

A phenomic and GWAS adventure studying the causes of genetic variation in amenability to regeneration and transformation in black cottonwood (*Populus trichocarpa*)

Steve Strauss
Oregon State University



Michael Nagle,
PhD student and
postdoc

Summarizing main results from a major NSF grant

NSF Plant Genome Research Program - “Analysis of Genes Affecting Plant Regeneration and Transformation (RT) in Poplar”

Steve Strauss, Li Fuxin, Yuan Jiang, Troy Hall, Jay Well, Wellington Muchero



- Explore a variety of RT methods to maximize variation in RT responses
- Develop new phenomic tools, including an image capture and generalizable machine-vision system, to determine *in vivo* and *vitro* phenotypes
- Using GWAS, map sets of alleles that are associated with variation in RT frequency
- Study cognitive processes with respect to GE crops, develop case studies and new teaching materials, deliver them to rural and underserved communities including via social media

[Home](#) > [In Vitro Cellular & Developmental Biology - Plant](#) > [Article](#)

Factors affecting *in vitro* regeneration in the model tree *Populus trichocarpa* I. Medium, environment, and hormone controls on organogenesis

[Home](#) > [In Vitro Cellular & Developmental Biology - Plant](#) > [Article](#)

Factors affecting *in vitro* regeneration in the model tree *Populus trichocarpa*: II. Heritability estimates, correlations among explant types, and genetic interactions with treatments among wild genotypes

Scientific results mostly published, in four papers

Horticulture Research

Volumes Advance Access Collections Publish Alerts About Horticulture Research



JOURNAL ARTICLE

GWAS identifies candidate genes controlling adventitious rooting in *Populus trichocarpa*

Michael F Nagle ✉, Jialin Yuan, Damanpreet Kaur, Cathleen Ma, Ekaterina Peremyslova, Yuan Jiang, Bahiya Zahl, Alexa Niño de Rivera, Wellington Muchero, Li Fuxin ... Show more

Horticulture Research, Volume 10, Issue 8, August 2023, uhad125, <https://doi.org/10.1093/hr/uhad125>

Volume 10, Issue 8
August 2023

G3 Genes | Genomes | Genetics

Issues More Content Series Submit Alerts About G3 Genes|Genomes|Ge



JOURNAL ARTICLE FEATURED

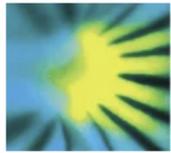
GWAS supported by computer vision identifies large numbers of candidate regulators of *in planta* regeneration in *Populus trichocarpa*

Michael F Nagle, Jialin Yuan, Damanpreet Kaur, Cathleen Ma, Ekaterina Peremyslova, Yuan Jiang, Alexa Niño de Rivera, Sara Jawdy, Jin-Gui Chen, Kai Feng ... Show more
Author Notes

G3 Genes|Genomes|Genetics, Volume 14, Issue 4, April 2024, jkae026, <https://doi.org/10.1093/g3journal/jkae026>

Volume 14, Issue 4
April 2024

Article Contents



New Phytologist

Full paper |  Open Access |  

Genome-wide association study and network analysis of *in vitro* transformation in *Populus trichocarpa* support key roles of diverse phytohormone pathways and cross talk

Michael F. Nagle, Jialin Yuan, Damanpreet Kaur, Cathleen Ma, Ekaterina Peremyslova, Yuan Jiang, Greg S. Goralogia, Anna Magnuson, Jia Yi Li, Wellington Muchero, Li Fuxin, Steven H. Strauss ✉

First published: 22 April 2024 | <https://doi.org/10.1111/nph.19737> | Citations: 2

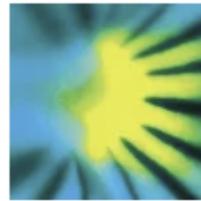
Agenda

- I. Background
Motivation, genetic resources
- II. High-throughput phenomics
RGB/hyperspectral methods developed
- III. Association mapping methods & results
Power, precision, and selected results
- IV. Gene candidate roles/interpretation
From SNPs to genes, genes to pathways

Agricultural and forest genetic engineering, including gene editing, severely limited by social and technical challenges

- Social: Key are regulatory barriers and market bans – such as by “green” certification systems
- Technical: Key is recalcitrance, and variation, in amenability to transformation and editing within and among species

The basis of this variation is essentially unknown



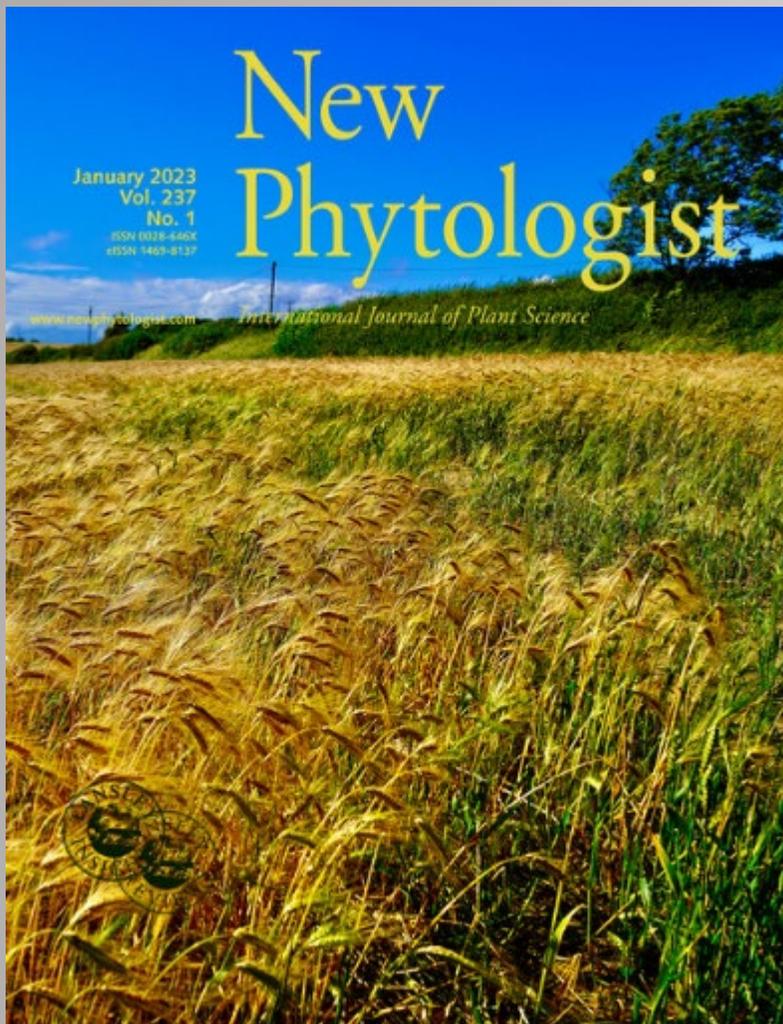
New Phytologist

Viewpoints |  Free Access

Social and biological innovations are essential to deliver transformative forest biotechnologies

Wout Boerjan , Steven H. Strauss 

First published: 27 May 2024 | <https://doi.org/10.1111/nph.19855>



New section of journal on “transformative plant biotechnology”

- Joined as an editor (again)
- All kinds of plant biotechnologies welcomed, emphasizing....
 - Scientific novelty
 - Social impact/value

Poplar (genus *Populus*) is the model tree for genomics and biotechnology

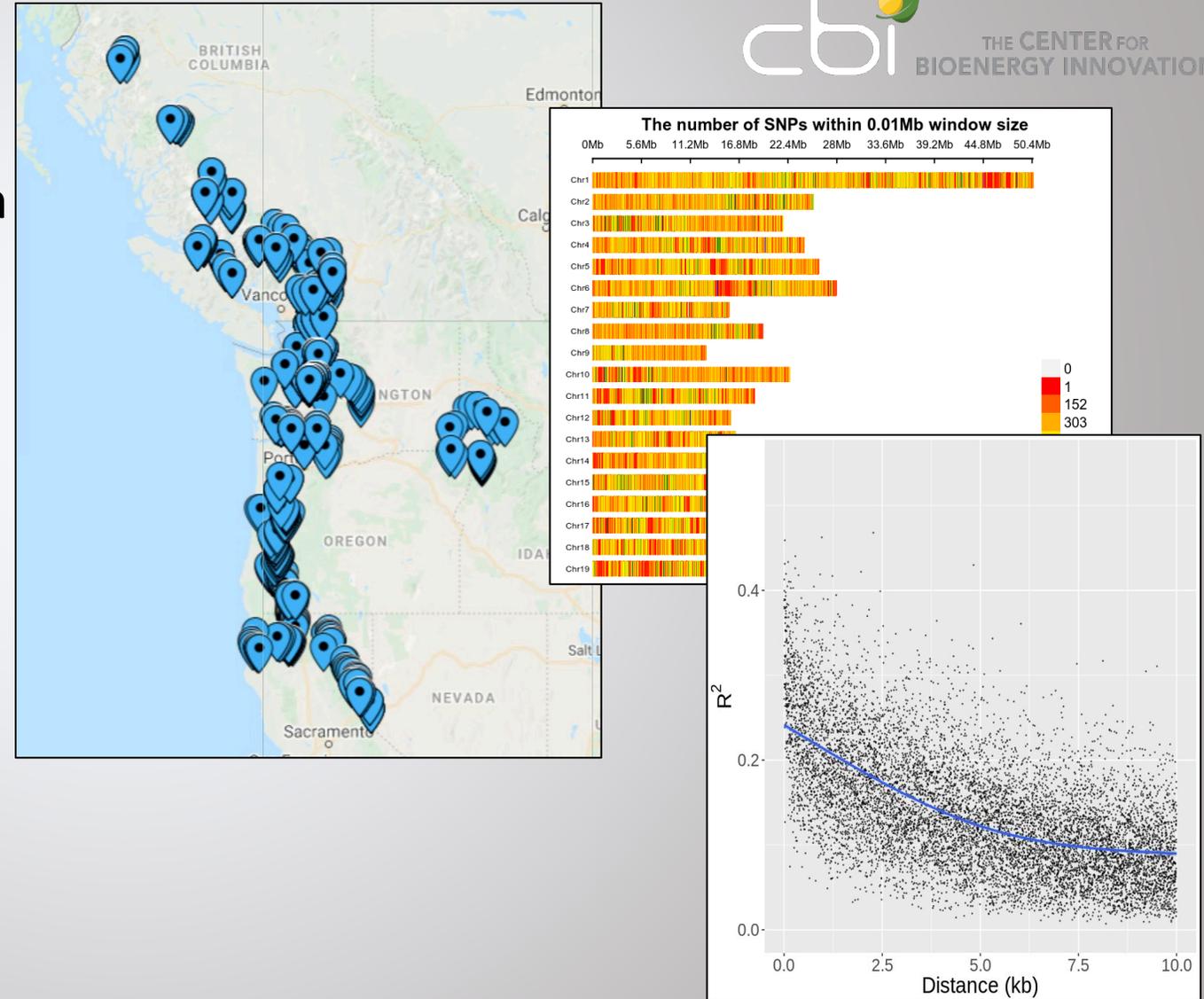
- Easy to clone
- Grows rapidly
- Readily hybridizes
- Several genotypes readily transformed
- Modest sized genome (esp. compared to conifers)
- High quality reference genome and many other omic resources
 - *US Department of Energy model bioenergy tree*



