↓ membrane leakage ↓ ROS production	Ke et al., 2016
FMO (flavin monooxygenase-like) *YUCCA6*	*A. thaliana*	auxin	SWPA2 (oxidative stress inducible)	*Populus alba x glandulosa*	drought	↑ auxin production ↑ drought tolerance ↑ PSII activity ↓ membrane leakage ↓ main root growth ↓ membrane leakage	Ke et al., 2015

(Continued)

TABLE 2 | Continued

Genes	Sources	Abiotic and chemical elicitors	Functional characterization				
			Promoter	Transgenic species	Phenotypes of stress tolerance	Functional traits of transgenics under stress conditions	References
GS (glutamine synthase) GS1	*Pinus sp.*	drought nitrogen	CaMV 35S	*Hybrid poplar*	drought	↑ WUE ↑ NUE ↑ glutamine, GABA, putrescin, hydroxyrpoline	Molina-Rueda and Kirby, 2015
STS (Stilbene synthase) MaSTS1 MaSTS2 MaSTS3 MaSTS4	*Morus atropurpurea*	SA ABA wounding NaCl	CaMV 35S	*Nicotiana tabacum*	heat NaCl PEG	↑ trans-resveratrol levels ↑ drought tolerance ↑ NaCl tolerance	Wang et al., 2017b
CYP450 CYP714A3	*Populus trichocarpa*	GA	Eui (GA mutant)	*Oryza sativa*	NaCl	↑ NaCl tolerance ↓ GA levels ↓ excessive shoot growth of *eui* mutant is compensated	Wang C. et al., 2016
Aquaporin TIP4;1-1	*Phyllostachys edullis*	drought NaCl	CaMV 35S	*A. thaliana*	drought NaCl	↑ antioxidant enzymes ↑ photosynthetis ↓ MDA	Sun et al., 2017
UDP-galactose-4-epimerase PeUGE	*Phyllostachys edulis*	drought NaCl Water stress	CaMV 35S	*A. thaliana*	Drought NaCl	↑ chlorophyll fluorescence	Sun et al., 2016
Vacuolar H$^+$-pyrophosphatase AVP1	*Populus trichocarpa*	NaCl drought pH	CaMV 35S	*A. thaliana* *Populus davidiana × bolleana*	NaCl	↑ prevents NaCl accumulation ↑ higher ion efflux	Yang et al., 2015
Galactinol synthase GolS2 **SNF1-related protein kinase** SRK2C	*A. thaliana*	PEG ABA NaCl	CaMV 35S	*Populus tremula × tremuloides*	NaCl	↑ abiotic stress tolerance ↓ OEX did not induce gene expression in poplar as in *A. thaliana*	Yu X. et al., 2017
Hydrolases superfamily protein PAD4	*Populus tremula x tremuloides*	SA drought UV-light root hypoxia	RNAi	*A.thaliana Populus tremula x tremuloides*	drought	↑ drought sensitivity ↓ water use	Szechynska-Hebda et al., 2016

Overexpression (OEX), Polyethyleneglycol (PEG), 1-naphthaleneacetic acid (NAA), 6-benzyl aminopurine (6-BA), gamma-Aminobutyric acid (GABA), gibberellic acid (GA), abscisic acid (ABA), superoxide dismutase (SOD), ascorbate peroxidase (APX), Salicylic Acid (SA), peroxidase (POD), malondialdehyde (MDA), Cauliflower mosaic virus (CaMV), Electrolyte leakage (EL), ↑, Increase; ↓, Decrease.
Data were searched between 2013 and 2018 with the key words: Tree and overexpression and drought in Web of Science.

Hevea brasiliensis (rubber), *Picea* and *Pinus* (conifers), and *Cicer arietium* (herbaceous legume crop such chikpea). The target species were model species such as poplars, *Arabidopsis* and *Nicotiana tabacum*, but also crops such as alfalfa, cotton, lotus, tomato and rice. Transformation of non-model tree species for enhanced stress tolerance is still rare but recent results showed promise. Overexpression of *DREB2A*, a gene that forms a hub for drought-stress related gene expression in *Robinia pseudoacacia* resulted in enhanced drought resistance (Xiu et al., 2016). The drought resistant phenotype was mediated by the formation of deeper roots and decreased oxidative stress, and most likely mediated by effects on the phytohormone balance of the plants (Xiu et al., 2016, **Table 1**).

Succulence, which occurs in many drought or salt resistant species, is a complex trait that may prove to be useful for drought resistance. Leaf thickness and water content increase with increasing salinity and aridity (Ottow et al., 2005; Nguyen et al., 2017). Succulent leaves exhibit a significant water storage capacity and dilute intrinsic salt concentrations (Ottow et al., 2005; Scholz et al., 2011; Han et al., 2013; Ishii et al., 2014). Overexpression of a putative xyloglucan endotransglucosylase/hydrolase from *P. euphratica* (*PeXTH*) contributed to salt-induced leaf succulence (Han et al., 2013) by improving cell wall properties to cope with water deficit and high salinity (Cho et al., 2006). Overexpression of a hot pepper (*Capsicum annuum*) *CaXTH3* in guard cells reduced transpiration under dehydration stress, thus, supporting a role of *XTHs* in drought resistance (Choi et al., 2011).

As highlighted before, the acclimatory responses of trees to drought invoke a multitude of molecular and biochemical changes. Consequently, a focus of many recent genetic approaches was on genes encoding protein kinases and transcription factors to potentially target whole signaling and biochemical pathways instead of single gene products. Overexpression of *CPK* (calcium-dependent protein kinase, Chen et al., 2013), *bZIP* (Basic leucine zipper protein, Wang et al., 2010), *DREB* (dehydration-responsive element-binding protein, Chen et al., 2009, 2011; Zhou et al., 2012; Yang et al., 2017), *CBL* (calcineurin B-like protein, Li et al., 2013), *NAC* [no apical meristem (NAM, Wang L. et al., 2016, 2017), *ATAF* (*Arabidopsis* transcription activation factor), *CUC* (cup-shaped cotyledon) superfamily, Wang et al., 2013; Lu et al., 2017], *ERF* (ethylene response factor, Wang et al., 2014; Yao et al., 2016), *WRKY* (Zheng et al., 2013), and *ZFP* (zinc finger proteins, Zang et al., 2015) often resulted in enhanced photosynthesis, higher WUE, higher activity of antioxidative enzymes, lower oxidative damage and improved growth under osmotic stress (**Table 1**). Examples are still rare where drought and salt responses are not congruent (**Table 1**). For example, overexpression of a *MAPK1* of the MAPK C family resulted in more salt tolerant but less drought and heat resistant plants but the underlying mechanisms for this difference are speculative (Liu et al., 2017).

ABA is crucial in mediating plant drought responses. Most of the signal transduction and response factors used for stress amelioration are regulated by ABA (**Table 1**). The receptor RCAR is the first target of ABA and forms a complex with PP2C for stress signaling (Fujita et al., 2011; de Zelicourt et al., 2016). The situation is even more complex because multiple RCARs and PP2Cs exist that are forming combinatorial interaction

networks (Tischer et al., 2017). *Arabidopsis* and hybrid poplar overexpressing *RCARs* from *P. trichocarpa* were more drought tolerant than the wildtype through decreased water loss and increased osmotic and antioxidative protection (Yu J. et al., 2016, 2017). However, there is also fitness trade-off because germination of the overexpressed *Arabidopsis* seeds is inhibited (Yu J. et al., 2016). The transgenic poplar phenotype was normal and the biomass gain under strong drought stress was higher than that of the controls (Yu J. et al., 2017). HAB1 was suggested to be the ortholog of *Arabidopsis* PP2C, which is a negative regulator of ABA signaling and acts as co-receptor for RCARs. *Arabidopsis* overexpressing *HAB1* gene from *P. euphratica* lost ABA sensitivity and became more drought sensitive than the wildtype (Chen et al., 2015). Overall, functional characterization of these genes indicated that the core ABA signaling pathway is conserved in poplar and may be a suitable target for genetic engineering.

Studies applying novel gene editing methods (CRISPR/Cas9) to improve drought tolerance are still in their infancy but hold promise for new discoveries. For example, lignin deposition was reduced in poplars in which *Myb170* expression was abolished by CRISPR/Cas9 (Xu et al., 2017). Surprisingly, heterologous expression of *Myb170* in *Arabidopsis* uncovered its presence in guard cells, which showed stronger stomatal closure at night and thereby, enhanced drought protection (Xu et al., 2017). This study illustrates that novel functions of genes can be detected by combining CRISPR/Cas9 and overexpression.

CONCLUSIONS AND RESEARCH NEEDS

Trees are capable of responding to drought stress through a wide variety of cellular and physiological acclimation strategies, which form the basis for genetic improvements of drought tolerance. In particular, overexpression of drought sensing, signal transduction, and drought responsive transcription factors can enhance drought tolerance in a variety of model systems and some economically important woody species. Our overview on transgenic modifications revealed that modifications at the cellular level were the main targets, often using genes from drought or salt tolerant woody species for overexpression. However, systematic studies to clarify if these genes perform better than those from drought sensitive species are lacking. Comparative studies suggest that amplification of distinct gene families such as the SOS pathway in *P. euphratica*, gene duplication, and evolutionary recruitment of distinct metabolites such as ABA for stomatal regulation could also be important avenues for future research. Furthermore, long-term studies under field conditions are still scarce. There is obviously a strong need for testing genetically modified trees in their natural environment because the combination of stress factors such as heat and drought together may overrule the effects of single stressors present under laboratory conditions.

At a wider scale, we have to assert that our mechanistic understanding of the interplay among osmotic regulation, hydraulic adjustment and uptake systems for water and nutrients is still in its infancy. In particular, the root-to-shoot communication that sets off a suite of responses leading to morphological changes of the root system is not clear.

Therefore, an important future task will be to uncover the genetic basis for an optimized resource allocation between biochemical defenses and production of new structures such as deep rooting systems under stressful climatic conditions. Next-generation genomics and phenomics approaches will facilitate a better understanding of phenotype-genotype maps and help to formulate genomic-assisted breeding strategies in forest trees for resistance to drought stress and other osmotic cues.

AUTHOR CONTRIBUTIONS

FUNDING

REFERENCES

Al Afas, N., Marron, N., and Ceulemans, R. (2006). Clonal variation in stomatal characteristics related to biomass production of 12 poplar (*Populus*) clones in a short rotation coppice culture. *Environ. Exp. Bot.* 58, 279–286. doi: 10.1016/j.envexpbot.2005.09.003

Allen, C. D., Breshears, D. D., and McDowell, N. G. (2015). On underestimation of global vulnerability to tree mortality and forest die-off from hotter drought in the Anthropocene. *Ecosphere* 6:155. doi: 10.1890/ES15-00203.1

Allen, C. D., Macalady, A. K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., et al. (2010). A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecol. Manag.* 259, 660–684. doi: 10.1016/j.foreco.2009.09.001

Almeida-Rodriguez, A. M., Hacke, U. G., and Laur, J. (2011). Influence of evaporative demand on aquaporin expression and root hydraulics of hybrid poplar. *Plant Cell Environ.* 34, 1318–1331. doi: 10.1111/j.1365-3040.2011.02331.x

Anderegg, W. R. (2015). Spatial and temporal variation in plant hydraulic traits and their relevance for climate change impacts on vegetation. *New. Phytol.* 205, 1008–1014. doi: 10.1111/nph.12907

Aranda, I., Cano, F. J., Gasc,ó, A., Cochard, H., Nardini, A., Mancha, J. A., et al. (2015). Variation in photosynthetic performance and hydraulic architecture across European beech (*Fagus sylvatica L.*) populations supports the case for local adaptation to water stress. *Tree Physiol.* 35, 34–46. doi: 10.1093/treephys/tpu101

Araújo, W. L., Fernie, A. R., and Nunes-Nesi, A. (2011). Control of stomatal aperture: a renaissance of the old guard. *Plant Signal. Behav.* 6, 1305–1311. doi: 10.4161/psb.6.9.16425

Ariani, A., Francini, A., Andreucci, A., and Sebastiani, L. (2016). Over-expression of AQUA1 in *Populus alba Villafranca* clone increases relative growth rate and water use efficiency, under Zn excess condition. *Plant Cell Rep.* 35, 289–301. doi: 10.1007/s00299-015-1883-9

Bao, A. K., Du, B. Q., Touil, L, Kang P., Wang, Q. L. and Wang, S. M. (2015). Co-expression of tonoplast Cation/H^+ antiporter and H^+-pyrophosphatase from xerophyte *Zygophyllum xanthoxylum* improves alfalfa plant growth under salinity, drought and field conditions. *Plant Biotechnol. J.* 14, 964–975. doi: 10.1111/pbi.12451

Barberon, M., Vermeer, J. E., De Bellis, D., Wang, P., Naseer, S., Andersen, T. G., et al. (2016). Adaptation of root function by nutrient-induced plasticity of endodermal differentiation. *Cell* 164, 447–459. doi: 10.1016/j.cell.2015.12.021

Barchet, G. L., Dauwe, R., Guy, R. D., Schroeder, W. R., Soolanayakanahally, R. Y., Campbell, M. M., et al. (2014). Investigating the drought-stress response of hybrid poplar genotypes by metabolite profiling. *Tree Physiol.* 34, 1203–1219. doi: 10.1093/treephys/tpt080

