



Oregon State University



Genome-wide association studies of regeneration in *Populus* with machine vision and hyperspectral phenomics

Michael Nagle¹, Jialin Yuan², Zheng Zhou³, Nihar A. Doshi⁴, Fuxin Li², Cathleen Ma¹, Anna Magnuson¹, Chris Draves⁵, Howland Jones⁵, Natalie Crothers⁵, Gabor Kemeny⁵, David Braumann⁵, Yuan Jiang⁶, Ritesh Mewalal¹, and Steven H. Strauss¹
¹Department of Forest Ecosystems & Society, Oregon State University, Corvallis OR; ²Department of Computer Science, Oregon State University, Corvallis OR; ³Tencent Inc., Shenzhen, China; ⁴University of Southern California, Los Angeles, CA; ⁵Middleton Spectral Vision, Middleton WI; ⁶Department of Statistics, Oregon State University, Corvallis OR

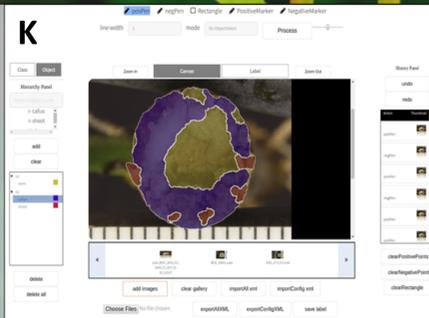


Introduction

- The recalcitrance of many plant species to *in vitro* regeneration methods prevents researchers from obtaining non-chimaeric transformants following genetic transformation.
- Populus trichocarpa* is an ideal model organism for studying regeneration due to amenability to *in vitro* and *in vivo* regeneration protocols.
- Genome-wide association studies (GWAS) can benefit from advances in phenotyping, particularly high-throughput phenotyping, hyperspectral imaging and machine vision, to capture a breadth and depth of data that would be practically unobtainable by humans unassisted by machines.
- To obtain both RGB and hyperspectral data on a high-throughput scale for *in vitro* plant tissue cultures, we are using the *macroPhor*TM Array, a custom instrument from Middleton Spectral Vision, and the accompanying software suite *KemoQuant*TM to deconvolute spectra into individual components, including chlorophylls, fluorescent reporters and spectral shifts.
- Deep learning models using convolutional neural networks are trained to segment images by tissue type (i.e. callus, shoot) after learning from images annotated by a graphical user interface, thus enabling high-throughput and precise analysis of images.

Machine vision

(K) Through a graphical user interface, specific tissues of interest are annotated by the user with assistance by an edge detection algorithm. These annotated images are then used to train convolutional neural networks for image segmentation. Project aims includes deployment of this interface on a web server to assist researchers using machine vision.

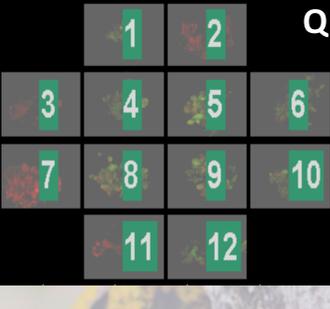
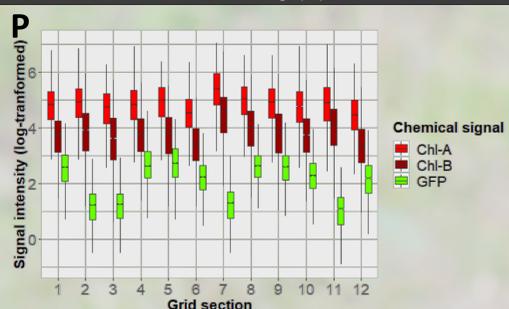
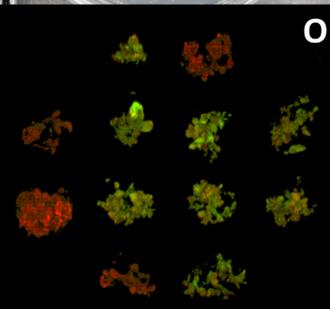
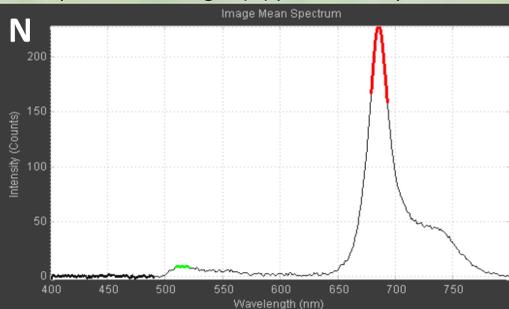


After learning from user annotations, deep learning models classify sections of images by the type of tissue recognized (L). The accuracy is measurable by comparison to user-annotated images not included in the training set and depends on the type of deep learning architectures used (VGG-16 and PSPNet). Images below are segmented into callus (blue), shoot (green) stem (blue) and background.