Genetic resources for study

Power and precision supported by DOE poplar resources

- Wild, highly diverse parents
 - 1,323 genotypes, among largest in plant GWAS
 - Cloned into plantations/clone banks (Corvallis)
- Very high SNP density
 - 10 million common SNPs/tree
- Very low linkage disequilibrium
- Continually improving reference genome/s



Agenda

I. Background

Motivation, genetic resources

II. High-throughput phenomics

RGB/hyperspectral methods developed

III. Association mapping methods & results

Power, precision, and selected results

IV. Gene candidate roles/interpretation

From SNPs to genes, genes to pathways

Transformation, regeneration (TR)

Multiple stages, each with complex and hard to quantify phenotypes over time



Co-cultivation with
Agrobacterium



Callus induction



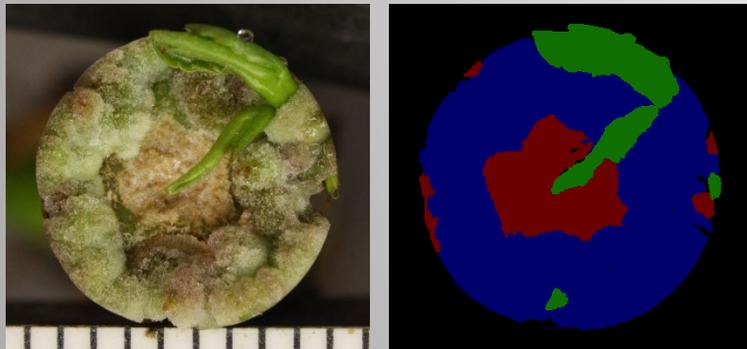
Shoot induction



Adventitious
rooting

Distinct machine-vision phenomics systems needed for four GWAS studies

In planta
shoot regeneration



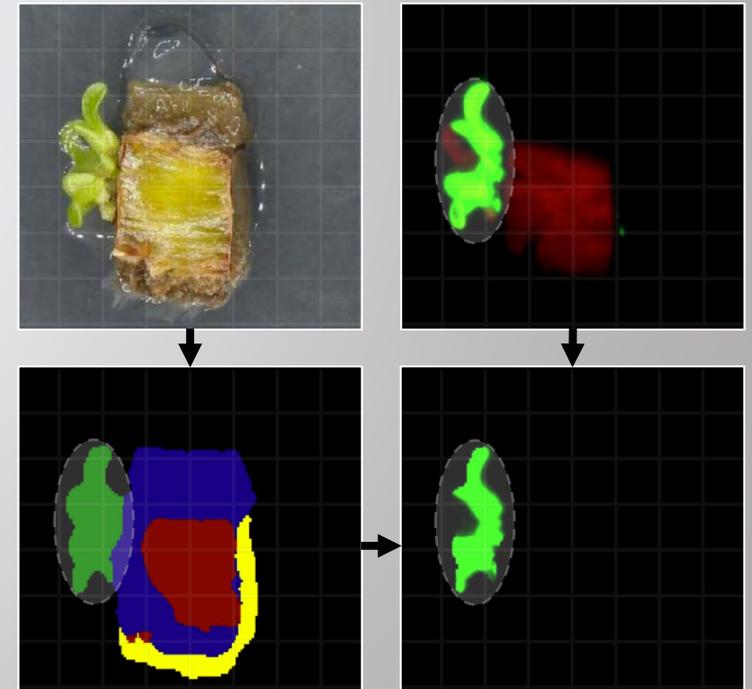
In planta
adventitious rooting



In vitro
regeneration



+ transformation

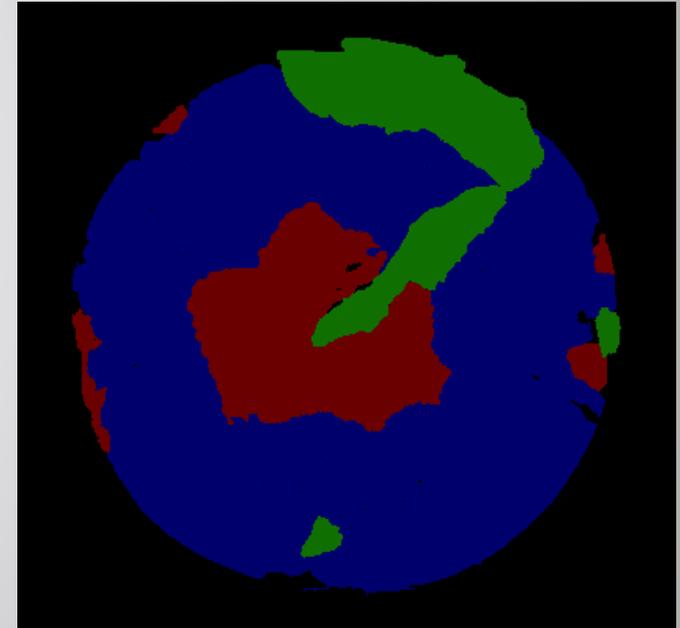
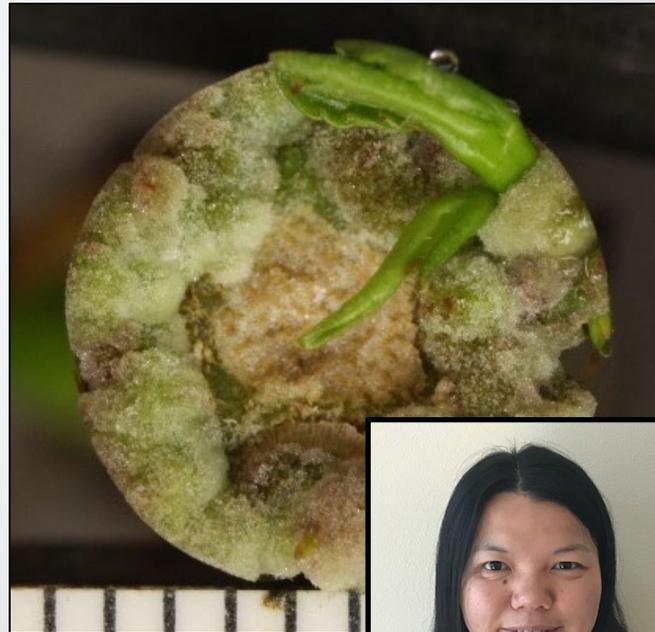


In planta callus/shoot segmentation

Stem cuttings decapitated, treated with cytokinin, and left in water



Deep convolutional neural network (DCNN) to segment tissues



Jialin Yuan
PhD Student,
Machine Vision

Graphical User Interface (GUI) for annotation of training set: "IDEAS"



Damanpreet Kaur
Master's Student,
Machine Vision



Jialin Yuan
PhD Student,
Machine Vision

The screenshot displays the IDEAS GUI with several highlighted areas:

- Toolkit:** A circled area containing 'posPen', 'negPen', and 'Rectangle' options.
- Class Panel:** A circled area containing a list of classes: 'Stem', 'Callus', and 'Shoot'.
- Canvas:** A large central area showing a plant image with segmentation masks and a red dashed line indicating a selection.
- History Panel:** A panel on the right showing a list of actions (all 'negPen') and their corresponding thumbnails.

Other visible elements include a 'Class' dropdown, 'Zoom In/Out' buttons, a 'Process' button, and a 'Class Panel' with an 'add' button and a 'clear' button.