Barrs, H. D. (1971). Cyclic variations in stomatal aperture, transpiration, and leaf water potential under constant environmental conditions. *Annu. Rev. Plant. Physiol.* 22, 223–236. doi: 10.1146/annurev.pp.22.060171.001255

Berta, M., Giovannelli, A., Sebastiani, F., Camussi, A., and Racchi, M. L. (2010). Transcriptome changes in the cambial region of poplar (*Populus alba L.*) in response to water deficit. *Plant Biology* 12, 341–354. doi: 10.1111/j.1438-8677.2009.00320.x

Bogeat-Triboulot, M. B., Brosché, M., Renaut, J., Jouve, L., Thiec, D. L., Fayyaz, P., et al. (2007). Gradual soil water depletion results in reversible changes of gene expression, protein profiles, ecophysiology, and growth performance in *Populus euphratica*, a poplar growing in arid regions. *Plant Physiol.* 143, 876–892. doi: 10.1104/pp.106.088708

Bréda, N., Huc, R., Granier, A., and Dreyer, E. (2006). Temperate forest trees and stands under severe drought: a review of ecophysiological responses, adaptation processes and long-term consequences. *Ann. For. Sci.* 63, 625–644. doi: 10.1051/forest:2006042

Brignolas, F., Thierry, C., Guerrier, G., and Boudouresque, É. (2000). Compared water deficit response of two *Populus × euramericana* clones, Luisa Avanzo and Dorskamp. *Ann. For. Sci.* 57, 261–266. doi: 10.1051/forest:2000117

Brinker, M., Brosché, M., Vinocur, B., Abo-Ogiala, A., Fayyaz, P., Janz, D., et al. (2010). Linking the salt transcriptome with physiological responses of a salt-resistant *Populus* species as a strategy to identify genes important for stress acclimation. *Plant Physiol.* 154, 1697–1709. doi: 10.1104/pp.110.164152

Brodribb, T. J., Field, T. S., and Sack, L. (2010). Viewing leaf structure and evolution from a hydraulic perspective. *Func. Plant Biol.* 37, 488–498. doi: 10.1071/FP10010

Brodribb, T. J., and Holbrook, N. M. (2003). Stomatal closure during leaf dehydration, correlation with other leaf physiological traits. *Plant Physiol.* 132, 2166–2173. doi: 10.1104/pp.103.023879

Broschè, M., Vinocur, B., Alatalo, E. R., Lamminmäki, A., Teichmann, T., Ottow, E. A., et al. (2005). Gene expression and metabolite profiling of *Populus euphratica* growing in the Negev desert. *Genome Biol.* 6:R101. doi: 10.1186/gb-2005-6-12-r101

Brunner, I., Herzog, C., Dawes, M. A., Arend, M., and Sperisen, C. (2015). How tree roots respond to drought. *Front. Plant Sci.* 6:547. doi: 10.3389/fpls.2015.00547

Buckley, T. N. (2005). The control of stomata by water balance. *New. Phytol.* 168, 275–292. doi: 10.1111/j.1469-8137.2005.01543.x

Cao, S., Du, X.-H., Li, L.-H., Liu, Y.-D., Zhang, L., Pan, X., et al. (2017). Overexpression of *Populus tomentosa* cytosolic ascorbate peroxidase enhances abiotic stress tolerance in tobacco plants. *Russ. J. Plant Physiol.* 64, 224–234. doi: 10.1134/S1021443717020029

Cao, Y., Xiang, X., Geng, M., You, Q., and Huang, X. (2017). Effect of *HbDHN1* and *HbDHN2* genes on abiotic stress responses in *Arabidopsis*. *Front. Plant Sci.* 8:470. doi: 10.3389/fpls.2017.00470

Carsjens, C., Nguyen Ngoc, Q., Guzy, J., Knutzen, F., Meier, I. C., Müller, M., et al. (2014). Intra-specific variations in expression of stress-related genes in beech progenies are stronger than drought-induced responses. *Tree Physiol.* 34, 1348–1361. doi: 10.1093/treephys/tpu093

Chaves, M. M., Maroco, J. P., and Pereira, J. S. (2003). Understanding plant responses to drought — from genes to the whole plant. *Functional Plant Biol.* 30, 239–264. doi: 10.1071/fp02076

Chen, J., Xia, X., and Yin, W. (2009). Expression profiling and functional characterization of a DREB2-type gene from *Populus euphratica*. *Biochem. Biophys. Res. Commun.* 378, 483–487. doi: 10.1016/j.bbrc.2008.11.071

Chen, J., Xia, X., and Yin, W. (2011). A poplar DRE-binding protein gene, PeDREB2L, is involved in regulation of defense response against abiotic stress. *Gene* 483, 36–42. doi: 10.1016/j.gene.2011.05.010

Chen, J., Xue, B., Xia, X., and Yin, W. (2013). A novel calcium-dependent protein kinase gene from *Populus euphratica*, confers both drought and cold stress tolerance. *Biochem. Biophys. Res. Commun.* 441, 630–636. doi: 10.1016/j.bbrc.2013.10.103

Chen, J., Zhang, D., Zhang, C., Xia, X., Yin, W., and Tian, Q. (2015). A putative PP2C-encoding gene negatively regulates ABA signaling in *Populus euphratica*. *PLoS ONE* 10:e0139466. doi: 10.1371/journal.pone.0139466

Chen, S., Li, J., Fritz, E., Wang, S., and Hüttermann, A. (2002c). Sodium and chloride distribution in roots and transport in three poplar genotypes under increasing NaCl stress. *Forest Ecol. Manag.* 168, 217–230. doi: 10.1016/S0378-1127(01)00743-5

Chen, S., Li, J., Wang, S., Fritz, E., Hüttermann, A., and Altman, A. (2003). Effects of NaCl on shoot growth, transpiration, ion compartmentation, and transport in regenerated plants of *Populus euphratica* and *Populus tomentosa*. *Can. J. For. Res.* 33, 967–975. doi: 10.1139/x03-066

Chen, S., Li, J., Wang, S., Hüttermann, A., and Altman, A. (2001). Salt, nutrient uptake and transport, and ABA of *Populus euphratica*; a hybrid in response to increasing soil NaCl. *Trees* 15, 186–194. doi: 10.1007/s004680100091

Chen, S., Li, J., Wang, T., Wang, S., Polle, A., and Hüttermann, A. (2002b). Osmotic stress and ion-specific effects on xylem abscisic acid and the relevance to salinity tolerance in poplar. *J. Plant Growth Regul.* 21, 224–233. doi: 10.1007/s00344-002-1001-4

Chen, S., and Polle, A. (2010). Salinity tolerance of *Populus*. *Plant Biol.* 12, 317–333. doi: 10.1111/j.1438-8677.2009.00301.x

Chen, S., Wang, S., Aloys Hüttermann Altman, A. (2002a). Xylem abscisic acid accelerates leaf abscission by modulating polyamine and ethylene synthesis in water-stressed intact poplar. *Trees*, 16, 16–22. doi: 10.1007/s00468-001-0138-2

Chen, S., Wang, S., Altman, A., and Hüttermann, A. (1997). Genotypic variation in drought tolerance of poplar in relation to abscisic acid. *Tree Physiol.* 17, 797–803.

Cheng, H., Chen, X., Zhu, J., and Huang, H. (2016). Overexpression of a *Hevea brasiliensis* ErbB-3 Binding protein 1 gene increases drought tolerance and organ size in *Arabidopsis*. *Front. Plant Sci.* 7:1703. doi: 10.3389/fpls.2016.01703

Cho, S. K., Kim, J. E., Park, J.-A., Eom, T. J., and Kim, W. T. (2006). Constitutive expression of abiotic stress-inducible hot pepper *CaXTH3*, which encodes a xyloglucan endotransglucosylase/hydrolase homolog, improves drought and salt tolerance in transgenic *Arabidopsis* plants. *FEBS Lett.* 580, 3136–3144. doi: 10.1016/j.febslet.2006.04.062

Choi, J. Y., Seo, Y. S., Kim, S. J., Kim, W. T., and Shin, J. S. (2011). Constitutive expression of *CaXTH3*, a hot pepper xyloglucan endotransglucosylase/hydrolase, enhanced tolerance to salt and drought stresses without phenotypic defects in tomato plants (*Solanum lycopersicum* cv. Dotaerang). *Plant Cell Rep.* 30, 867–877. doi: 10.1007/s00299-010-0989-3

Chu, Y., Zhang, W., Wu, B., Huang, Q., Zhang, B., and Su, X. (2016). Overexpression of the novel *Zygophyllum xanthoxylum* C2H2-type zinc finger gene ZxZF improves drought tolerance in transgenic *Arabidopsis* and poplar. *Biologia* 71, 769–776. doi: 10.1515/biolog-2016-0093

Claeys, H., and Inzé, D. (2013). The agony of choice: how plants balance growth and survival under water-limiting conditions. *Plant Physiol.* 162, 1768–1779. doi: 10.1104/pp.113.220921

Cocozza, C., Cherubini, P., Regier, N., Saurer, M., Frey, B., and Tognetti, R. (2010). Early effects of water deficit on two parental clones of *Populus nigra* grown under different environmental conditions. *Funct. Plant Biol.* 37, 244–254. doi: 10.1071/FP09156

Cohen, D., Bogeat-Triboulot, M. B., Tisserant, E., Balzergue, S., Martin-Magniette, M.-L., Lelandais, G., et al. (2010). Comparative transcriptomics of drought responses in *Populus*: a meta-analysis of genome-wide expression profiling in mature leaves and root apices across two genotypes. *BMC Genomics* 11:630. doi: 10.1186/1471-2164-11-630

Couso, L. L., and Fernández, R. J. (2012). Phenotypic plasticity as an index of drought tolerance in three Patagonian steppe grasses. *Ann. Bot.* 110, 849–857. doi: 10.1093/aob/mcs147

Cowan, I. R. (1978). "Stomatal Behaviour and Environment," in *Advances in Botanical Research*, eds. R. D. Preston and H. W. Woolhouse (London: Academic Press), 117–228.

Daszkowska-Golec, A., and Szarejko, I. (2013). Open or close the gate – stomata action under the control of phytohormones in drought stress conditions. *Front. Plant Sci.* 4:138. doi: 10.3389/fpls.2013.00138

Davies, W. J., and Bacon, M. A. (2003). "Adaptation of Roots to Drought," in *Root Ecology Ecological Studies*, eds H. de Kroon and E. T. W. Visser (Berlin, Heidelberg: Springer), 173–192.

de Zelicourt, A., Colcombet, J., and Hirt, H. (2016). The role of MAPK modules and ABA during abiotic stress signaling. *Trends Plant Sci.* 21, 677–685. doi: 10.1016/j.tplants.2016.04.004

Dietz, S., von Bülow, J., Beitz, E., and Nehls, U. (2011). The aquaporin gene family of the ectomycorrhizal fungus *Laccaria bicolor*: lessons for symbiotic functions. *New. Phytol.* 190, 927–940. doi: 10.1111/j.1469-8137.2011.03651.x

Dluzniewska, P., Gessler, A., Dietrich, H., Schnitzler, J.-P., Teuber, M., and Rennenberg, H. (2007). Nitrogen uptake and metabolism in *Populus* × *canescens* as affected by salinity. *New. Phytol.* 173, 279–293. doi: 10.1111/j.1469-8137.2006.01908.x

Ehlting, B., Dluzniewska, P., Dietrich, H., Selle, A., Teuber, M., Hänsch, R., et al. (2007). Interaction of nitrogen nutrition and salinity in Grey poplar (*Populus tremula* × *alba*). *Plant Cell Environ.* 30, 796–811. doi: 10.1111/j.1365-3040.2007.01668.x

Ellison, D., Morris, C. E., Locatelli, B., Sheil, D., Cohen, J., Murdiyarso, D., et al. (2017). Trees, forests and water: Cool insights for a hot world. *Glob. Environ. Change* 43, 51–61. doi: 10.1016/j.gloenvcha.2017.01.002

FAO (2016). *Global Forest Resources Assessment 2015: How Are the World's Forests Changing?* Avaialable online at: http://www.fao.org/resources/infographics/infographics-details/en/c/325836/

Fischer, U., and Polle, A. (2010). "*Populus* Responses to Abiotic Stress," in *Genetics and Genomics of Populus Plant Genetics and Genomics: Crops and Models*, eds S. Jansson, R. P. Bhalerao, and A. T. Groover (New York, NY: Springer), 225–246. doi: 10.1007/978-1-4419-1541-2_11

Fox, A. R., Maistriaux, L. C., and Chaumont, F. (2017). Toward understanding of the high number of plant aquaporin isoforms and multiple regulation mechanisms. *Plant Sci.* 264, 179–187. doi: 10.1016/j.plantsci.2017.07.021

Fujii, H., Chinnusamy, V., Rodrigues, A., Rubio, S., Antoni, R., Park, S. Y., et al. (2009). *In vitro* reconstitution of an abscisic acid signalling pathway. *Nature* 462, 660–664. doi: 10.1038/nature08599

Fujita, Y., Fujita, M., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2011). ABA-mediated transcriptional regulation in response to osmotic stress in plants. *J. Plant Res.* 124, 509–525. doi: 10.1007/s10265-011-0412-3

Gebre, G. M., Kuhns, M. R., and Brandle, J. R. (1994). Organic solute accumulation and dehydration tolerance in three water-stressed *Populus deltoides* clones. *Tree Physiol.* 14, 575–587.