Hyperspectral analysis

The *macroPhor*TM Array (middletonspectralvision.com) is used to capture RGB (M) and hyperspectral images (N-Q) for trays of petri dishes. Deep learning models, while depending on user annotation of RGB images, will also incorporate channels for spectral components provided by hyperspectral analysis. False color applied to GFP and chlorophyll spectral peaks (N) enables qualitative inspection of spectral components of images (O) prior to analysis.

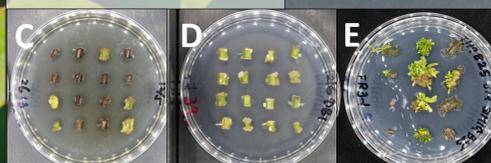


With the software *KemoQuant*TM, images are segmented by explant (Q) and spectral components are identified by multivariate curve resolution and quantified by least squares linear regression (P). For each component, signal intensities are computed per pixel (providing additional image channels to be used for machine vision training and prediction of tissue type) and per explant (to be used directly in GWAS).

Regeneration

In vitro regeneration

Optimized callus and shoot induction hormone treatments for *in vitro* regeneration from stem and petiole were selected following heritability testing to determine which cause the widest range in response across genotypes (data not shown). Tissue culturing and plant phenotyping has begun and will run through 2019, with 1200 genotypes to be studied.



(A) Stem explant undergoing callus induction and (B) shoot regeneration; petri dishes of stem explants undergoing (C) necrosis, (D) various stages of callus induction, and (E) later stages of callus growth and shoot regeneration

Stem regeneration

To enhance the natural wounding and regeneration response, 0.5mg/mL thiazuron (a synthetic cytokinin) was applied to the tips of cut stems. Phenotyping has been completed for 590 of 902 genotypes to be included in the stem regeneration study, and this partial dataset is being used to test and refine GWAS methods.

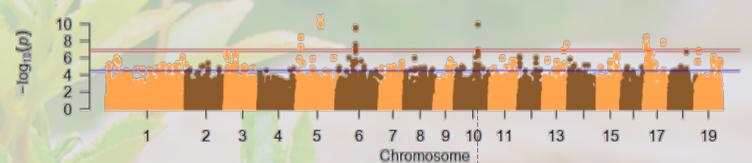


(F-G) Side view of shoots regenerating from tips of cut stems placed in water; View of stem tip showing several phenotypes recognized by machine vision: (H) green callus; (I) green callus and shoot; (J) red callus and shoot

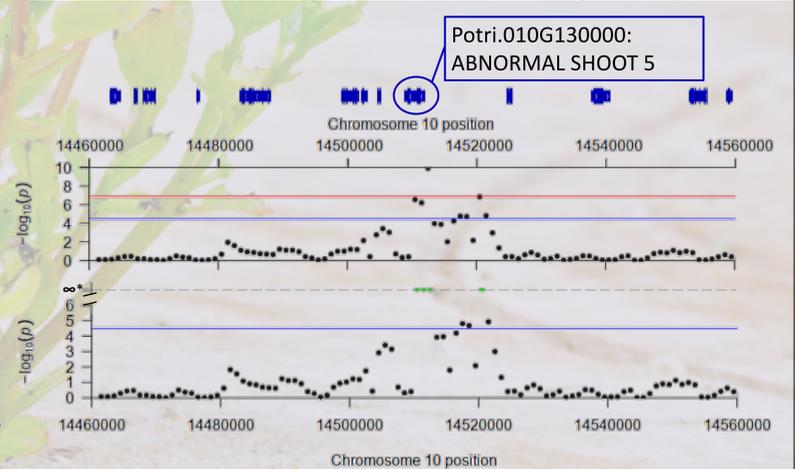
Association testing

The GWAS method Sequence Kernel Association Test (SKAT) was used to collapse SNPs into 3kb windows based on physical location and test for associations with traits of interest. To control for non-normality of data, resampling was applied. This test reveals a possible association of shoot area with a homolog of a known shoot regulator in *Arabidopsis*.

Genetic markers associated with shoot area: Genome-wide view



View zoomed to chromosome 10 subsection, aligned to gene track



Without resampling

With resampling (up to 1 million permutations)

*The lowest p-values (green) will require more resampling to validate

Outlook

- Convolutional neural networks and hyperspectral imaging provide new opportunities for genetic discovery by enabling precise, high-throughput phenotyping of complex traits.
- While currently, most genes in forest species and other non-model plants remain uncharacterized, the advent of high-throughput hyperspectral phenotyping with machine vision may enable rapid elucidation of these genes via improved power and accuracy for experiments to study gene-function relationships.
- Characterization of the genetic basis of regeneration offers opportunities for converting poor responders to regeneration into efficient responders via overexpression or knockdown/out of developmental regulators. This may enable robust transformation of genotypes and species which genetic engineering methods cannot be efficiently applied to yet.

Acknowledgements



We thank the National Science Foundation Plant Genome Research Program for support (IOS # 1546900, Analysis of genes affecting plant regeneration and transformation in poplar), and members of the Tree Biosafety and Genomics Research Cooperative at OSU for its long-term support of the Strauss laboratory.