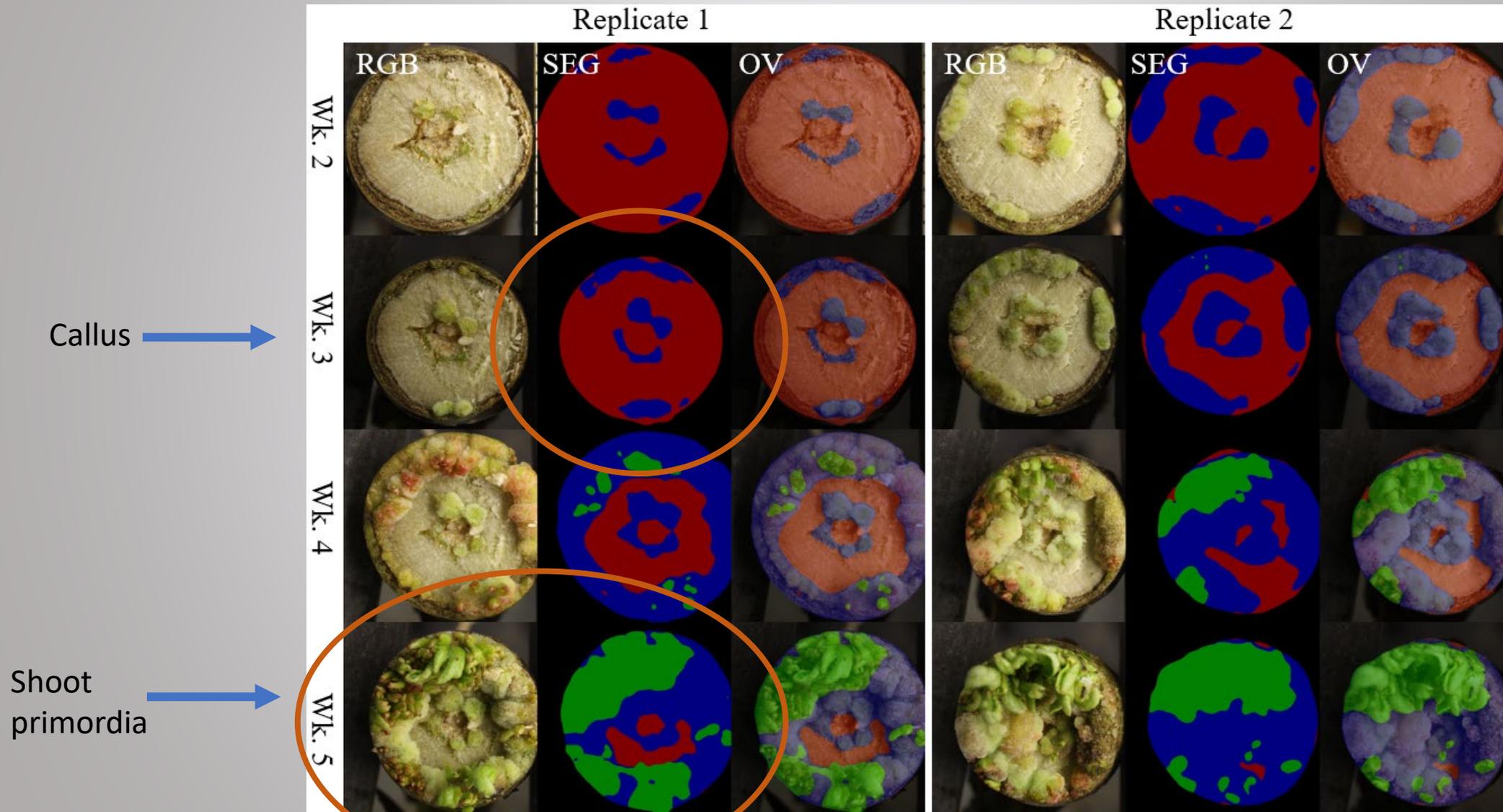
Plant Phenomics
Volume 2022, 2022, 9893639

Database/Software Article

Robust High-Throughput Phenotyping with Deep Segmentation Enabled by a Web-Based Annotator

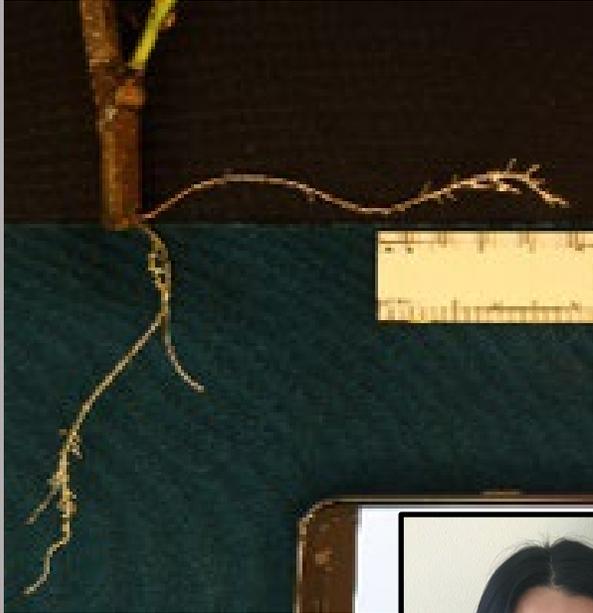
Jialin Yuan ¹, Damanpreet Kaur ¹, Zheng Zhou ¹, Michael Nagle ¹, Nicholas George Kiddle ¹, Nihar A. Doshi ², Ali Behnoudfar ¹, Ekaterina Peremyslova ¹, Cathleen Ma ¹, Steven H. Strauss ¹, Fuxin Li ¹