Geiger, D., Maierhofer, T., Al-Rasheid, K. A., Scherzer, S., and Mumm, P., Liese, A., et al. (2011). Stomatal closure by fast abscisic acid signaling is mediated by the guard cell anion channel SLAH3 and the Receptor RCAR1. *Sci. Signal.* 4:ra32. doi: 10.1126/scisignal.2001346

Giovannelli, A., Deslauriers, A., Fragnelli, G., Scaletti, L., Castro, G., Rossi, S., et al. (2007). Evaluation of drought response of two poplar clones (*Populus* × *canadensis* Mönch 'I-214' and *P. deltoides* Marsh. 'Dvina') through high resolution analysis of stem growth. *J. Exp. Bot.* 58, 2673–2683. doi: 10.1093/jxb/erm117

González-Carranza, Z. H., Lozoya-Gloria, E., and Roberts, J. A. (1998). Recent developments in abscission: shedding light on the shedding process. *Trends Plant Sci.* 3, 10–14.

Hacke, U. G., Sperry, J. S., Pockman, W. T., Davis, S. D., and McCulloh, K. A. (2001). Trends in wood density and structure are linked to prevention of xylem implosion by negative pressure. *Oecologia* 126, 457–461. doi: 10.1007/s004420100628

Hacke, U. G., Sperry, J. S., Wheeler, J. K., and Castro, L. (2006). Scaling of angiosperm xylem structure with safety and efficiency. *Tree Physiol.* 26, 689–701. doi: 10.1093/treephys/26.6.689

Hadley, J. L., and Smith, W. K. (1990). Influence of leaf surface wax and leaf area to water content ratio on cuticular transpiration in western conifers, U.S.A. *Can. J. For. Res.* 20, 1306–1311. doi: 10.1139/x90-173

Hamanishi, E. T., and Campbell, M. M. (2011). Genome-wide responses to drought in forest trees. *Forestry* 84, 273–283. doi: 10.1093/forestry/cpr012

Hamanishi, E. T., Thomas, B. R., and Campbell, M. M. (2012). Drought induces alterations in the stomatal development program in *Populus*. *J. Exp. Bot.* 63, 4959–4971. doi: 10.1093/jxb/ers177

Han, Y., Han, S., Ban, Q., He, Y., Jin, M., and Rao, J. (2017). Overexpression of persimmon *DkXTH1* enhanced tolerance to abiotic stress and delayed fruit softening in transgenic plants. *Plant Cell. Rep.* 36, 583–596. doi: 10.1007/s00299-017-2105-4

Han, Y., Wang, W., Sun, J., Ding, M., Zhao, R., Deng, S., et al. (2013). *Populus euphratica XTH* overexpression enhances salinity tolerance by the development of leaf succulence in transgenic tobacco plants. *J. Exp. Bot.* 64:4225. doi: 10.1093/jxb/ert229

Hanin, M., Brini, F., Ebel, C., Toda, Y., Takeda, S., and Masmoudi, K. (2011). Plant dehydrins and stress tolerance. *Plant Signal. Behav.* 6, 1503–1509. doi: 10.4161/psb.6.10.17088

Harfouche, A., Meilan, R., and Altman, A. (2014). Molecular and physiological responses to abiotic stress in forest trees and their relevance to tree improvement. *Tree Physiol.* 34, 1181–1198. doi: 10.1093/treephys/tpu012

Hasanuzzaman, M., Nahar, K., Gill, S. S., and Fujita, M. (2013). "Drought stress responses in plants, oxidative stress, and antioxidant defense," in *Climate Change and Plant Abiotic Stress Tolerance*, eds N. Tuteja and S. S. Gill (Weinheim: Wiley-VCH Verlag GmbH & Co), 209–250.

Hennig, A. (2016). *Drought Stress Response Of Tetraploid Hybrid Aspen (Populus tremula L. x P. tremuloides Michx.) of Protoplast Fusion Experiments*. Available online at: https://ediss.uni-goettingen.de/handle/11858/00-1735-0000-0028-8754-5

Hennig, A., Kleinschmit, J. R., Schoneberg, S., Löffler, S., Janßen, A., and Polle, A. (2015). Water consumption and biomass production of protoplast fusion lines of poplar hybrids under drought stress. *Front. Plant Sci.* 6:330. doi: 10.3389/fpls.2015.00330

Ishii, H. R., Azuma, W., Kuroda, K., and Sillett, S. C. (2014). Pushing the limits to tree height: could foliar water storage compensate for hydraulic constraints in *Sequoia sempervirens*? *Funct. Ecol.* 28, 1087–1093. doi: 10.1111/1365-2435.12284

Iuchi, S., Kobayashi, M., Taji, T., Naramoto, M., Seki, M., Kato, T., et al. (2001). Regulation of drought tolerance by gene manipulation of 9-cis-epoxycarotenoid dioxygenase, a key enzyme in abscisic acid biosynthesis in *Arabidopsis*. *Plant J.* 27, 325–333. doi: 10.1046/j.1365-313x.2001.01096.x

Janz, D., Behnke, K., Schnitzler, J. P., Kanawati, B., Schmitt-Kopplin, P., and Polle, A. (2010). Pathway analysis of the transcriptome and metabolome of salt sensitive and tolerant poplar species reveals evolutionary adaption of stress tolerance mechanisms. *BMC Plant Biol.* 10:150. doi: 10.1186/1471-2229-10-150

Janz, D., Lautner, S., Wildhagen, H., Behnke, K., Schnitzler, J. P., Rennenberg, H., et al. (2012). Salt stress induces the formation of a novel type of "pressure wood" in two *Populus* species. *New. Phytol.* 194, 129–141. doi: 10.1111/j.1469-8137.2011.03975.x

Janz, D., and Polle, A. (2012). Harnessing salt for woody biomass production. *Tree Physiol.* 32, 1–3. doi: 10.1093/treephys/tpr127

Jarvis, P. G., and Jarvis, M. S. (1963). The water relations of tree seedlings.: IV. some aspects of the tissue water relations and drought resistance. *Physiol. Plant.* 16, 501–516. doi: 10.1111/j.1399-3054.1963.tb08327.x

Jin, X., Zimmermann, J., Polle, A., and Fischer, U. (2015). Auxin is a long-range signal that acts independently of ethylene signaling on leaf abscission in *Populus*. *Front. Plant. Sci.* 6:634. doi: 10.3389/fpls.2015.00634

Jones, H. G. (1993). "Drought tolerance and water-use efficiency", in *Water deficits: Plant responses from cell to community*, eds J. A. C. Smith and H. Griffiths (Oxford: BIOS Scientific Publishers), 193–203.

Junghans, U., Polle, A., Düchting, P., Weiler, E., Kuhlman, B., Gruber, F., et al. (2006). Adaptation to high salinity in poplar involves changes in xylem anatomy and auxin physiology. *Plant Cell Environ.* 29, 1519–1531

Kanno, Y., Hanada, A., Chiba, Y., Ichikawa, T., Nakazawa, M., Matsui, M., et al. (2012). Identification of an abscisic acid transporter by functional screening using the receptor complex as a sensor. *Proc. Natl. Acad. Sci. U. S. A.* 109, 9653–9658. doi: 10.1073/pnas.1203567109

Ke, Q., Wang, Z., Ji, C. Y., Jeong, J. C., Lee, H.-S., Li, H., et al. (2015). Transgenic poplar expressing *Arabidopsis YUCCA6* exhibits auxin-overproduction

phenotypes and increased tolerance to abiotic stress. *Plant Physiol. Biochem.* 94, 19–27. doi: 10.1016/j.plaphy.2015.05.003

Ke, Q., Wang, Z., Ji, C. Y., Jeong, J. C., Lee, H.-S., Li, H., et al. (2016). Transgenic poplar expressing *codA* exhibits enhanced growth and abiotic stress tolerance. *Plant Physiol. Biochem.* 100, 75–84. doi: 10.1016/j.plaphy.2016.01.004

Klein, T. (2014). The variability of stomatal sensitivity to leaf water potential across tree species indicates a continuum between isohydric and anisohydric behaviours. *Funct. Ecol.* 28, 1313–1320. doi: 10.1111/1365-2435.12289

Koiwai, H., Nakaminami, K., Seo, M., Mitsuhashi, W., Toyomasu, T., and Koshiba, T. (2004). Tissue-specific localization of an abscisic acid biosynthetic enzyme, AAO3, in *Arabidopsis*. *Plant Physiol.* 134, 1697–1707. doi: 10.1104/pp.103.036970

Kuromori, T., Miyaji, T., Yabuuchi, H., Shimizu, H., Sugimoto, E., Kamiya, A., et al. (2010). ABC transporter AtABCG25 is involved in abscisic acid transport and responses. *Proc. Natl. Acad. Sci. U. S. A.* 107, 2361–2366. doi: 10.1073/pnas.0912516107

Kuromori, T., Seo, M., and Shinozaki, K. (2018). ABA transport and plant water stress responses. *Trends Plant Sci.* 23, 513–522. doi: 10.1016/j.tplants.2018.04.001

Laur, J., and Hacke, U. G. (2014). The role of water channel proteins in facilitating recovery of leaf hydraulic conductance from water stress in *Populus trichocarpa*. *PLoS ONE* 9:e111751. doi: 10.1371/journal.pone.0111751

Laurance, W. F. (1998). A crisis in the making: responses of Amazonian forests to land use and climate change. *Trends Ecol. Evol.* 13, 411–415. doi: 10.1016/S0169-5347(98)01433-5

Laurance, W. F. (2004). Forest-climate interactions in fragmented tropical landscapes. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 359, 345–352. doi: 10.1098/rstb.2003.1430

Leuschner, C., and Meier, I. C. (2018). The ecology of Central European tree species: Trait spectra, functional trade-offs, and ecological classification of adult trees. *Perspect. Plant Ecol. Syst.* 33, 89–103. doi: 10.1016/j.ppees.2018.05.003

Levitt, J. (1980). Responses of Plants to Environmental Stress: *Chilling, Freezing, and High Temperature Stresses. Responses of Plants to Environmental Stress, 2nd Edn, Vol. 1* Available online at: https://www.cabdirect.org/cabdirect/abstract/19802605739

Li, D.-D., Xia, X.-L., Yin, W.-L., and Zhang, H.-C. (2013). Two poplar calcineurin B-like proteins confer enhanced tolerance to abiotic stresses in transgenic *Arabidopsis thaliana*. *Biol. Plant.* 57, 70–78. doi: 10.1007/s10535-012-0251-7

Li, T., Sun, J., Bi, Y., and Peng, Z. (2016). Overexpression of an MYB-Related gene *FvMYB1* from *Fraxinus velutina* increases tolerance to salt stress in transgenic tobacco. *J. Plant Growth Regul.* 35, 632–645. doi: 10.1007/s00344-015-9565-y

Li, Y., Zhao, M., Motesharrei, S., Mu, Q., Kalnay, E., and Li, S. (2015). Local cooling and warming effects of forests based on satellite observations. *Nat. Commun.* 6:6603. doi: 10.1038/ncomms7603

Liu, C., Wei, C., Zhang, M., Xu, Y., Xiang, Z., and Zhao, A. (2017). Mulberry *MnMAPK1*, a group C mitogen-activated protein kinase gene, endowed transgenic *Arabidopsis* with novel responses to various abiotic stresses. *Plant Cell Tiss. Organ Cult.* 131, 151–162. doi: 10.1007/s11240-017-1272-x

Liu, S., Wang, C., Jia, F., An, Y., Liu, C., Xia, X., et al. (2016). Secretory peptide PdEPF2 enhances drought tolerance by modulating stomatal density and regulates ABA response in transgenic *Arabidopsis thaliana*. *Plant Cell Tiss. Organ Cult.* 125, 419–431. doi: 10.1007/s11240-016-0957-x

Liu, X.-Q., Liu, C.-Y., Guo, Q., Zhang, M., Cao, B.-N., Xiang, Z.-H., et al. (2015). Mulberry transcription factor MnDREB4A confers tolerance to multiple abiotic stresses in transgenic tobacco. *PLoS ONE* 10:e0145619. doi: 10.1371/journal.pone.0145619

Lorenz, W. W., Alba, R., Yu, Y.-S., Bordeaux, J. M., Simões, M., and Dean, J. F. (2011). Microarray analysis and scale-free gene networks identify candidate regulators in drought-stressed roots of loblolly pine (*P. taeda L.*). *BMC Genomics* 12:264. doi: 10.1186/1471-2164-12-264

Lu, X., Zhang, X., Duan, H., Lian, C., Liu, C., Yin, W., et al. (2017). Three stress-responsive NAC transcription factors from *Populus euphratica* differentially regulate salt and drought tolerance in transgenic plants. *Physiol. Plant.* 162, 73–97. doi: 10.1111/ppl.12613

Luo, Z. B., Janz, D., Jiang, X., Göbel, C., Wildhagen, H., Tan, Y., et al. (2009). Upgrading root physiology for stress tolerance by ectomycorrhizas: insights

from metabolite and transcriptional profiling into reprogramming for stress anticipation. *Plant Physiol.* 151, 1902–1917. doi: 10.1104/pp.109.143735

Ma, Q., Yue, L.-J., Zhang, J.-L., Wu, G.-Q., Bao, A.-K., and Wang, S.-M. (2012). Sodium chloride improves photosynthesis and water status in the succulent xerophyte *Zygophyllum xanthoxylum*. *Tree Physiol.* 32, 4–13. doi: 10.1093/treephys/tpr098