Examples of segmented stem sections

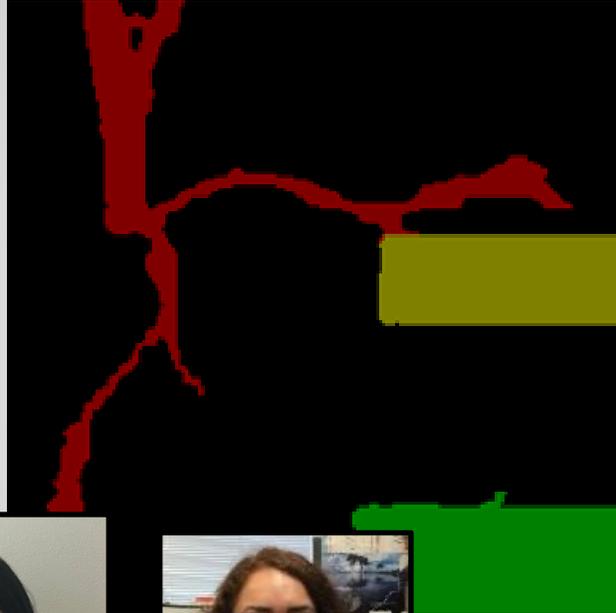


Stepwise segmentation of roots – a separate segmentation process developed

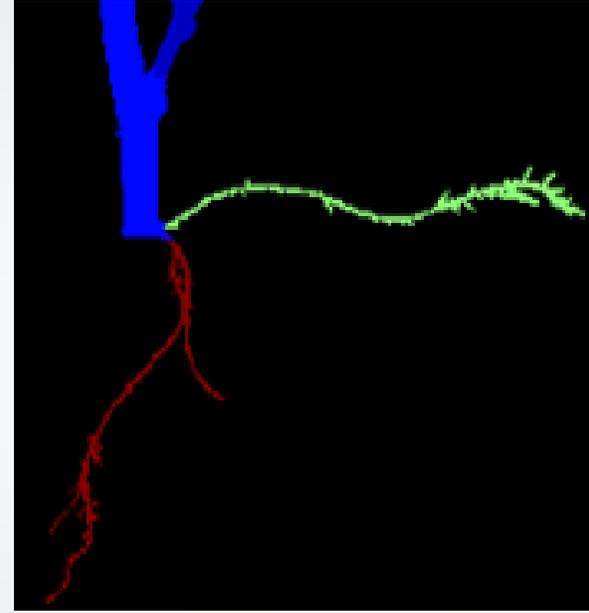
RGB image



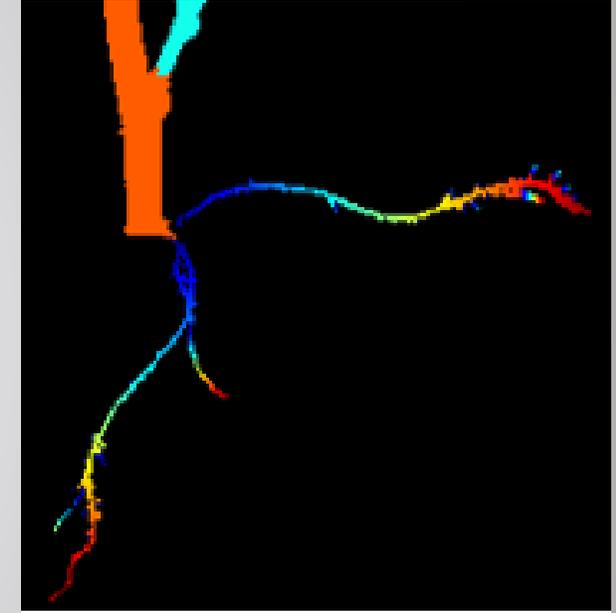
First deep model



Second deep model



Root metrics computed

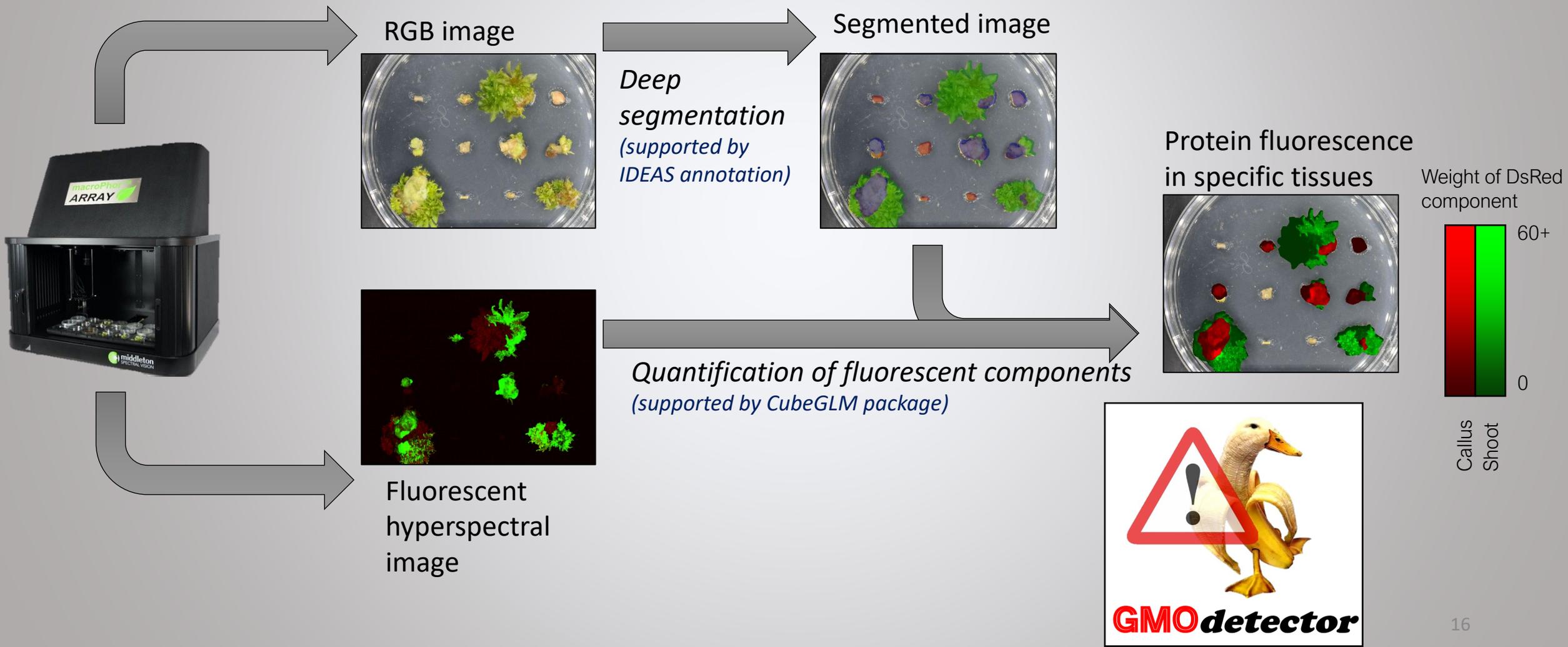


Jialin Yuan
PhD Student,
Machine Vision



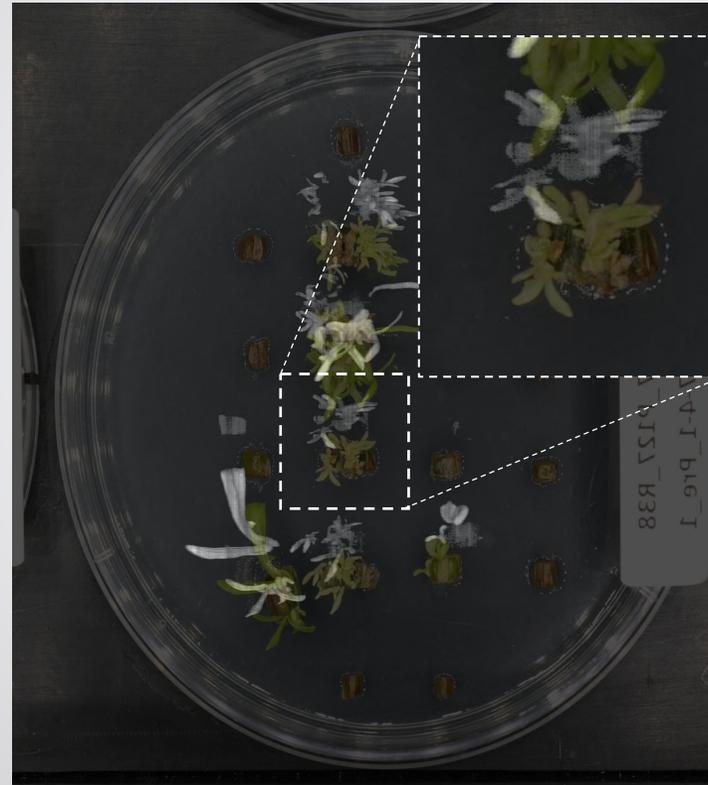
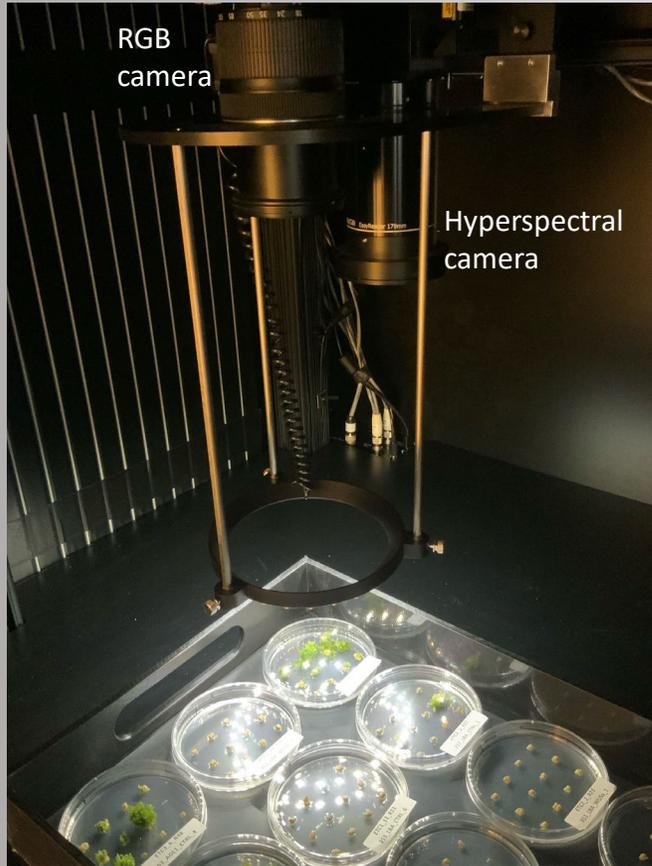
Bahiya Zahl
Undergraduate

Strategy to integrate RGB, hyperspectral data to measure tissue-specific transformation + regeneration

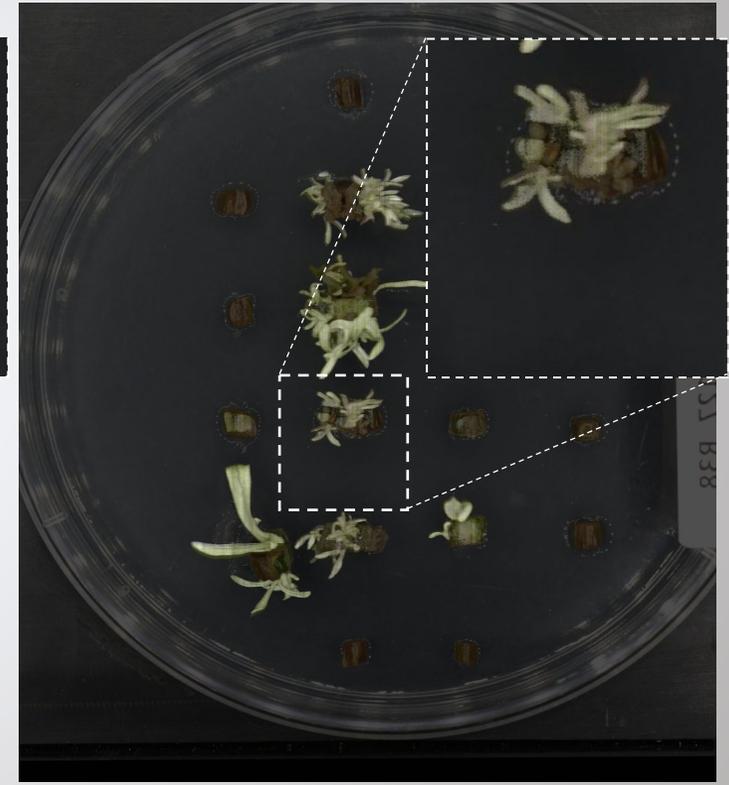


Other important “tweaks”

RGB and hyperspectral image layers differ in resolutions, proportions, frame, angle, explant location, and boundaries



Unaligned image channels



Aligned image channels

Dealing with contamination

Mostly a result of using greenhouse-source explants

Heavily contaminated samples detected by binary classification model, completely excluded from downstream analysis



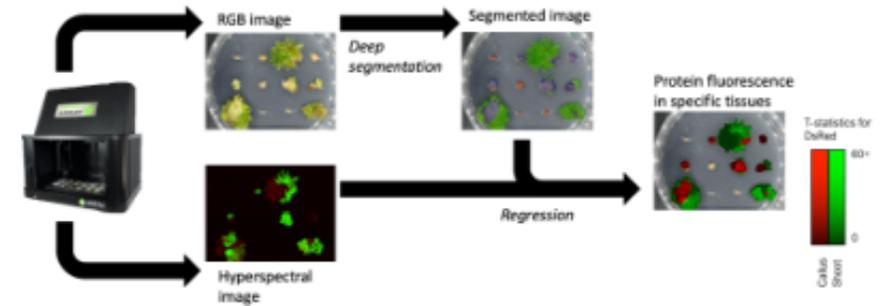
User-friendly and high-throughput deployment of entire workflow

- Many R, Python modules in workflow strung together
- Jupyter notebook templates for running workflow over a single dataset
- High-throughput mode: User provides spreadsheets of data paths, parameters



GMOdetector notebook: Template to analyze a new batch of images

Notebook template for applying routine hyperspectral/segmentation cross-analysis phenomics workflow over new datasets (v.0.6.2)



In this workflow, images taken with the macroPhor Array dual RGB/hyperspectral imaging platform are analyzed by a workflow in which regression quantifies fluorescent signals in hyperspectral images, deep learning segments RGB images into different tissues, and these datasets are cross-referenced to produce statistics on growth of transgenic callus and shoot.

Parameters for analysis:

The below variables must be modified appropriately every time this workflow is run over new images.

Data location

The `data` variable below provides the **complete** path to the folder containing data to be analyzed. This should include all folders and subfolders in which the data of interest is organized by. For the organizational system used for our lab's data, this should follow the format `"/Experiment/Subexperiment/Timepoint/"`

```
data="ENTER_DATA_PATH"
```

Phenomics methods available for download

- Highly general, widely applicable
 - **CubeGLM:** Hyperspectral analysis Python package
https://github.com/naglemi/gmodetector_py
 - **IDEAS:** Segmentation annotation GUI
<https://ideas.eecs.oregonstate.edu/>
- Specialized, tailored to study plant regeneration and transformation
 - **GMOLabeler:** RGB/hyperspectral channel cross-referencing
<https://github.com/naglemi/GMOLabeler>
 - **GMOdetector:** High-level workflow integrating all modules
<https://github.com/naglemi/GMOnotebook>

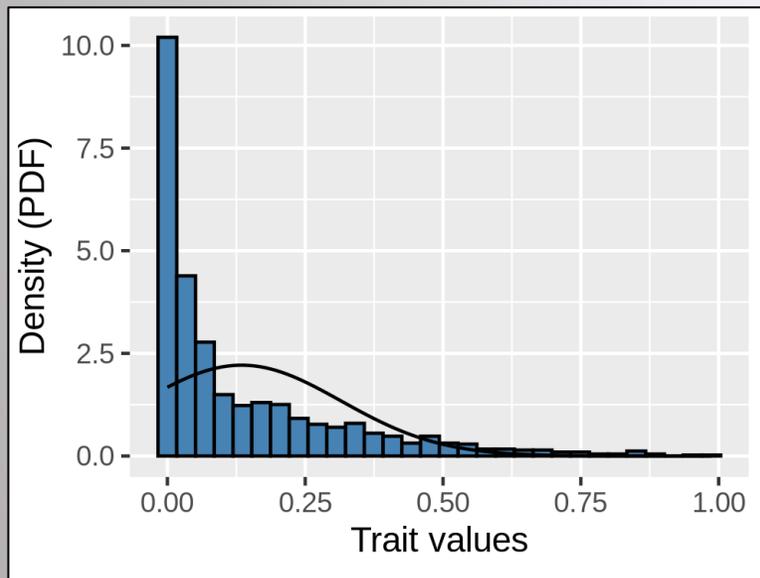
Agenda

- I. Background
Motivation, genetic resources
- II. High-throughput phenomics
RGB/hyperspectral methods developed
- III. Association mapping methods & results
Power, precision, and selected results
- IV. Gene candidate roles/interpretation
From SNPs to genes, genes to pathways

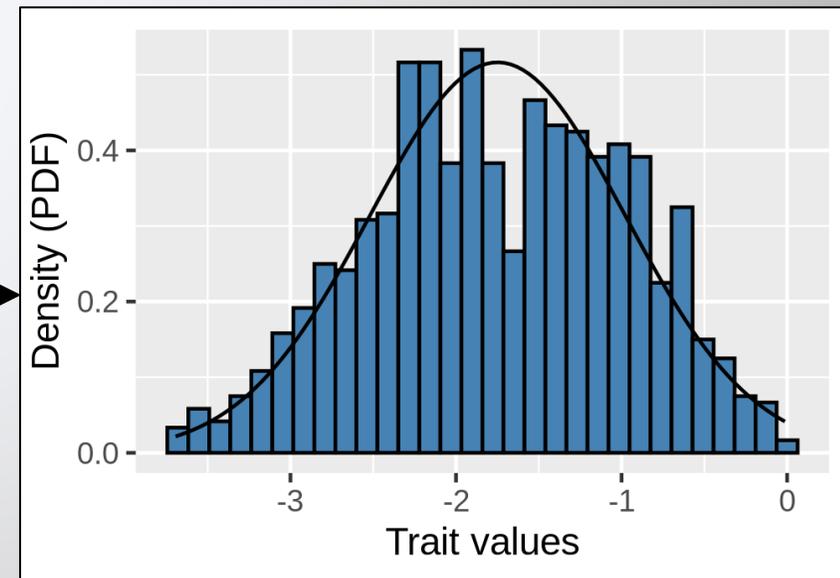
Non-normality of traits common

Transformations - A way to avoid violating assumption of residual normality, though can overcorrect, reduce power

Resampling is the other major option – we employed both



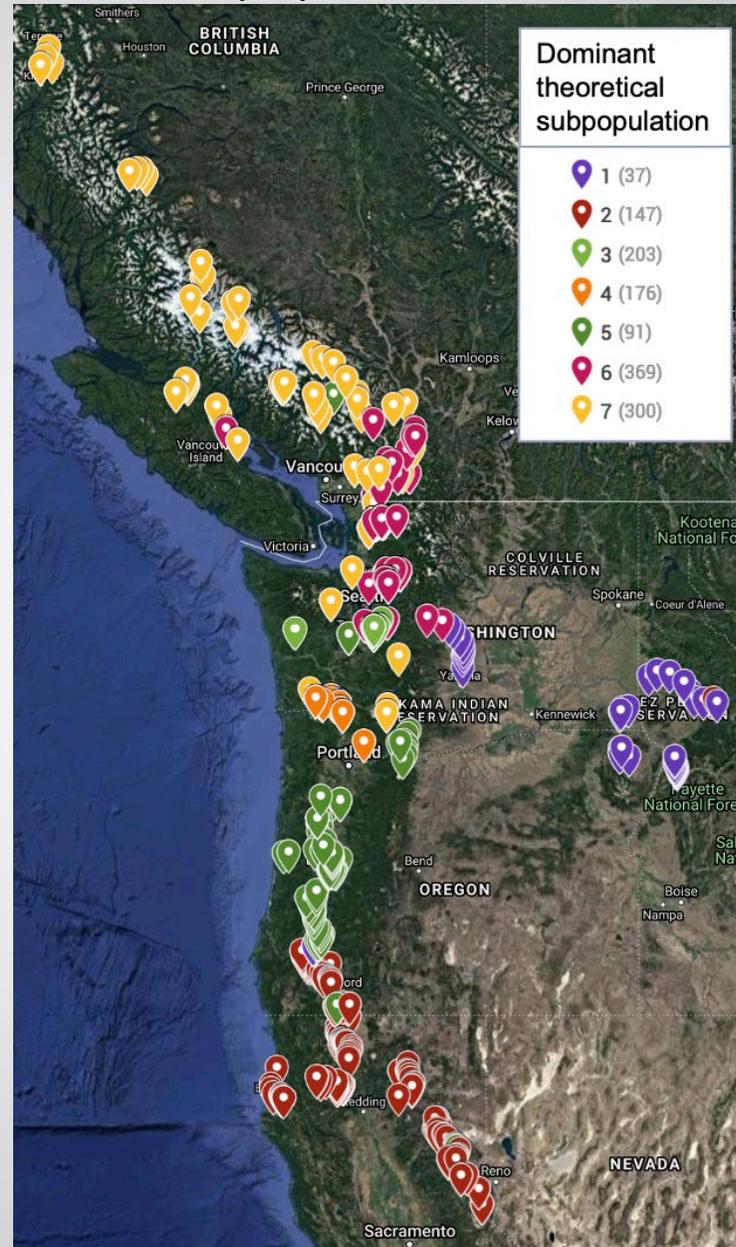
Box-Cox
transformation



fastSTRUCTURE results display stratification

Population stratification

Inflates false positives if not controlled



Low genotype number and low heritability required combined variant approaches

AJHG ARTICLE

Sequence Kernel Association Tests for the Combined Effect of Rare and Common Variants

Iuliana Ionita-Laza,^{1,6,*} Seunggeun Lee,^{2,6} Vlad Makarov,¹ Joseph D. Buxbaum,^{3,4,5} and Xihong Lin^{2,*}

Detecting Weak Signals by Combining Small P-Values in Genetic Association Studies

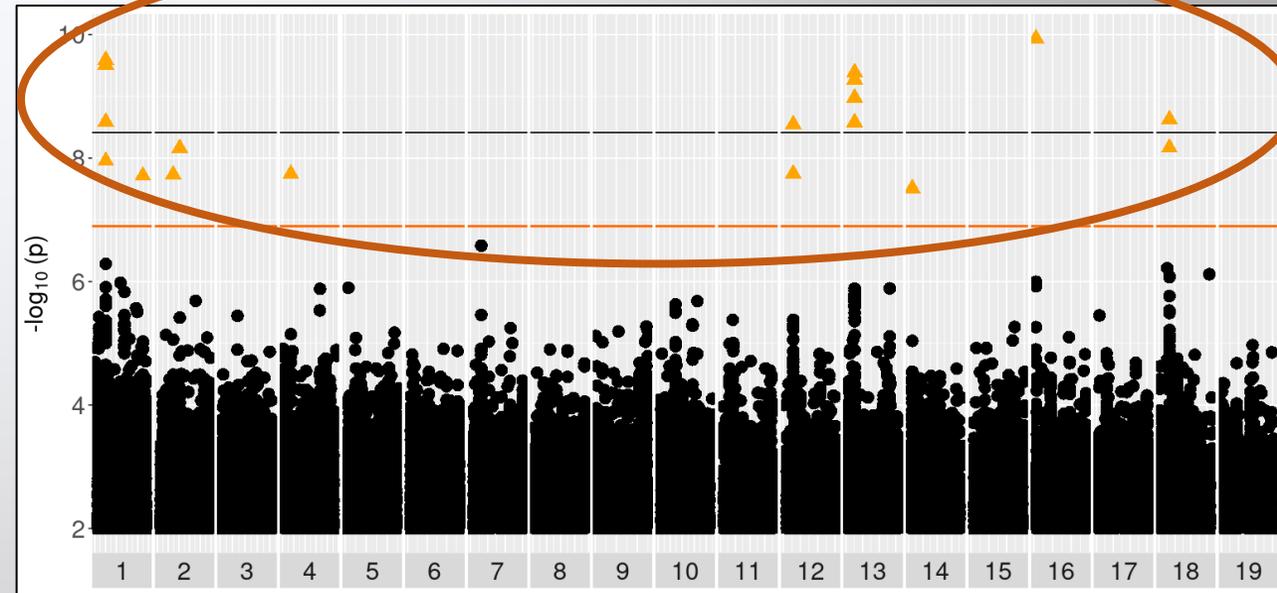
Olga A. Vsevolozhskaya¹, Fengjiao Hu² and Dmitri V. Zaykin^{1*}

frontiers in Genetics

SNP-set Kernel Association Test (SKAT) and Multi-Threaded Monte Carlo SKAT (MTMC-SKAT)



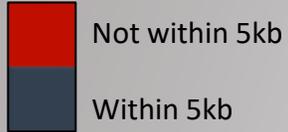
Augmented Rank Truncation (ART)



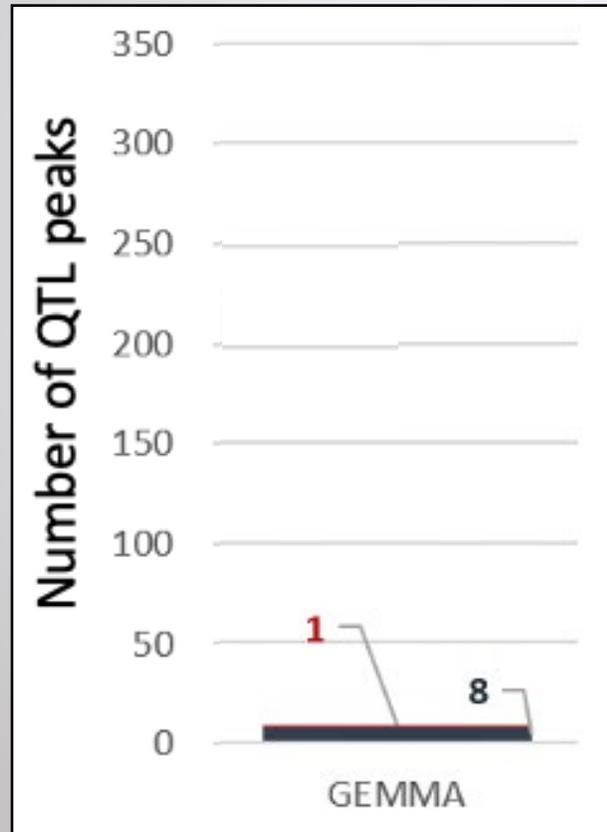
Most associations found via SKAT or ART

Example shown for *in vitro* transf./regen.