Ma, T., Wang, J., Zhou, G., Yue, Z., Hu, Q., Chen, Y., et al. (2013). Genomic insights into salt adaptation in a desert poplar. *Nat. Commun.* 4:2797. doi: 10.1038/ncomms3797

Malone, S. L. (2017). Monitoring changes in water use efficiency to understand drought induced tree mortality. *Forests* 8:365. doi: 10.3390/f8100365

Marron, N., Delay, D., Petit, J.-M., Dreyer, E., Kahlem, G., Delmotte, F. M., et al. (2002). Physiological traits of two *Populus × euramericana* clones, Luisa Avanzo and Dorskamp, during a water stress and re-watering cycle. *Tree Physiol.* 22, 849–858. doi: 10.1093/treephys/22.12.849

Marron, N., Dreyer, E., Boudouresque, E., Delay, D., Petit, J.-M., Delmotte, F. M., et al. (2003). Impact of successive drought and re-watering cycles on growth and specific leaf area of two *Populus x canadensis* (Moench) clones, "Dorskamp" and "Luisa_Avanzo." *Tree Physiol.* 23, 1225–1235. doi: 10.1093/treephys/23.18.1225

Martorell, S., Medrano, H., Tomàs, M., Escalona, J. M., Flexas, J., and Diaz-Espejo, A. (2015). Plasticity of vulnerability to leaf hydraulic dysfunction during acclimation to drought in grapevines: an osmotic-mediated process. *Physiol. Plant.* 153, 381–391. doi: 10.1111/ppl.12253

McDowell, N., Pockman William, T., Allen Craig, D., Breshears David, D., Cobb, N., Kolb, T., et al. (2008). Mechanisms of plant survival and mortality during drought: why do some plants survive while others succumb to drought? *New. Phytol.* 178, 719–739. doi: 10.1111/j.1469-8137.2008.02436.x

Mishra, B., Gupta, D. K., Pfenninger, M., Hickler, T., Langer, E., Nam, B., et al. (2018). A reference genome of the European beech (*Fagus sylvatica L.*). *Gigascience* 7:giy063. doi: 10.1093/gigascience/giy063

Molina-Rueda, J. J., and Kirby, E. G. (2015). Transgenic poplar expressing the pine GS1a show alterations in nitrogen homeostasis during drought. *Plant Physiol. Biochem.* 94, 181–190. doi: 10.1016/j.plaphy.2015.06.009

Monclus, R., Dreyer, E., Villar, M., Delmotte, F. M., Delay, D., Petit, J.-M., et al. (2006). Impact of drought on productivity and water use efficiency in 29 genotypes of *Populus deltoides×Populus nigra*. *New. Phytol.* 169, 765–777. doi: 10.1111/j.1469-8137.2005.01630.x

Movahedi, A., Zhang, J. X., Gao, P. H., Yang, Y., Wang, L. K., Yin, T .M., et al. (2015). Expression of the chickpea *CarNAC3* gene enhances salinity and drought tolerance in transgenic poplars. *Plant Cell Tissue Organ Cult.* 120, 141–154. doi: 10.1007/s11240-014-0588-z

Munné-Bosch, S., and Alegre, L. (2004). Die and let live: leaf senescence contributes to plant survival under drought stress. *Funct. Plant Biol.* 31, 203–216. doi: 10.1071/fp03236

Naser, L., Kourosh, V., Bahman, K., and Reza, A. (2010). Soluble sugars and proline accumulation play a role as effective indices for drought tolerance screening in Persian walnut (*Juglans regia L.*) during germination. *Fruits* 65, 97–112. doi: 10.1051/fruits/20010005

Nguyen, A., and Lamant, A. (1989). Variation in growth and osmotic regulation of roots of water-stressed maritime pine (*Pinus pinaster Ait.*) provenances. *Tree Physiol.* 5, 123–133.

Nguyen, H. T., Meir, P., Sack, L., Evans, J. R., Oliveira, R. S., and Ball, M. C. (2017). Leaf water storage increases with salinity and aridity in the mangrove *Avicennia marina*: integration of leaf structure, osmotic adjustment and access to multiple water sources. *Plant Cell Environ.* 40, 1576–1591. doi: 10.1111/pce.12962

NSIDC (2018). *National Snow Ice and Data Center.* Available online at: https://nsidc.org/arcticseaicenews/

Orlovic, S., Guzina, V., Krstic, B., and Merkulov, L. (1998). Genetic variability in anatomical, physiological and growth characteristics of hybrid poplar (*Populus × euramericana Dode (Guinier)*) and eastern cottonwood (*Populus deltoides Bartr.*) clones. *Silvae Genet.* 47, 183–190.

Ottow, E. A., Brinker, M., Teichmann, T., Fritz, E., Kaiser, W., Brosch,é, M., et al. (2005). *Populus euphratica* displays apoplastic sodium accumulation, osmotic adjustment by decreases in calcium and soluble carbohydrates, and develops leaf succulence under salt stress. *Plant Physiol.* 139:1762. doi: 10.1104/pp.105.069971

Pallara, G., Giovannelli, A., Traversi, M. L., Camussi, A., and Racchi, M. L. (2012). Effect of water deficit on expression of stress-related genes in the cambial region of two contrasting poplar clones. *J. Plant Growth Regul.* 31, 102–112. doi: 10.1007/s00344-011-9223-y

Parent, B., Hachez, C., Redondo, E., Simonneau, T., Chaumont, F., and Tardieu, F. (2009). Drought and abscisic acid effects on aquaporin content translate into changes in hydraulic conductivity and leaf growth rate: a trans-scale approach. *Plant Physiol.* 149, 2000–2012. doi: 10.1104/pp.108.130682

Parker, J. (1956). Drought resistance in woody plants. *Bot. Rev.* 22, 241–289.

Passioura, J. B. (2002). Soil conditions and plant growth. *Plant Cell Environ.* 25, 311–318. doi: 10.1046/j.0016-8025.2001.00802.x

Paul, S., Wildhagen, H., Janz, D., and Polle, A. (2018). Drought effects on the tissue- and cell-specific cytokinin activity in poplar. *AoB PLANTS* 10:plx067. doi: 10.1093/aobpla/plx067

Perdiguero, P., Collada, C., Barbero, M., del, C., García Casado, G., Cervera, M. T., et al. (2012). Identification of water stress genes in *Pinus pinaster* Ait. by controlled progressive stress and suppression-subtractive hybridization. *Plant Physiol. Biochem.* 50, 44–53. doi: 10.1016/j.plaphy.2011.09.022

Peter, M., Kohler, A., Ohm, R. A., Kuo, A., Krützmann, J., Morin, E., et al. (2016). Ectomycorrhizal ecology is imprinted in the genome of the dominant symbiotic fungus *Cenococcum geophilum*. *Nat. Commun.* 7:12662. doi: 10.1038/ncomms12662

Plomion, C., Aury, J.-M., Amselem, J., Leroy, T., Murat, F., Duplessis, S., et al. (2018). Oak genome reveals facets of long lifespan. *Nat. Plants* 4, 440–452. doi: 10.1038/s41477-018-0172-3

Polle, A., Altman, A., and Jiang, X. (2006). "Towards genetic engineering for drought tolerance in trees," in *Tree Transgenics*, eds M. Fladung and D. Ewald (Berlin, Heidelberg: Springer), 275–297.

Polle, A., and Chen, S. (2015). On the salty side of life: molecular, physiological and anatomical adaptation and acclimation of trees to extreme habitats. *Plant Cell Environ.* 38, 1794–1816. doi: 10.1111/pce.12440

Rennenberg, H., Loreto, F., Polle, A., Brilli, F., Fares, S., Beniwal, R. S., et al. (2006). Physiological responses of forest trees to heat and drought. *Plant Biol.* 8, 556–571. doi: 10.1055/s-2006-924084

Reyer, C. P. O., Brouwers, N., Rammig, A., Brook, B. W., Epila, J., Grant, R. F., et al. (2015). Forest resilience and tipping points at different spatio-temporal scales: approaches and challenges. *J. Ecol.* 103, 5–15. doi: 10.1111/1365-2745.12337

Roman, D. T., Novick, K. A., Brzostek, E. R., Dragoni, D., Rahman, F., and Phillips, R. P. (2015). The role of isohydric and anisohydric species in determining ecosystem-scale response to severe drought. *Oecologia* 179, 641–654. doi: 10.1007/s00442-015-3380-9

Schaller, N., Sparrow, S. N., Massey, N. R., Bowery, A., Miller, J., Wilson, S., et al. (2018). Ensemble of European regional climate simulations for the winter of 2013 and 2014 from HadAM3P-RM3P. *Sci. Data* 5:180057. doi: 10.1038/sdata.2018.57

Scholz, F. G., Phillips, N. G., Bucci, S. J., Meinzer, F. C., and Goldstein, G. (2011). "Hydraulic capacitance: biophysics and functional significance of internal water sources in relation to tree size," in *Size- and Age-Related Changes in Tree Structure and Function Tree Physiology*, eds F. C. Meinzer, B. Lachenbruch, and T. E. Dawson (Dordrecht: Springer), 341–361.

Secchi, F., and Zwieniecki, M. A. (2011). Sensing embolism in xylem vessels: the role of sucrose as a trigger for refilling. *Plant Cell Environ.* 34, 514–524. doi: 10.1111/j.1365-3040.2010.02259.x

Shi, Y., Li, J., Shao, J., Deng, S., Wang, R., Li, N., et al. (2010). Effects of Stockosorb and Luquasorb polymers on salt and drought tolerance of *Populus popularis*. *Sci. Horticulturae* 124, 268–273. doi: 10.1016/j.scienta.2009.12.031

Skelton, R. P., West, A. G., and Dawson, T. E. (2015). Predicting plant vulnerability to drought in biodiverse regions using functional traits. *Proc. Natl. Acad. Sci. U. S. A.* 112, 5744–5749. doi: 10.1073/pnas.1503376112

Smart, L. B., Moskal, W. A., Cameron, K. D., and Bennett, A. B. (2001). MIP genes are down-regulated under drought stress in *Nicotiana glauca*. *Plant Cell Physiol.* 42, 686–693. doi: 10.1093/pcp/pce085

Sperry, J. S., and Love, D. M. (2015). What plant hydraulics can tell us about responses to climate-change droughts. *New. Phytol.* 207, 14–27. doi: 10.1111/nph.13354

Stålfelt, M. G. (1955). The stomata as a hydrophotic regulator of the water deficit of the plant. *Physiol. Plant.* 8, 572–593. doi: 10.1111/j.1399-3054.1955.tb07753.x

Sun, H., Li, L., Lou, Y., Zhao, H., Yang, Y., and Gao, Z. (2016). Cloning and preliminary functional analysis of *PeUGE* gene from moso bamboo (*Phyllostachys edulis*). *DNA Cell Biol.* 35, 706–714. doi: 10.1089/dna.2016.3389

Sun, H., Li, L., Lou, Y., Zhao, H., Yang, Y., Wang, S., et al. (2017). The bamboo aquaporin gene *PeTIP4;1–1*confers drought and salinity tolerance in transgenic *Arabidopsis*. *Plant Cell Rep.* 36, 597–609. doi: 10.1007/s00299-017-2106-3

Sussmilch, F. C., Brodribb, T. J., and McAdam, S. A. M. (2017). What are the evolutionary origins of stomatal responses to abscisic acid (ABA) in land plants? *J. Integr. Plant Biol.* 59, 240–260. doi: 10.1111/jipb.12523

Szechynska-Hebda, M., Czarnocka, W., Hebda, M., and Karpinski, S. (2016). *PAD4*, *LSD1* and *EDS1* regulate drought tolerance, plant biomass production, and cell wall properties. *Plant Cell Rep.* 35, 527–539. doi: 10.1007/s00299-015-1901-y

Tan, B.-C., Joseph, L. M., Deng, W.-T., Liu, L., Li, Q.-B., Cline, K., et al. (2003). Molecular characterization of the *Arabidopsis* 9-cis epoxycarotenoid dioxygenase gene family. *Plant J.* 35, 44–56. doi: 10.1046/j.1365-313X.2003.01786.x

Tardieu, F., and Simonneau, T. (1998). Variability among species of stomatal control under fluctuating soil water status and evaporative demand: modelling isohydric and anisohydric behaviours. *J. Exp. Bot.* 49, 419–432. doi: 10.1093/jxb/49.Special_Issue.419

Tian, Q., Chen, J., Wang, D., Wang, H.-L., Liu, C., Wang, S., et al. (2017). Overexpression of a *Populus euphratica CBF4* gene in poplar confers tolerance to multiple stresses. *Plant Cell Tiss. Organ Cult.* 128, 391–407. doi: 10.1007/s11240-016-1118-y

Tischer, S. V., Wunschel, C., Papacek, M., Kleigrewe, K., Hofmann, T., Christmann, A., et al. (2017). Combinatorial interaction network of abscisic acid receptors and coreceptors from *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U. S. A.* 114, 10280–10285. doi: 10.1073/pnas.1706593114

Tung, S. A., Smeeton, R., White, C. A., Black, C. R., Taylor, I. B., Hilton, H. W., et al. (2008). Over-expression of *LeNCED1* in tomato (*Solanum lycopersicum L.*) with the rbcS3C promoter allows recovery of lines that accumulate very high levels of abscisic acid and exhibit severe phenotypes. *Plant Cell Environ.* 31, 968–981. doi: 10.1111/j.1365-3040.2008.01812.x