Distance from gene



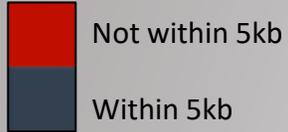
FDR ($\alpha = 0.10$)



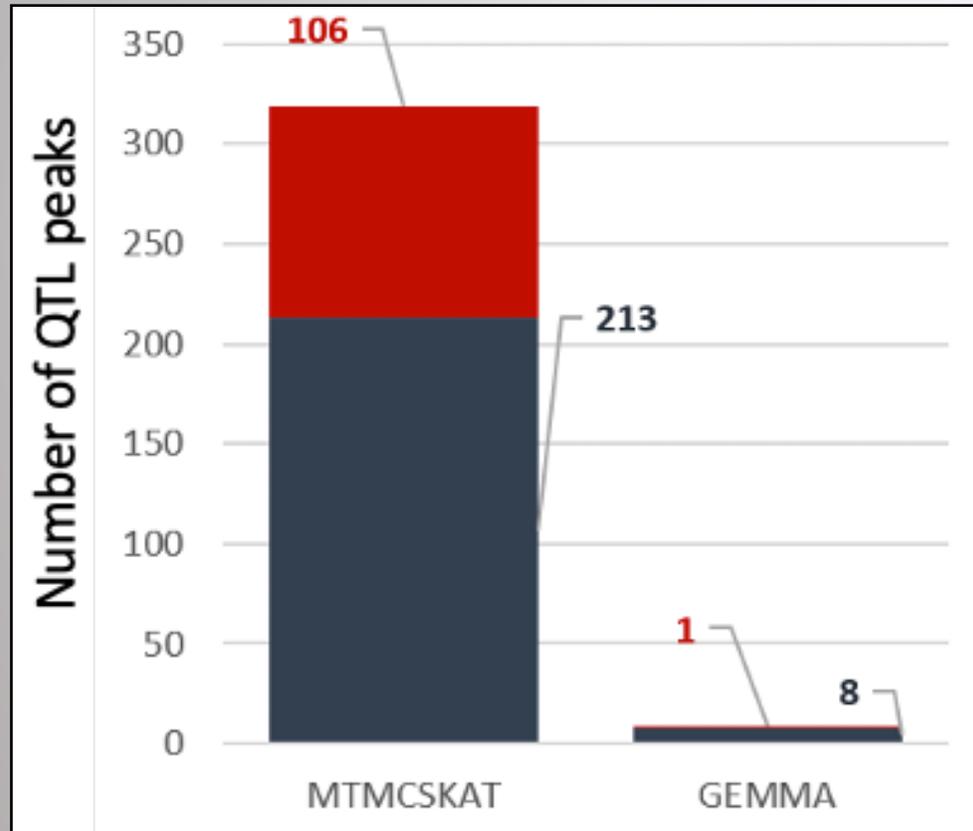
Most associations found via SKAT or ART

Example shown for *in vitro* transf./regen.

Distance from gene



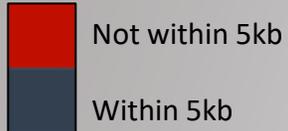
FDR ($\alpha = 0.10$)



Most associations found via SKAT or ART

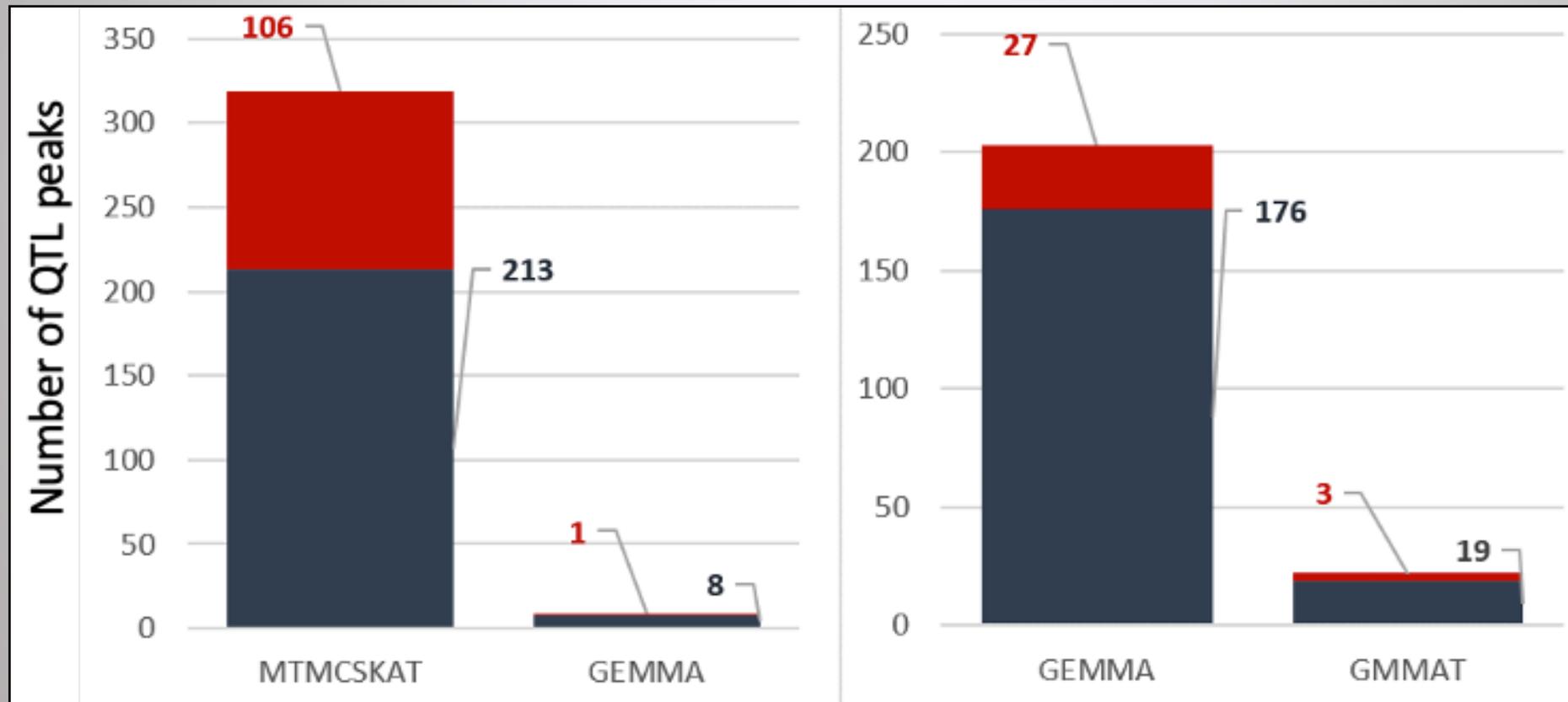
Example shown for *in vitro* transf./regen.

Distance from gene



FDR ($\alpha = 0.10$)

ART-Bonferroni



Association mapping with MTMC-SKAT, ART can be applied for diverse GWAS datasets

- MTMC-SKAT R extension available on GitHub
<https://github.com/naglemi/mtmcskat>
- ART theory and algorithm previously published (Vsevolozhskaya, et al., 2019; Front. Genet.), our implementation available upon request

Agenda

I. Background

Motivation, genetic resources

II. High-throughput phenomics

RGB/hyperspectral methods developed

III. Association mapping methods & results

Power, precision, and selected results

IV. Gene candidate roles/interpretation

From SNPs to genes, genes to pathways

Steps in gene candidate interpretation

What genes are near to SNPs?

- Genome annotation, automated mining
- SNP peak shape, annotation
- Manual inspection in relation to likely regulatory regions (synonymous)

What is gene function?

- Literature review, gene and close homologs
- Sequence alignment

What biological processes are implicated by multiple genes?

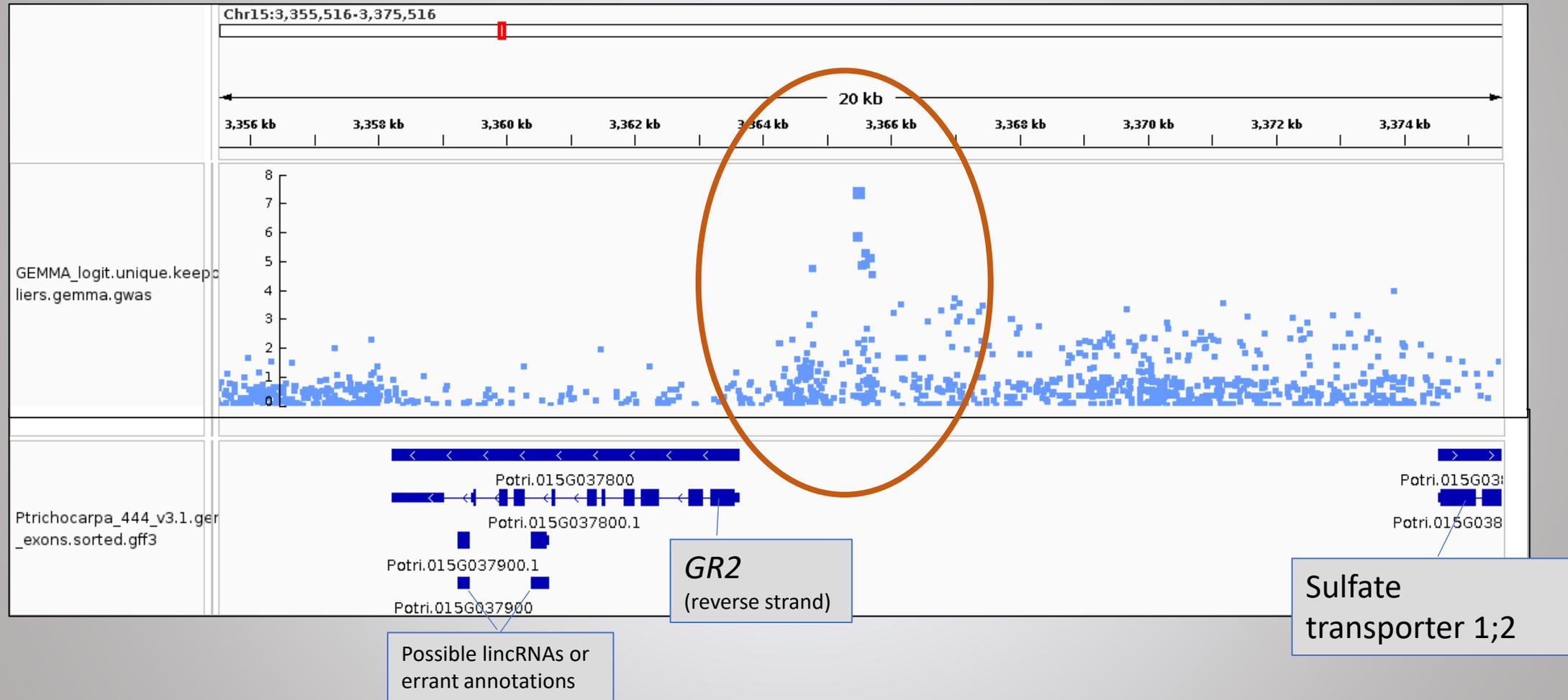
- Deeper literature review
- Gene ontology
- KEGG
- Database integration
- Network analysis

Validation, value for biotechnology

- Mutant studies
- Functional association mapping

Peaks very often in promoter region

Example: ~2kb upstream of GLUTATHIONE REDUCTASE 2 (GR2)

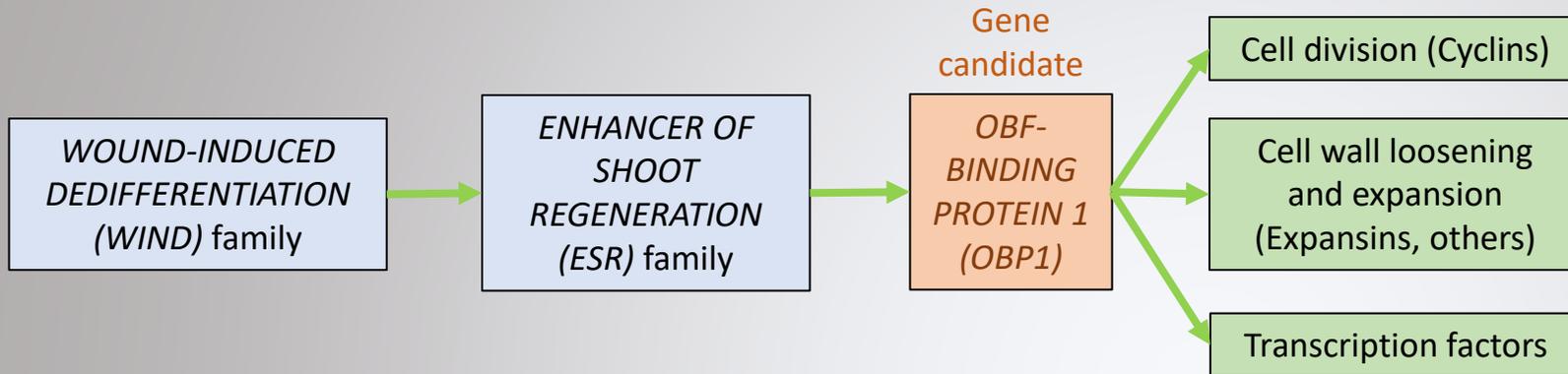


Many processes represented prominently by gene candidates across callus/shoot/root regeneration

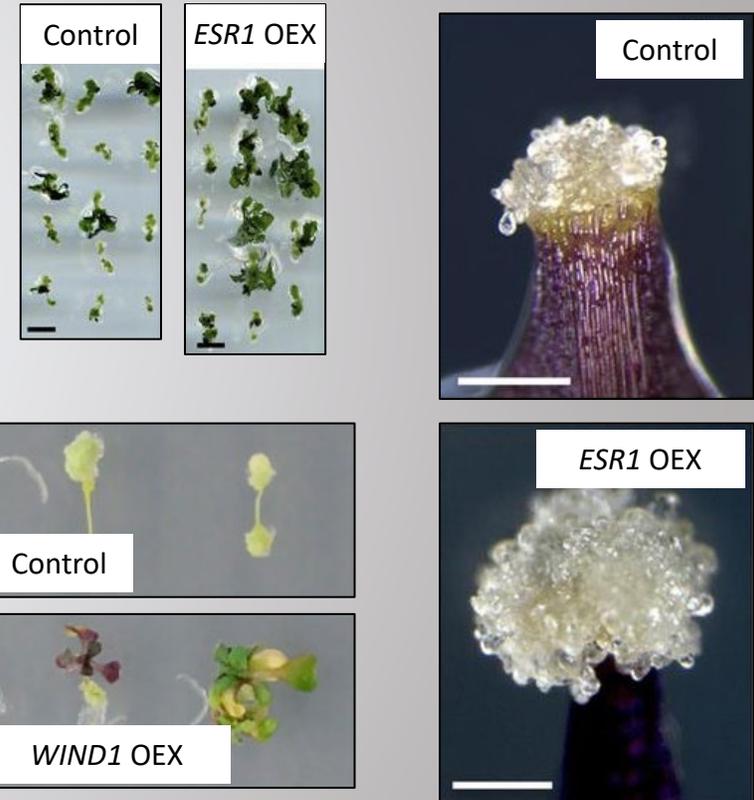
- Hormone signaling
 - Auxin, cytokinin control of cell expansion and division
 - Jasmonate, ethylene, salicylic acid control of wounding, stress
- Cell wall structural genes
- Cell cycle control
- Reactive oxygen species (ROS) regulators
- Numerous genes of unknown function, unclear homology to characterized genes
- Numerous SNPs distant from, or in unusual locations, relative to annotated genes

Building functional networks: *OBP1* example

OCS-binding-protein forms a putative hub bridging wound response, callus, and shoot formation



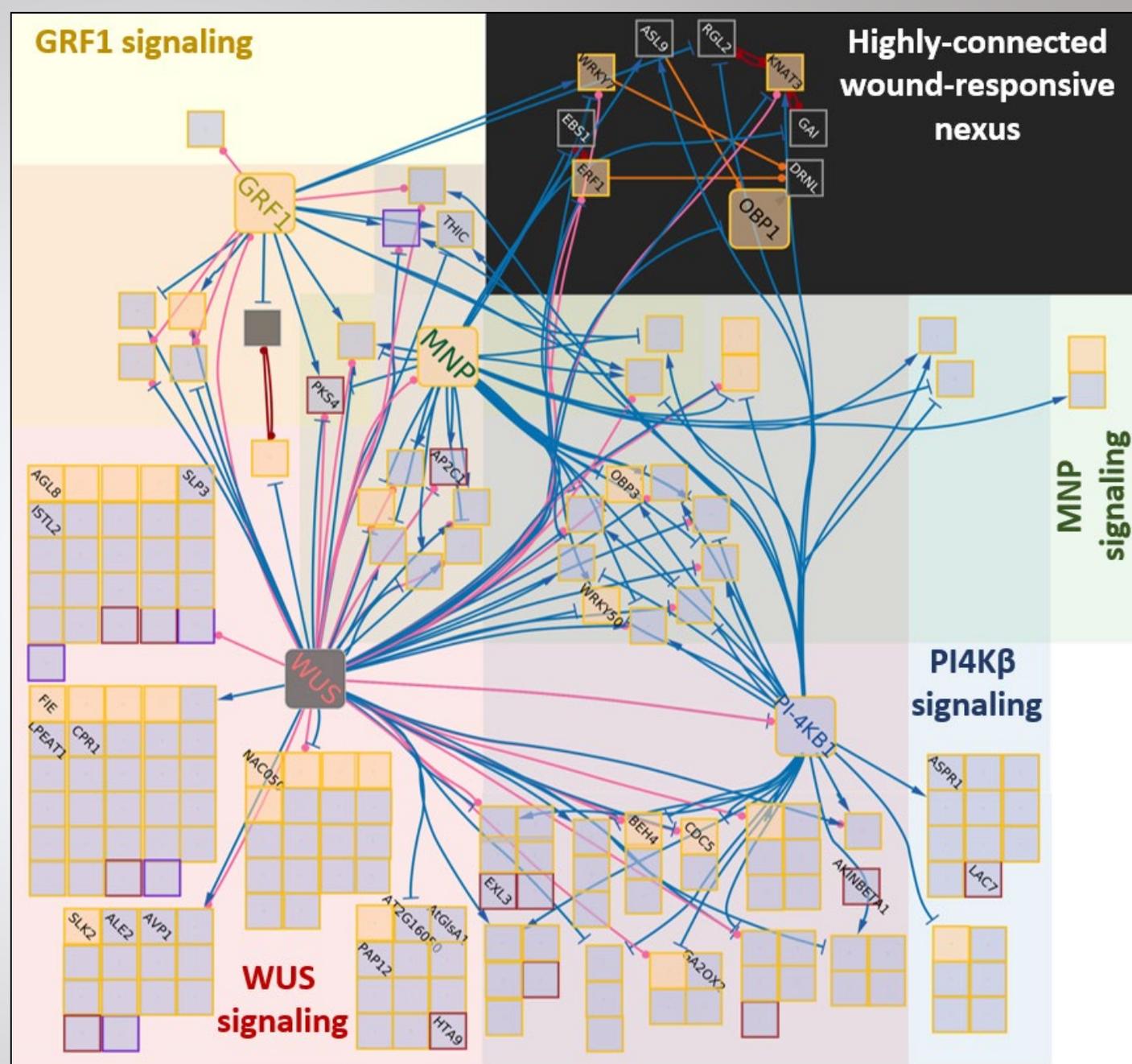
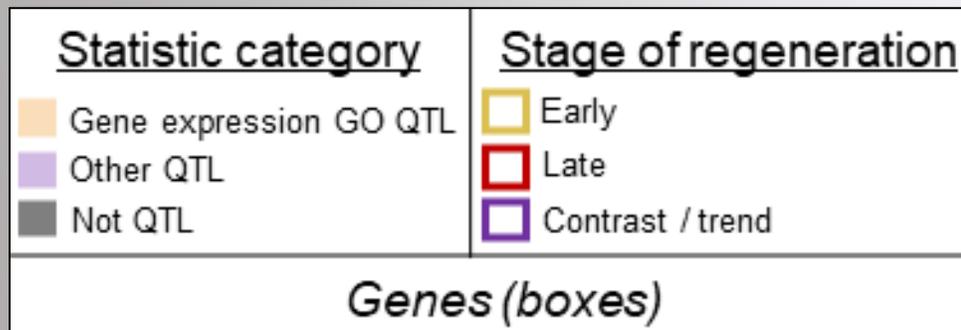
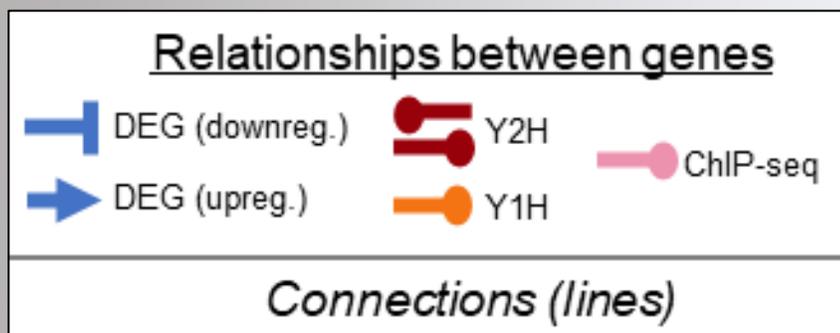
Overexpression of genes upstream of *OBP1* enhance callus, shoots [3-5]