Tyree, M. T., and Ewers, F. W. (1991). The hydraulic architecture of trees and other woody plants. *New. Phytologist* 119, 345–360. doi: 10.1111/j.1469-8137.1991.tb00035.x

van der Werf, G. R., Dempewolf, J., Trigg, S. N., Randerson, J. T., Kasibhatla, P. S., Giglio, L., et al. (2008). Climate regulation of fire emissions and deforestation in equatorial Asia. *Proc. Natl. Acad. Sci. U. S. A.* 105, 20350–20355. doi: 10.1073/pnas.0803375105

Wang, C., Wang, L., Yang, C., and Wang, Y. (2017a). Identification, phylogeny, and transcript profiling of aquaporin genes in response to abiotic stress in *Tamarix hispida*. *Tree Genet. Genomes* 13:81. doi: 10.1007/s11295-017-1163-7

Wang, C., Yang, Y., Wang, H., Ran, X., Li, B., Zhang, J., et al. (2016). Ectopic expression of a cytochrome P450 monooxygenase gene *PtCYP714A3* from *Populus trichocarpa* reduces shoot growth and improves tolerance to salt stress in transgenic rice. *Plant Biotechnol. J.* 14, 1838–1851. doi: 10.1111/pbi.12544

Wang, C., Zhi, S., Liu, C., Xu, F., Zhao, A., Wang, X., et al. (2017b). Characterization of stilbene synthase genes in mulberry (*Morus atropurpurea*) and metabolic engineering for the production of resveratrol in *Escherichia coli*. *J. Agric. Food Chem.* 65, 1659–1668. doi: 10.1021/acs.jafc.6b05212

Wang, J. Y., Wang, J. P., and He-Yuan (2013). A *Populus euphratica* NAC protein regulating Na$^+$/K$^+$ homeostasis improves salt tolerance in *Arabidopsis thaliana*. *Gene* 521, 265–273. doi: 10.1016/j.gene.2013.03.068

Wang, L., Li, Z., Lu, M., and Wang, Y. (2017). ThNAC13, a NAC transcription factor from *Tamarix hispida*, confers salt and osmotic stress tolerance to transgenic *Tamarix* and *Arabidopsis*. *Front. Plant Sci.* 8:635. doi: 10.3389/fpls.2017.00635

Wang, L., Zhao, H., Chen, D., Li, L., Sun, H., Lou, Y., et al. (2016). Characterization and primary functional analysis of a bamboo NAC gene targeted by miR164b. *Plant Cell Rep.* 35, 1371–1383. doi: 10.1007/s00299-016-1970-6

Wang, S., Yao, W., Wei, H., Jiang, T., and Zhou, B. (2014). Expression patterns of *ERF* genes underlying abiotic stresses in di-haploid *Populus simonii* × *P. nigra*. *Sci. World J.* 2014:745091. doi: 10.1155/2014/745091

Wang, W., Vinocur, B., and Altman, A. (2003). Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. *Planta* 218, 1–14. doi: 10.1007/s00425-003-1105-5

Wang, Y., Gao, C., Liang, Y., Wang, C., Yang, C., and Liu, G. (2010). A novel *bZIP* gene from *Tamarix hispida* mediates physiological responses to salt stress in tobacco plants. *J. Plant Physiol.* 167, 222–230. doi: 10.1016/j.jplph.2009.09.008

Wasilewska, A., Vlad, F., Sirichandra, C., Redko, Y., Jammes, F., Valon, C., et al. (2008). An update on abscisic acid signaling in plants and more. *Mol. Plant* 1, 198–217. doi: 10.1093/mp/ssm022

Wildhagen, H., Paul, S., Allwright, M., Smith, H. K., Malinowska, M., Schnabel, S. K., et al. (2018). Genes and gene clusters related to genotype and drought-induced variation in saccharification potential, lignin content and wood anatomical traits in *Populus nigra*. *Tree Physiol.* 38, 320–339. doi: 10.1093/treephys/tpx054

Wilkins, O., Waldron, L., Nahal, H., Provart, N. J., and Campbell, M. M. (2009). Genotype and time of day shape the *Populus* drought response. *Plant J.* 60, 703–715. doi: 10.1111/j.1365-313X.2009.03993.x

Wolfe, B. T., Sperry, J. S., and Kursar, T. A. (2016). Does leaf shedding protect stems from cavitation during seasonal droughts? A test of the hydraulic fuse hypothesis. *New. Phytol.* 212, 1007–1018. doi: 10.1111/nph.14087

Xing, H. T., Guo, P., Xia, X. L., and Yin, W. L. (2011). PdERECTA, a leucine-rich repeat receptor-like kinase of poplar, confers enhanced water use efficiency in *Arabidopsis*. *Planta* 234, 229–241. doi: 10.1007/s00425-011-1389-9

Xiu, Y., Iqbal, A., Zhu, C., Wu, G., Chang, Y., Li, N., et al. (2016). Improvement and transcriptome analysis of root architecture by overexpression of *Fraxinus pennsylvanica* DREB2A transcription factor in *Robinia pseudoacacia L. 'Idaho.'* *Plant Biotechnol. J.* 14, 1456–1469. doi: 10.1111/pbi.12509

Xu, C., Fu, X., Liu, R., Guo, L., Ran, L., Li, C., et al. (2017). *PtoMYB170* positively regulates lignin deposition during wood formation in poplar and confers drought tolerance in transgenic *Arabidopsis*. *Tree Physiol.* 37, 1713–1726. doi: 10.1093/treephys/tpx093

Xu, H., Kemppainen, M., El Kayal, W., Lee, S. H., Pardo, A. G., Cooke, J. E. K., et al. (2015). Overexpression of *Laccaria bicolor* aquaporin JQ585595 alters root water transport properties in ectomycorrhizal white spruce (*Picea glauca*) seedlings. *New. Phytol.* 205, 757–770. doi: 10.1111/nph.13098

Yang, G., Yu, L., Zhang, K., Zhao, Y., Guo, Y., and Gao, C. (2017). A *ThDREB* gene from *Tamarix hispida* improved the salt and drought tolerance of transgenic tobacco and *T. hispida*. *Plant Physiol. Biochem.* 113, 187–197. doi: 10.1016/j.plaphy.2017.02.007

Yang, Y., Tang, R. J., Li, B., Wang, H. H., Jin, Y. L., Jiang, C. M., et al. (2015). Overexpression of a *Populus trichocarpa* H$^+$-pyrophosphatase gene PtVP1.1 confers salt tolerance on transgenic poplar. *Tree Physiol.* 35, 663–677. doi: 10.1093/treephys/tpv027

Yao, W., Wang, S., Zhou, B., and Jiang, T. (2016). Transgenic poplar overexpressing the endogenous transcription factor ERF76 gene improves salinity tolerance. *Tree Physiol.* 36, 896–908. doi: 10.1093/treephys/tpw004

Yu, J., Ge, H., Wang, X., Tang, R., Wang, Y., Zhao, F., et al. (2017). Overexpression of pyrabactin resistance-like abscisic acid receptors enhances drought, osmotic, and cold tolerance in transgenic poplars. *Front. Plant Sci.* 8:1752. doi: 10.3389/fpls.2017.01752

Yu, J., Yang, L., Liu, X., Tang, R., Wang, Y., Ge, H., et al. (2016). Overexpression of poplar pyrabactin resistance-like abscisic acid receptors promotes abscisic acid sensitivity and drought resistance in transgenic *Arabidopsis*. *PLoS ONE* 11:e0168040. doi: 10.1371/journal.pone.0168040

Yu, L. H., Wu, S. J., Peng, Y. S., Liu, R. N., Chen, X., Zhao, P., et al. (2016). *Arabidopsis EDT1/HDG11* improves drought and salt tolerance in cotton and poplar and increases cotton yield in the field. *Plant Biotechnol. J* 14, 72–84. doi: 10.1111/pbi.12358

Yu, X., Ohtani, M., Kusano, M., Nishikubo, N., Uenoyama, M., Umezawa, T., et al. (2017). Enhancement of abiotic stress tolerance in poplar by overexpression of key *Arabidopsis* stress response genes, AtSRK2C and AtGolS2. *Mol. Breeding* 37:57. doi: 10.1007/s11032-016-0618-0

Zang, D., Wang, C., Ji, X., and Wang, Y. (2015). *Tamarix hispida* zinc finger protein ThZFP1 participates in salt and osmotic stress tolerance by increasing proline content and SOD and POD activities. *Plant Sci.* 235, 111–121. doi: 10.1016/j.plantsci.2015.02.016

Zhang, H., Zhu, H., Pan, Y., Yu, Y., Luan, S., and Li, L. (2014). A DTX/MATE-type transporter facilitates abscisic acid efflux and modulates ABA sensitivity and drought tolerance in *Arabidopsis*. *Mol. Plant* 7, 1522–1532. doi: 10.1093/mp/ssu063

Zhang, J., Movahedi, A., Sang, M., Wei, Z., Xu, J., Wang, X., et al. (2017). Functional analyses of NDPK2 in *Populus trichocarpa* and overexpression of PtNDPK2 enhances growth and tolerance to abiotic stresses in transgenic poplar. *Plant Physiol. Biochem.* 117, 61–74. doi: 10.1016/j.plaphy.2017.05.019

Zhang, T., Zhang, D., Liu, Y., Luo, C., Zhou, Y., and Zhang, L. (2015). Overexpression of a NF-YB3 transcription factor from *Picea wilsonii* confers tolerance to salinity and drought stress in transformed *Arabidopsis thaliana*. *Plant Physiol. Biochem.* 94, 153–164. doi: 10.1016/j.plaphy.2015.05.001

Zheng, L., Chen, S., Xie, L., Lu, Z., Liu, M., Han, X., et al. (2018). Overexpression of cysteine protease gene from *Salix matsudana* enhances salt tolerance in transgenic *Arabidopsis*. *Environ. Exp. Bot.* 147, 53–62. doi: 10.1016/j.envexpbot.2017.11.008

Zheng, L., Liu, G., Meng, X., Liu, Y., Ji, X., Li, Y., et al. (2013). A WRKY gene from *Tamarix hispida*, *ThWRKY4*, mediates abiotic stress responses by modulating reactive oxygen species and expression of stress-responsive genes. *Plant Mol. Biol.* 82, 303–320. doi: 10.1007/s11103-013-0063-y

Zhou, M. L., Ma, J. T., Zhao, Y. M., Wei, Y. H., Tang, Y. X., and Wu, Y. M. (2012). Improvement of drought and salt tolerance in *Arabidopsis* and *Lotus corniculatus* by overexpression of a novel DREB transcription factor from *Populus euphratica*. *Gene* 506, 10–17. doi: 10.1016/j.gene.2012.06.089

Cyberinfrastructure to Improve Forest Health and Productivity: The Role of Tree Databases in Connecting Genomes, Phenomes, and the Environment

Jill L. Wegrzyn[1]*, Margaret A. Staton[2]*, Nathaniel R. Street[3]*, Dorrie Main[4]*, Emily Grau[1]*, Nic Herndon[1], Sean Buehler[1], Taylor Falk[1], Sumaira Zaman[1], Risharde Ramnath[1], Peter Richter[1], Lang Sun[1], Bradford Condon[2], Abdullah Almsaeed[2], Ming Chen[2], Chanaka Mannapperuma[3], Sook Jung[4] and Stephen Ficklin[4]

[1] Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT, United States, [2] Department of Entomology and Plant Pathology, University of Tennessee, Knoxville, Knoxville, TN, United States, [3] Umeå Plant Science Centre, Department of Plant Physiology, Umeå University, Umeå, Sweden, [4] Department of Horticulture, Washington State University, Pullman, WA, United States

Despite tremendous advancements in high throughput sequencing, the vast majority of tree genomes, and in particular, forest trees, remain elusive. Although primary databases store genetic resources for just over 2,000 forest tree species, these are largely focused on sequence storage, basic genome assemblies, and functional assignment through existing pipelines. The tree databases reviewed here serve as secondary repositories for community data. They vary in their focal species, the data they curate, and the analytics provided, but they are united in moving toward a goal of centralizing both data access and analysis. They provide frameworks to view and update annotations for complex genomes, interrogate systems level expression profiles, curate data for comparative genomics, and perform real-time analysis with genotype and phenotype data. The organism databases of today are no longer simply catalogs or containers of genetic information. These repositories represent integrated cyberinfrastructure that support cross-site queries and analysis in web-based environments. These resources are striving to integrate across diverse experimental designs, sequence types, and related measures through ontologies, community standards, and web services. Efficient, simple, and robust platforms that enhance the data generated by the research community, contribute to improving forest health and productivity.

Keywords: database, content management system, forest tree, bioinformatics, web services

INTRODUCTION

Starting in the Sanger sequencing era, significant investments were made to catalog genetic resources in primary repositories (Frishman et al., 1998). EMBL (the European Molecular Biology Laboratory), DDBJ (the DNA Data Bank of Japan), and NCBI (the National Center for Biotechnology Information) GenBank were initiated between 1980 and 1992, and remain freely

accessible and federally funded (Benson et al., 1997; Tateno and Gojobori, 1997). The vast majority of data for these large, sequence-centric databases is sourced directly from researcher submissions that are encouraged through peer review journals. These primary resources have evolved with the data collection and curation needs of today, expanding in terms of both the sequence source and the associated metadata (Sayers et al., 2019). All three specialize in generating persistent identifiers to track a single sequence over an extensive network of resources. A genic identifier, as an example, may link a reference genome in NCBI's Genome, an expression value in the Gene Expression Omnibus (GEO), and support for a UniRef90 cluster. These uniquely accessioned resources are increasingly integrated into secondary and tertiary repositories that subset or enhance these accessions with data specific to the communities they serve (Herrero et al., 2016).

As the data types and experimental designs contributing to these repositories diversified, a plethora of model organism databases (MODs) or clade organism databases (CODs) emerged. These databases sought to provide unique resources for the research communities they serve, through layered curation and specialized integration. The AAtDB (An *Arabidopsis thaliana* Database), developed in 1991 to support the first model plant system, has since evolved into the widely accessed, Arabidopsis Information Resource (TAIR) (Flanders et al., 1998; Huala et al., 2001). Around the same time, USDA-ARS funds were dedicated to developing some of the first informatic portals for economically important crop species, including RiceGenes (Cartinhour, 1997), GrainGenes (Triticeae) (Carollo et al., 2005), MaizeGDB (Lawrence et al., 2004), SoyBase (Grant et al., 2009), and the Dendrome Project for forest trees (Wegrzyn et al., 2008). Some of these databases remain independent funded entities, while others have merged into larger repositories or broadened their scope. There are hundreds of plant-focused organismal databases acting as secondary repositories today (Lai et al., 2012; Chen et al., 2018). The vast majority have moved beyond genetics and genomics data, providing advanced integration through stock centers, phenotypic evaluation, breeding resources, and metabolomic pathway integration.

Forest trees are unique among species represented in crop databases. The vast majority are long-lived, outcrossing, with extensive natural distributions represented by large, diverse and locally adapted populations (Holliday et al., 2017). They represent species of economic importance and are used for paper, pulp, biofuels, food, and timber production. At the same time, they serve as a foundation for watersheds, biodiversity, and contribute substantially to carbon sequestration with forests covering roughly 30% of the earth's surface (Houghton, 2005). Like many plants, forest trees have complex genomes with challenges associated with ploidy and repetitive content. Additionally, gymnosperm tree genomes are exceptionally large, ranging from 10 to 40 Gbp in size (De La Torre et al., 2014). This, in combination with the need to broadly sample these large and diverse populations, yields limited full genome representation. Of the nearly 60,000 forest tree species, less than 35 are associated with an assembled and annotated reference genome (Neale et al., 2013; Plomion et al., 2016). A view into our primary

databases reveal that over 2,000 species are associated with genetic information that is of value to the research community (Sayers et al., 2019).

In the era of high throughput technologies that assess genotype, phenotype, and environmental metrics much faster than we can conceive, the need for well structured, efficient, and well-connected databases is apparent. In a recent survey of over 700 investigators in the biological sciences, access to analytical frameworks, long-term storage, and the ability to integrate data across disparate sources, were of primary concern (Barone et al., 2017). The long-lived nature of trees, their pivotal role in local economies, and role in ecosystem health requires an integrated approach that must leverage datasets from a variety of sources. To meet the needs of the research community, forest tree databases are moving away from independent database structures and toward integrated Content Management Systems (CMS) that support specific, shareable modules for query and analysis (Ficklin et al., 2011; Sundell et al., 2015). There is less focus on standardized database backends since web services allow users to expose their data, and data definitions. The ability to provide this, and cross-site query, relies heavily on the curation of ontologies to describe data housed in these frameworks. Initiatives that curate woody plant ontologies, to describe plant structures and measured traits, are critical components of forest tree database cross-talk (Jaiswal et al., 2005; Lens et al., 2012; Cooper et al., 2018). When these terms are provided within the framework of recommended standards for data collection, such as the Minimal Information About a Plant Phenotyping Experiment (MIAPPE), opportunities for analytical pipelines to evaluate complex data becomes a reality (Krajewski et al., 2015). In addition to these standards, all tree databases are integrated with existing analytical frameworks, such as Galaxy, that support and expose common bioinformatic workflows (Afgan et al., 2018). In this review, four tree databases are described, including their history, scope, current resources, and analytic tools (**Figure 1**).

TREEGENES

TreeGenes[1], previously known as Dendrome, was initially constructed to provide access to genetic data for forest trees in a relational framework. Early development included curation of molecular markers, genetic maps, Expressed Sequence Tags (ESTs), and species range maps. TreeGenes remained in a custom database schema and website, adopting components of the Generic Model Organism (GMOD) framework for housing genetic maps (cMAP) and genome assemblies (JBrowse) (Stein et al., 2002). Later development focused on the integration of genotyping resources, phenotypes, expression studies, and additional reference genome sequences (Wegrzyn et al., 2012; Falk et al., 2018).

TreeGenes currently represents just over 1700 species from 16 orders and 124 genera (Falk et al., 2018). TreeGenes has 1200 registered users with associated colleague accounts that enable access to data submission and analytical pipelines.

---

[1] http://www.treegenesdb.org

FIGURE 1 | PlantGenIE, TreeGenes, and Hardwood Genomics Project represent integrated web-based frameworks that rely on a combination of primary repositories, secondary plant comparative databases, and user submissions to provide further value through data curation, integration, and analytics.
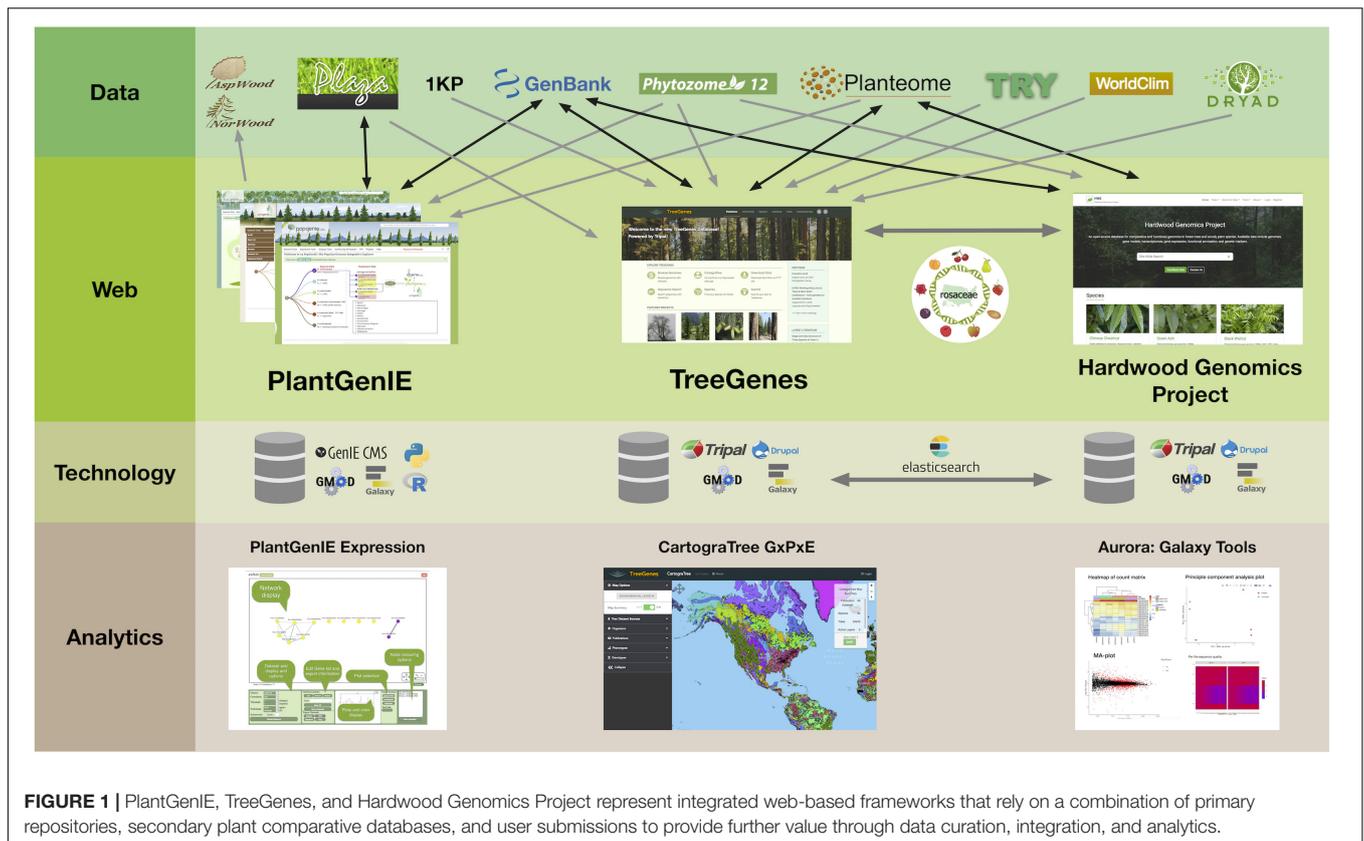
The database contains 27 reference genomes, 100 genetic maps, 36.7 M genotypes, 303 species with transcriptomes, 40 species with TreeGenes' Unigenes, 306 unique phenotype measures and 935,596 phenotypic measures. Genomic data is sourced primarily from GenBank, 1KP, Phytozome, and PLAZA (Goodstein et al., 2012; Matasci et al., 2014; Proost et al., 2015; Sayers et al., 2019). Phenotypic data is integrated from TRY-DB and Dryad, but primarily comes from direct user submissions (Kattge et al., 2011). Environmental data is extracted from imported layers, including temperature, precipitation, and solar radiation from WorldClim (Fick and Hijmans, 2017), and a variety of metrics from the Harmonized World Soil Survey Database (FAO et al., 2009).

TreeGenes is running on Tripal v3 which integrates a content management system known as Drupal with the Genetic Model Organism Database's (GMOD) relational schema, known as Chado. This conversation in 2017, aligned TreeGenes for the first time with over 30, primarily plant, databases (Ficklin et al., 2011; Sanderson et al., 2013). Recent focused development in Tripal, led by the tree and legume community, enabled cross-site communication, access to efficient data transfer, and the ability to interface with a local installation of Galaxy (Watts and Feltus, 2017; Wytko et al., 2017). Galaxy is an independent framework that provides an API to abstract command line informatic software, develop workflows, and connect to high performance computing resources. Conversion

into Tripal resulted in a complete overhaul of the database, and has enabled the development of several analytical modules that allow researchers to search, filter, and funnel data directly into supported workflows.

Following conversion, TreeGenes released a set of Tripal modules that can be utilized by researchers visiting the site or installed and customized for any Tripal supported databases. Tripal Sequence Similarity Search (TSeq) provides access to genomes, transcriptomes, and curated TreeGenes unigenes through traditional NCBI BLAST or optimized Diamond protein searches (Boratyn et al., 2013; Buchfink et al., 2015). The Tripal Plant PopGen Submit (TPPS) module presents a framework for researchers to submit their association genetics, landscape genomics, and related population studies by collecting any combination of molecular markers, phenotypes, and environmental measures. This module implements MIAPPE standards and the associated ontologies to ensure data integrity. The Tripal OrthoQuery module provides a framework for curating unigenes, executing OrthoFinder (Emms and Kelly, 2015), and generating interactive visualizations of gene families in a phylogenetic context. OrthoQuery enables both real-time orthologous gene family analysis and functional assessment of the resulting orthogroups.

Current development in TreeGenes focuses on CartograTree, which enables integration of genotype, phenotype and environmental data for georeferenced forest trees (Herndon et al., 2018). This module provides a robust framework

to query publication datasets, species, phenotypes, genotypes, and associations based on metadata collected at the time of submission. The data and metadata exposed in CartograTree is derived from published population level studies submitted to TreeGenes via TPPS or curated from Dryad. Landscape genomics, association genetics, and population structure analysis is executed through the Galaxy framework.

HARDWOOD GENOMICS PROJECT

The Hardwood Genomics Project (HWG) provides access to genomic resources generated from angiosperm trees, including forest and urban trees of ecological and agricultural significance. The resource originated from the Fagaceae Genomics Web, built in 2007, to house transcriptomes, genomes and genetic maps. As new collaborators joined the effort and the scope of species extended beyond Fagaceae, the site was rebuilt in 2011 as the HWG. HWG's mission is to host reference genomes and transcriptomes that are either not accessible elsewhere, or only available as raw files without an associated and searchable, functional annotation. In addition, HWG accepts molecular markers, genetic maps, germplasm and population descriptions, and community project descriptions. Current resources support species associated with pest or pathogen threats, including green ash, European ash, American chestnut, American beech, black walnut, and redbay, as well as trees with significant economic value, including white oak, black cherry, sugar maple, and tulip poplar.

For species with an available reference genome, HWG provides a workspace for accessing the annotation. This provides value to these sequence resources by performing and hosting functional annotation, including: identification of Open Reading Frames (ORFs) from transcripts, BLAST annotations derived from the Uniprot Swiss-Prot/TrEMBL plant protein databases, InterProScan domain searches (Jones et al., 2014), Gene Ontology (GO) term assignments (Ashburner et al., 2000), and predictions for Simple Sequence Repeats (SSRs) and primers. Researchers can download flat files, explore the spatial context of the assembly with JBrowse (Buels et al., 2016), search functional annotation for genes, and explore assigned GO terms through the ontology graphs. Additional genome specific data, such as alternative splicing, variants, and molecular markers are added to the JBrowse viewer when available.