Related experimental evidence [1-5]

1. Skirycz, Aleksandra, et al. "The DOF transcription factor OBP1 is involved in cell cycle regulation in *Arabidopsis thaliana*." *The Plant Journal* 56.5 (2008): 779-792.
2. Ikeuchi, Momoko, et al. "A gene regulatory network for cellular reprogramming in plant regeneration." *Plant and Cell Physiology* 59.4 (2018): 770-782.
3. Banno, Hiroharu, et al. "Overexpression of *Arabidopsis* ESR1 induces initiation of shoot regeneration." *The Plant Cell* 13.12 (2001): 2609-2618.
4. Iwase, Akira, et al. "WIND1 promotes shoot regeneration through transcriptional activation of ENHANCER OF SHOOT REGENERATION1 in *Arabidopsis*." *The Plant Cell* 29.1 (2017): 54-69.
5. Iwase, Akira, et al. "The AP2/ERF transcription factor WIND1 controls cell dedifferentiation in *Arabidopsis*." *Current Biology* 21.6 (2011): 508-514.

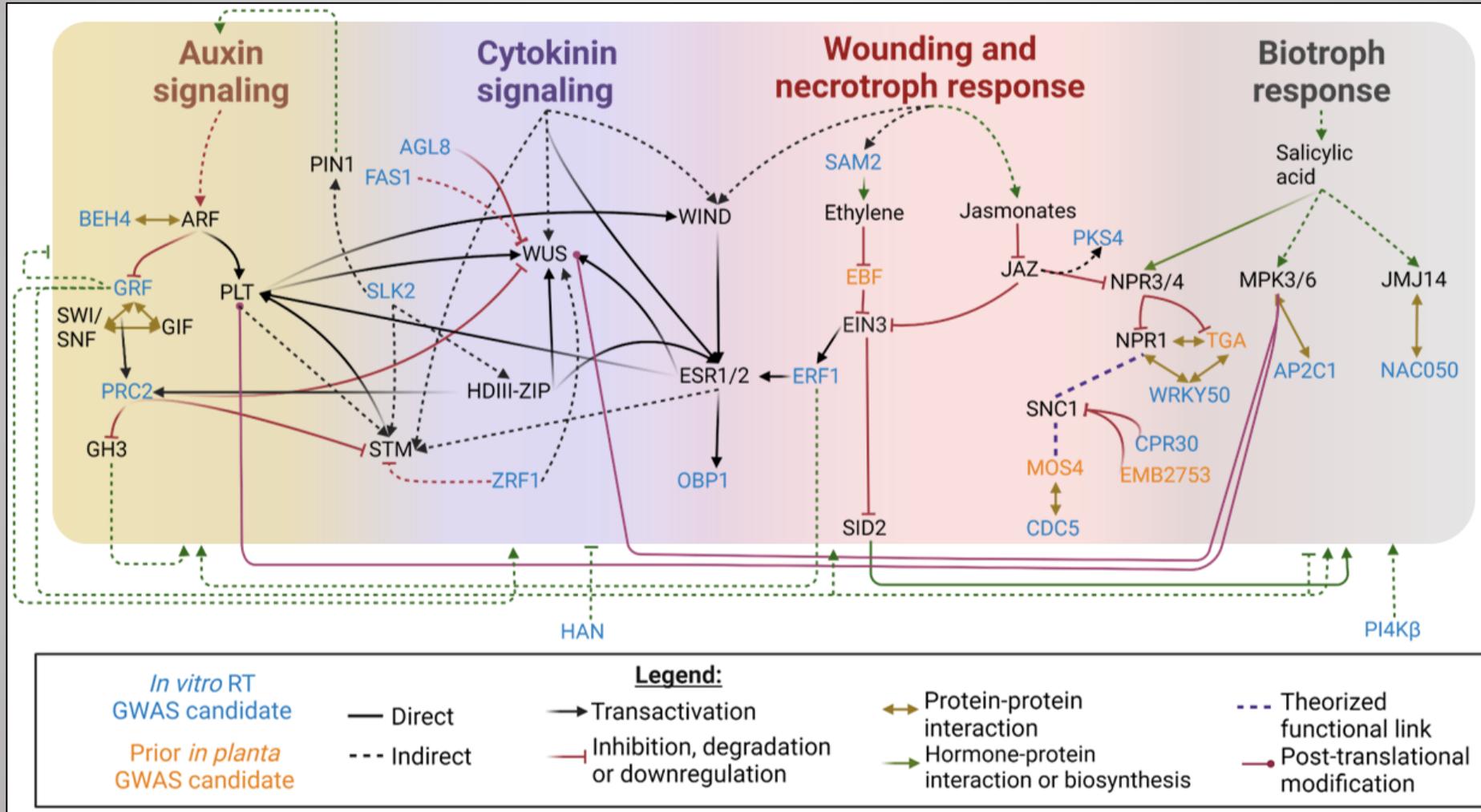
Networks highly interconnected - span auxin-cytokinin, wound-stress subnetworks



Mechanistic connections between gene candidate Arabidopsis homologs from *in vitro*, *in planta* experiments across ~80 studies

A network summary

Not only auxin/cytokinin -- extensive wound/stress processes



Summary

- Powerful GWAS genetic resource employed
- Novel phenomic systems developed
 - Visible and hyperspectral imager built
 - Image analysis algorithms, machine vision pipelines
- Multilocus GWAS methods adapted, essential for detection
 - More than one-thousand candidates
- Results underline the critical importance and connections to wound/stress signaling
 - Importance of common wounding, pathogen challenge, hormone treatments in RT
- Highly polygenic and diversified genes and pathways among, root, shoot, and *in vitro* experiments
- Responses are so variable because of extensive genetic variation in numerous genes, from diverse and interacting pathways, in determining RT responses
- Numerous pathways, and specific genes that modulate them, candidates for functional analysis and use in biotechnology

Thanks to NSF, DOE, & GREAT TREES industry consortium, and to a great research team, that made the work possible

Strauss Lab
Tissue culture,
transformation,
phenotyping,
clone management

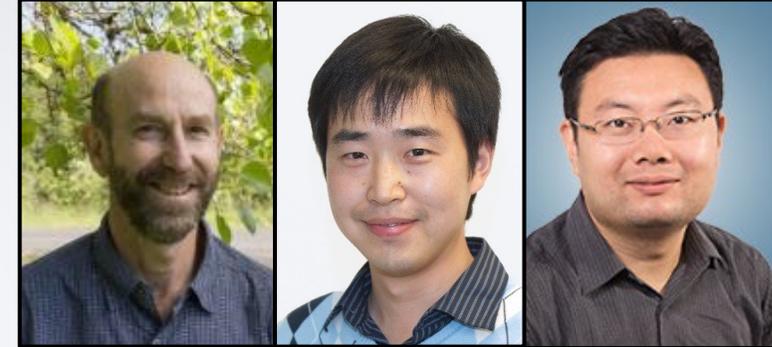


Cathleen Ma
Sr. FRA

Kate Peremyslova
FRA

Bahiya Zahl
Undergraduate
student

Amanda Goddard
Lab manager



Prof. Steve Strauss
PI

Prof. Yuan Jiang
Co-PI, Statistics

Prof. Fuxin Li
Co-PI, machine vision

Leadership

Li lab
RGB deep
segmentation and
classification,
annotation
interface



Jialin Yuan
PhD Student

Damanpreet Kaur
MSc student

Jia Yi Li
Undergraduate
student



OAK RIDGE
National Laboratory