Hardwood Genomics Project is currently running Tripal v3, and like TreeGenes, is responsible for the development of custom modules that can be installed on other Tripal-enabled sites. RNASeq data is poorly integrated in plant community databases despite the widespread use of expression studies to examine responses to biotic and abiotic stressors in plant systems. To address this limitation, HWG launched a framework devoted to the integration and analysis of gene expression experiments. BioSamples imported from GenBank, with the metadata describing the tissues, treatments, experimental design, and informatic methods, can be explored and compared. Each transcript, examined as part of an RNASeq experiment, has expression values that can be interrogated through interactive

visualizations or downloaded for further analysis. The expression data displays can be customized interactively, grouping BioSamples by their tagged metadata values. A tool for comparing gene expression is also available, allowing the user to provide their own gene list and generate a heat map comparing expression of those genes across the relevant BioSamples (Chen et al., 2017). Current development in HWG is focused on supporting bioinformatic workflows, through Galaxy, to allow users to load their own datasets for analysis. HWG has also developed an Elastic Search module that enables search engine style cross-site query. This enables the discovery of relevant datasets within and across Tripal-enabled websites. The Aurora Galaxy Tripal module allows informatic tools to be wrapped in R Markdown which makes it possible to generate Galaxy workflow outputs as static websites.

GENOME DATABASE FOR ROSACEAE

The Genome Database for Rosaceae (GDR[2], Jung et al., 2013) was initiated in 2003 to provide curated and integrated, genomic, genetics and breeding (GGB) data alongside analysis tools to enable basic, translational and applied research. Rosaceae is an economically, nutritionally and biologically important plant family that includes the majority of tree fruits (apple, apricot, blackberry, cherry, peach, plum), nuts (almond), and ornamentals (pear, crab apple). While not specifically focused on forest trees, GDR is included here for its role in developing Tripal modules for breeding and the comparative genomics utility with forest hardwoods.

GDR contains 21 genome assemblies and annotations for 14 species. A total of 528,890 genes, reference transcriptomes for all major species, 14,411 germplasm records, 313 genetic maps, 3.3 M molecular markers, 402,559 phenotype measurements, 3,902 QTL/MTL for 392 agronomic traits, 10.8 M genotypes, and 7,449 publications are housed in the database. GDR provides access to breeding management and analysis tools, pathway analysis through PlantCyc and Pathway Inspector, flexible front-end querying, genome annotations through JBrowse, and sequence similarity search functionality (Jung et al., 2016, 2017). GDR is participating in the development of new Tripal modules; visualization and analysis of genetic maps is available through the new Tripal Map Viewer module while whole genome alignments can be executed through the Tripal Synteny Viewer (Jung et al., 2018). GDR is currently expanding the analytic capabilities of their Breeding Information Management System (BIMS) and developing reference genome integration for the Tripal Map Viewer module.

PLANT GENOME INTEGRATIVE EXPLORER

The Plant Genome Integrative Explorer (PlantGenIE) began as The Populus Integrative Genome Explorer (PopGenIE;

---

[2]https://www.rosaceae.org

Sjödin et al., 2009), to overcome a lack of tools for routine tasks such as annotating gene lists, converting among sequence identifiers, and visualizing transcript abundance on the basis of EST sequencing. The Populus version was expanded to include visualization of poplar microarray data using the concept of the *Arabidopsis* electronic fluorescent pictograph (eFP) browser (Winter et al., 2007), gene set enrichment tests for Gene Ontology (Ashburner et al., 2000), Pfam (Finn et al., 2010), genome synteny browsing alongside sequence similarity searching. Later, a complimentary comparative co-expression tool was developed to facilitate inference of functional orthologs on the basis of conserved co-expression (Netotea et al., 2014). The resulting networks were integrated within *Populus* and *Arabidopsis* GenIE sites. With the release of the Norway spruce (*Picea abies*) genome (Nystedt et al., 2013), the associated resources were made available in a Conifer database, ConGenIE, which also includes genomes for loblolly pine (*Pinus taeda*; Neale et al., 2014; Wegrzyn et al., 2014) and white spruce (*Picea glauca*; Birol et al., 2013).

The PlantGenIE umbrella resource (Sundell et al., 2015), which included the development of new and updated gene expression tools, together with an integrated gene family analysis, is now available for all species. The primary aim is visualization of gene expression data, primarily from forest tree species, but including related sites for plant models, such as *Arabidopsis*. As such, gene expression resources for aspen (AspWood; Sundell et al., 2017) and Norway spruce (NorWood; Jokipii-Lukkari et al., 2017) cryogenic tangential cuttings series profiling wood development are being integrated within the PopGenIE and ConGenIE sites, respectively. Dedicated sites have been developed to provide access to spatial transcriptomics (Giacomello et al., 2017) and laser capture microdissection (Canas et al., 2014) gene expression data. For a subset of species, community annotation is also provided via WebApollo (Stevens et al., 2016).

The GenIE sites were originally developed using the Drupal CMS with many of the tools from GMOD (Stein et al., 2002). More recently, an alternative open-source CMS has been developed (GenIE-CMS[3]) and the PlantGenIE resource is currently being updated in this platform. GenIE-CMS includes web services, enabling end users to access genomic information from external interfaces such as R and Python analysis scripts. Alongside this update, new and improved versions of gene expression visualization tools have been developed and made available as plugins to GenIE-CMS. The PlantGenIE update includes integration with the PLAZA resource (Proost et al., 2015), integration of cross-GenIE gene lists using PLAZA gene orthology inference methods, new integrative gene expression explorer tools, new and updated gene expression networks inferred using seidr (Schiffthaler et al., 2018), and an updated functional enrichment tool. In addition to updating existing reference genomes, new genomes are being added, including a dedicated eucalyptus site, EucGenIE. The development of GenIE-CMS enables rapid and easy implementation of new GenIE

resources and cross-linking among existing GenIE sites using PLAZA gene family and orthology information.

FUTURE DIRECTIONS

Tree database cyberinfrastructure is supporting comparative genomics, population genetics, expression profiling, and genome annotation. These resources focus on a combination of model and non-model systems and integrate with established comparative resources to deliver value added information. Despite their importance, the sustainability of cyberinfrastructure and the related activities of curating and importing scientific data is always in question. The databases described here are leveraging larger open-source projects as their base framework and sharing web-based applications for common functionality, such as genome browsing and sequence similarity searching. For the forest tree community, the majority of the functionality described here has been deployed within the last 3 years and represents the first coordinated effort across these resources. Frameworks like Tripal and PlantGenIE focus on efficient deployment, web services for cross-talk, data visualization, and analytics to provide a robust environment for end users. As an example, the Elastic Search module developed by HWG allows one to search a gene, genome, marker, and other indexed objects in one database and locate results in other Tripal databases without executing independent searches on each website. Sharing development across a larger community allows forest tree databases to focus on the specific needs of their users. Their independent value exists in the additional curation, metadata acquisition, indexing, analytics, and visualization that is not delivered from the primary repositories. TreeGenes and Hardwood Genomics Web, are focused on expression data integration for non-model trees and metadata retrieval and cross-study analytics for population genetics studies. GDR is focused on improving access and visualization of genetic maps as well as breeding tools. The PlantGenIE framework is providing a robust platform for species with a reference genome, and advanced visualization for expression data that integrates across studies. All of these databases are also seeking stronger connections to more broadly plant focused repositories, such as Phytozome, PLAZA, and Planteome, that provide genetic and ontological resources that improve the utility of cross-site querying.

While tremendous advancements have been made through recent and focused development on these pivotal frameworks, several challenges remain for the forest tree community. As datasets become larger and more integrative, it is increasingly difficult for small database teams to keep up with the data capture and curation. With increasing access to reference genomes, large-scale population studies, and high throughput environmental data, biological databases must develop more efficient metadata capture, storage, and query capacity. These repositories will be tasked with implementing advanced natural language processing, automated metadata capture, and ontological term assignment to span not only genetic data, but associated phenotypic and environmental data. These latter categories encompass an expansive range, from traditional growth traits to canopy

---

metrics, soil profiles, microbiomes, and metatranscriptomics. The biomedical community has paved the way for some of this technology but forest tree data, and the associated genetic resources, remain more heterogenous (Koleck et al., 2019). This heterogeneity is combined with high throughput technologies, such as remote sensing, that challenge existing cyberinfrastructure in terms of efficient transfer, storage, and query (Côté et al., 2018). Capturing data for large forest tree populations may involve storing millions of genotypes across thousands of individuals or hundreds of pangenomes. It will also rely on a combination of sequencing and phenotyping technologies that continue to evolve (Bolger et al., 2019). After the storage and minimal reporting requirements are established, the frameworks the databases are built upon will need to assist users in determining the most appropriate analytics and provide the required formatting for the queried data. While progress has been made in connecting data to workflows on high performance computing, such as Galaxy; systems that can recommend appropriate workflows are still in progress. The future of biological databases for all plants is reproducible workflows that represent the metadata associated with the original studies. Concerted efforts in this area and integration of new data types evolving from high throughput technologies will be key to advancing discovery for the forest tree community.

AUTHOR CONTRIBUTIONS

JW, MS, NS, DM, and SF designed the databases and software described. EG, NH, SB, TF, SZ, RR, PR, and LS developed the core TreeGenes and TreeGenes Tripal modules. MS, BC, AA, and MC developed the core HWG and HWG Tripal modules. NS and CM developed the core PlantGenIE. SJ developed the core GDR and GDR Tripal modules. SF developed the core Tripal. JW, MS, NS, and DM wrote the manuscript. All authors read and approved the manuscript.

FUNDING

REFERENCES

Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Čech, M., et al. (2018). The galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* 46, W537–W544. doi: 10.1093/nar/gky379

Ashburner, M., Ball, C., Blake, J., Botstein, D., Butler, H., Cherry, J., et al. (2000). Gene ontology: tool for the unification of biology. the gene ontology consortium. *Nat. Genet.* 25, 25–29.

Barone, L., Williams, J., and Micklos, D. (2017). Unmet needs for analyzing biological big data: a survey of 704 NSF principal investigators. *PLoS Comp. Biol.* 13:e1005755. doi: 10.1371/journal.pcbi.1005755

Benson, D. A., Boguski, M. S., Lipman, D. J., and Ostell, J. (1997). GenBank. *Nucleic Acids Res.* 25, 1–6. doi: 10.1093/nar/25.1.1

Birol, I., Raymond, A., Jackman, S. D., Pleasance, S., Coope, R., Taylor, G. A., et al. (2013). Assembling the 20 Gb white spruce (Picea glauca) genome from whole-genome shotgun sequencing data. *Bioinformatics* 29, 1492–1497. doi: 10.1093/bioinformatics/btt178

Bolger, A. M., Poorter, H., Dumschott, K., Bolger, M. E., Arend, D., Osorio, S., et al. (2019). Computational aspects underlying genome to phenome analysis in plants. *Plant J.* 97, 182–198. doi: 10.1111/tpj.14179

Boratyn, G. M., Camacho, C., Cooper, P. S., Coulouris, G., Fong, A., Ma, N., et al. (2013). BLAST: a more efficient report with usability improvements. *Nucleic Acids Res.* 41, W29–W33. doi: 10.1093/nar/gkt282

Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using diamond. *Nat. Methods* 12:59. doi: 10.1038/nmeth.3176

Buels, R., Yao, E., Diesh, C. M., Hayes, R. D., Munoz-Torres, M., Helt, G., et al. (2016). JBrowse: a dynamic web platform for genome visualization and analysis. *Genome Biol.* 17:66. doi: 10.1186/s13059-016-0924-1

Canas, R. A., Canales, J., Gomez-Maldonado, J., Avila, C., and Canovas, F. M. (2014). Transcriptome analysis in maritime pine using laser capture microdissection and 454 pyrosequencing. *Tree Physiol.* 34, 1278–1288. doi: 10.1093/treephys/tpt113

Carollo, V., Matthews, D. E., Lazo, G. R., Blake, T. K., Hummel, D. D., Lui, N., et al. (2005). GrainGenes 2.0. an improved resource for the small-grains community. *Plant Physiol.* 139, 643–651. doi: 10.1104/pp.105.064485

Cartinhour, S. W. (1997). Public informatics resources for rice and other grasses. *Plant Mol. Biol.* 35, 241–251. doi: 10.1007/978-94-011-5794-0_23

Chen, F., Dong, W., Zhang, J., Guo, X., Chen, J., Wang, Z., et al. (2018). The sequenced angiosperm genomes and genome databases. *Front. Plant Sci.* 9:418. doi: 10.3389/fpls.2018.00418

Chen, M., Henry, N., Almsaeed, A., Zhou, X., Wegrzyn, J., Ficklin, S., et al. (2017). New extension software modules to enhance searching and display of transcriptome data in tripal databases. *Database* 2017:bax052. doi: 10.1093/database/bax052

Cooper, L., Meier, A., Laporte, M.-A., Elser, J. L., Mungall, C., Sinn, B. T., et al. (2018). The planteome database: an integrated resource for reference ontologies, plant genomics and phenomics. *Nucleic Acids Res.* 46, D1168–D1180. doi: 10.1093/nar/gkx1152

Côté, J. F., Fournier, R. A., Luther, J. E., and van Lier, O. R. (2018). Fine-scale three-dimensional modeling of boreal forest plots to improve forest characterization with remote sensing. *Remote Sens. Environ.* 219, 99–114. doi: 10.1016/j.rse.2018.09.026

De La Torre, A., Birol, I., Bousquet, J., Ingvarsson, P., Jansson, S., Jones, S. J. M., et al. (2014). Insights into conifer giga-genomes. *Plant Physiol.* 166, 1724–1732. doi: 10.1104/pp.114.248708

Emms, D. M., and Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* 16:157. doi: 10.1186/s13059-015-0721-2

Falk, T., Herndon, N., Grau, E., Buehler, S., Richter, P., Zaman, S., et al. (2018). Growing and cultivating the forest genomics database, TreeGenes. *Database* 2018:bay084. doi: 10.1093/database/bay084

FAO, IIASA and ISSCAS, ISRIC (2009). *JRC Harmonized World Soil Database (Version 1.1)*. Rome: FAO.

Fick, S. E., and Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* 37, 4302–4315. doi: 10.1038/sdata.2018.254

Ficklin, S. P., Sanderson, L.-A., Cheng, C.-H., Staton, M. E., Lee, T., Cho, I.-H., et al. (2011). Tripal: a construction toolkit for online genome databases. *Database* 2011:bar044. doi: 10.1093/database/bar044

Finn, R., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J., et al. (2010). The Pfam protein families database. *Nucleic Acids Res.* 38, D211–D222. doi: 10.1093/nar/gkp985

Flanders, D. J., Weng, S., Petel, F. X., and Cherry, J. M. (1998). AtDB, the *Arabidopsis thaliana* database, and graphical-web-display of progress by the

Arabidopsis genome initiative. *Nucleic Acids Res.* 26, 80–84. doi: 10.1093/nar/26.1.80

Frishman, D., Heumann, K., Lesk, A., and Mewes, H. W. (1998). Comprehensive, comprehensible, distributed and intelligent databases: current status. *Bioinformatics* 14, 551–561. doi: 10.1093/bioinformatics/14.7.551

Giacomello, S., Salmén, F., Terebieniec, B. K., Vickovic, S., Navarro, J. F., Alexeyenko, A., et al. (2017). Spatially resolved transcriptome profiling in model plant species. *Nat. Plants* 3:17061. doi: 10.1038/nplants.2017.61

Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., et al. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40, D1178–D1186. doi: 10.1093/nar/gkr944

Grant, D., Nelson, R. T., Cannon, S. B., and Shoemaker, R. C. (2009). SoyBase, the USDA-ARS soybean genetics and genomics database. *Nucleic Acids Res.* 38, D843–D846. doi: 10.1093/nar/gkp798

Herndon, N., Falk, T., Grau, E. S., Jung, S., Ficklin, S., Main, D., et al. (2018). "Association mapping for forest trees with CartograTree," in *Application of Semantic Technologies in Biodiversity Science. Studies on the Semantic Web*, ed. A. E. Thessen (Amsterdam: IOS Press).

Herrero, J., Muffato, M., Beal, K., Fitzgerald, S., Gordon, L., Pignatelli, M., et al. (2016). Ensembl comparative genomics resources. *Database* 2016:bav096. doi: 10.1093/database/bav096

Holliday, J. A., Aitken, S. N., Cooke, J. E. K., Fady, B., Gonzalez-Martinez, S. C., Heuertz, M., et al. (2017). Advances in ecological genomics in forest trees and applications to genetic resources conservation and breeding. *Mol. Ecol.* 26, 706–717. doi: 10.1111/mec.13963

Houghton, R. A. (2005). Aboveground forest biomass and the global carbon balance. *Glob. Change Biol.* 11, 945–958. doi:10.1111/j.1365-2486.2005.00955.x

Huala, E., Dickerman, A. W., Garcia-Hernandez, M., Weems, D., Reiser, L., LaFond, F., et al. (2001). The *Arabidopsis* information resource (TAIR): a comprehensive database and web-based information retrieval, analysis, and visualization system for a model plant. *Nucleic Acids Res.* 29, 102–105. doi: 10.1093/nar/29.1.102

Jaiswal, P., Avraham, S., Ilic, K., Kellogg, E. A., McCouch, S., Pujar, A., et al. (2005). Plant ontology (PO): a controlled vocabulary of plant structures and growth stages. *Comp. Funct. Genomics* 6, 388–397. doi: 10.1002/cfg.496

Jokipii-Lukkari, S., Sundell, D., Nilsson, O., Hvidsten, T. R., Street, N. R., and Tuominen, H. (2017). NorWood: a gene expression resource for evo-devo studies of conifer wood development. *New Phytol.* 216, 482–494. doi: 10.1111/nph.14458

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., et al. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30, 1236–1240. doi: 10.1093/bioinformatics/btu031

Jung, S., Ficklin, S. P., Lee, T., Cheng, C.-H., Blenda, A., Zheng, P., et al. (2013). The Genome Database for Rosaceae (GDR): year 10 update. *Nucleic Acids Res.* 42, D1237–D1244. doi: 10.1093/nar/gkt1012

Jung, S., Lee, T., Cheng, C. H., Buble, K., Zheng, P., Yu, J., et al. (2018). 15 years of GDR: new data and functionality in the genome database for rosaceae. *Nucleic Acids Res.* 47, D1137–D1145. doi: 10.1093/nar/gky1000

Jung, S., Lee, T., Cheng, C.-H., Ficklin, S., Yu, J., Humann, J., et al. (2017). Extension modules for storage, visualization and querying of genomic, genetic and breeding data in tripal databases. *Database* 2017:bax092. doi: 10.1093/database/bax092

Jung, S., Lee, T., Ficklin, S., Yu, J., Cheng, C.-H., and Main, D. (2016). Chado use case: storing genomic, genetic and breeding data of *Rosaceae* and *Gossypium* crops in Chado. *Database* 2016:baw010. doi: 10.1093/database/baw010

Kattge, J., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Bönisch, G., et al. (2011). TRY – a global database of plant traits. *Glob. Change Biol.* 17, 2905–2935. doi: 10.1111/j.1365-2486.2011.02451.x

Koleck, T. A., Dreisbach, C., Bourne, P. E., and Bakken, S. (2019). Natural language processing of symptoms documented in free-text narratives of electronic health records: a systematic review. *J. Am. Med. Inform. Assoc.* 26, 364–379. doi: 10.1093/jamia/ocy173

Krajewski, P., Chen, D., Ćwiek, H., van Dijk, A. D. J., Fiorani, F., Kersey, P., et al. (2015). Towards recommendations for metadata and data handling in plant phenotyping. *J. Exp. Bot.* 66, 5417–5427. doi: 10.1093/jxb/erv271

Lai, K., Lorenc, M. T., and Edwards, D. (2012). Genomic databases for crop improvement. *Agronomy* 2, 62–73. doi: 10.3390/agronomy2010062

Lawrence, C. J., Dong, Q., Polacco, M. L., Seigfried, T. E., and Brendel, V. (2004). MaizeGDB, the community database for maize genetics and genomics. *Nucleic Acids Res.* 32, D393–D397. doi: 10.1093/nar/gkh011

Lens, F., Cooper, L., Gandolfo, M. A., Groover, A., Jaiswal, P., Lachenbruch, B., et al. (2012). An extension of the plant ontology project supporting wood anatomy and development research. *IAWA J.* 33, 113–117. doi: 10.1163/22941932-90000083

Matasci, N., Hung, L.-H., Yan, Z., Carpenter, E. J., Wickett, N. J., Mirarab, S., et al. (2014). Data access for the 1,000 Plants (1KP) project. *GigaScience* 3:17. doi: 10.1186/2047-217X-3-17

Neale, D. B., Langley, C. H., Salzberg, S. L., and Wegrzyn, J. L. (2013). Open access to tree genomes: the path to a better forest. *Genome Biol.* 14:120. doi: 10.1186/gb-2013-14-6-120

Neale, D. B., Wegrzyn, J. L., Stevens, K. A., Zimin, A. V., Puiu, D., Crepeau, M. W., et al. (2014). Decoding the massive genome of loblolly pine using haploid DNA and novel assembly strategies. *Genome Biol.* 15:R59. doi: 10.1186/gb-2014-15-3-r59

Netotea, S., Sundell, D., Street, N. R., and Hvidsten, T. R. (2014). ComPlEx: conservation and divergence of co-expression networks in *A. thaliana*, *Populus* and *O. sativa*. *BMC Genomics* 15:106. doi: 10.1186/1471-2164-15-106

Nystedt, B., Street, N. R., Wetterbom, A., Zuccolo, A., Lin, Y.-C., Scofield, D. G., et al. (2013). The Norway spruce genome sequence and conifer genome evolution. *Nature* 497, 579–584.

Plomion, C., Bastien, C., Bogeat-Triboulot, M.-B., Bouffier, L., Déjardin, A., Duplessis, S., et al. (2016). Forest tree genomics: 10 achievements from the past 10 years and future prospects. *Ann. For. Sci.* 73, 77–103. doi: 10.1007/s13595-015-0488-3

Proost, S., Van Bel, M., Vaneechoutte, D., Van de Peer, Y., Inzé, D., Mueller-Roeber, B., et al. (2015). PLAZA 3.0: an access point for plant comparative genomics. *Nucleic Acids Res.* 43, D974–D981. doi: 10.1093/nar/gku986

Sanderson, L.-A., Ficklin, S. P., Cheng, C.-H., Jung, S., Feltus, F. A., Bett, K. E., et al. (2013). Tripal v1.1: a standards-based toolkit for construction of online genetic and genomic databases. *Database* 2013:bat075. doi: 10.1093/database/bat075

Sayers, E. W., Cavanaugh, M., Clark, K., Ostell, J., Pruitt, K. D., and Karsch-Mizrachi, I. (2019). GenBank. *Nucleic Acids Res.* 47, D94–D99. doi: 10.1093/nar/gky989

Schiffthaler, B., Serrano, A., Delhomme, N., and Street, N. R. (2018). Seidr: a toolkit for calculation of crowd networks. *bioRxiv*

Sjödin, A., Street, N. R., Sandberg, G., Gustafsson, P., and Jansson, S. (2009). The populus genome integrative explorer (PopGenIE): a new resource for exploring the populus genome. *New Phytol.* 182, 1013–1025. doi: 10.1111/j.1469-8137.2009.02807.x

Stein, L. D., Mungall, C., Shu, S., Caudy, M., Mangone, M., Day, A., et al. (2002). The generic genome browser: a building block for a model organism system database. *Genome Res.* 12, 1599–1610. doi: 10.1101/gr.403602

Stevens, K. A., Wegrzyn, J. L., Zimin, A., Puiu, D., Crepeau, M., Cardeno, C., et al. (2016). Sequence of the sugar pine megagenome. *Genetics* 204, 1613–1626. doi: 10.1534/genetics.116.193227

Sundell, D., Mannapperuma, C., Netotea, S., Delhomme, N., Lin, Y.-C., Sjödin, A., et al. (2015). The plant genome integrative explorer resource: plantgenIE.org. *New Phytol.* 208, 1149–1156. doi: 10.1111/nph.13557

Sundell, D., Street, N. R., Kumar, M., Mellerowicz, E. J., Kucukoglu, M., Johnsson, C., et al. (2017). AspWood: high-spatial-resolution transcriptome profiles reveal uncharacterized modularity of wood

formation in *Populus tremula*. *Plant Cell* 29, 1585–1604. doi: 10.1105/tpc. 17.00153

Tateno, Y., and Gojobori, T. (1997). DNA data bank of Japan in the age of information biology. *Nucleic Acids Res.* 25, 14–17. doi: 10.1093/nar/25.1.14

Watts, N. A., and Feltus, F. A. (2017). Big Data Smart Socket (BDSS): a system that abstracts data transfer habits from end users. *Bioinformatics* 33, 627–628. doi: 10.1093/bioinformatics/btw679

Wegrzyn, J. L., Lee, J. M., Tearse, B. R., and Neale, D. B. (2008). TreeGenes: a forest tree genome database. *Int. J. Plant Genomics* 2008:412875. doi: 10.1155/2008/412875

Wegrzyn, J. L., Liechty, J. D., Stevens, K. A., Wu, L.-S., Loopstra, C. A., Vasquez-Gross, H. A., et al. (2014). Unique features of the loblolly pine (*Pinus taeda L.*) megagenome revealed through sequence annotation. *Genetics* 196, 891–909. doi: 10.1534/genetics.113.159996

Wegrzyn, J. L., Main, D., Figueroa, B., Choi, M., Yu, J., Neale, D. B., et al. (2012). Uniform standards for genome databases in forest and fruit trees. *Tree Genet. Genomes* 8:549. doi: 10.1007/s11295-012-0494-7

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V., and Provart, N. J. (2007). An 'electronic fluorescent pictograph' browser for exploring and analyzing large-scale biological data sets. *PLoS One* 2:e718. doi: 10.1371/journal.pone.0000718

Wytko, C., Soto, B., and Ficklin, S. P. (2017). blend4php: a PHP API for galaxy. *Database* 2017:baw154. doi: 10.1093/database/baw154

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read
for greatest visibility
and readership

90
FAST PUBLICATION
Around 90 days
from submission
to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative,
and constructive
peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers
acknowledged by name
on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: info@frontiersin.org | +41 21 510 17 00

REPRODUCIBILITY OF RESEARCH
Support open data
and methods to enhance
research reproducibility

DIGITAL PUBLISHING
Articles designed
for optimal readership
across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics
track visibility across
digital media

EXTENSIVE PROMOTION
Marketing
and promotion
of impactful research

LOOP RESEARCH NETWORK
Our network
increases your
article's readership