Analysis of Genes Affecting Plant Regeneration and Transformation

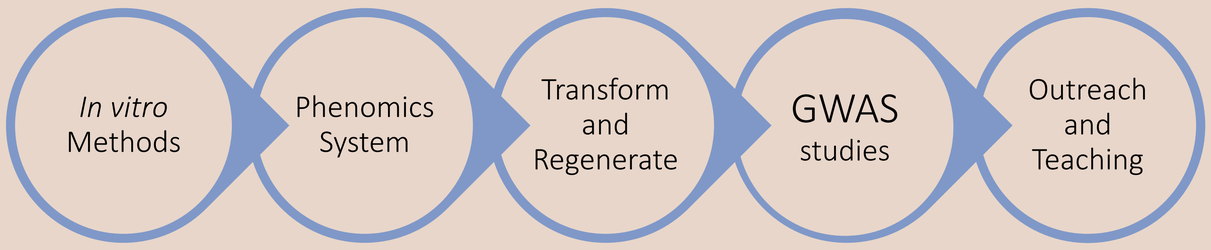
Brett R. Pierce[1], Cathleen Ma[1], Anna Magnuson[1], Amy Klocko[1], Michael Nagle[1], Wellington Muchero[2], Jialin Yuan[3], Yuan Jiang[4], Fuxin Li[3], and Steven H. Strauss[1]

[1]Department of Forest Ecosystems and Society, Oregon State University, Corvallis OR; [2]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge TN; [3]Department of Electrical Engineering and Computer Science, Oregon State University; and [4]Department of Statistics, Oregon State University

Project Overview

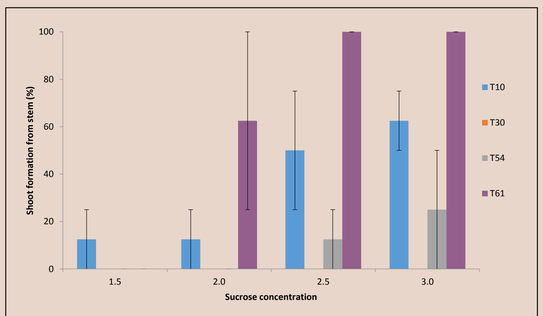
Regeneration of differentiated organisms from single cells is a critical need for functional genomics and for the production of genetically engineered organisms. The project is conducting a genome-wide investigation of the genes that control regenerability and transformation (RT) in *Populus*, which is one of the best studied crop species with respect to these traits. The project will identify genetic elements that control RT, develop novel phenomic methods based on image analysis, and develop new social science and education methods for teaching about genetic engineering to diverse high school students and teachers.

Our specific objectives are to (1) explore a variety of RT methods to maximize variation in RT responses; (2) develop new phenomic tools, including an image capture and generalizable machine-vision system, to precisely determine *in vitro* phenotypes; (3) using GWAS, map sets of alleles that are associated with variation in RT frequency; and (4) study cognitive processes with respect to GE crops, develop case studies and new teaching materials, deliver them to rural and underserved communities in the Pacific Northwest—and through publications, social media, and conferences share the project's insights and teaching modules internationally.

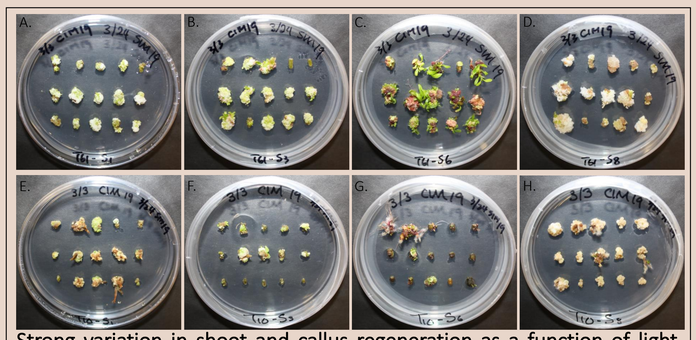


In Vitro Optimization Experiments

The goal of this phase of work is to identify a series of treatments that maximize genetic variance and minimize environmental variance in the GWAS population – in other words to identify treatments that give maximal heritability and thus enhance GWAS efficiency. In a factorial experiment, we tested 36 basal media types and a number of other key culture factors on a small number of genotypes, then a few selected treatments on ten to 16 genotypes. Factors tested included nitrogen and meso-nutrients, sucrose (1.5, 2.0, 2.5, and 3.0%), light spectra, and auxin types and concentrations. We have included two examples below. High sucrose concentration (3%) promotes shoot regeneration on most genotypes, and blue/red and full spectrum lighting give the best organogenesis.



Increased shoot regeneration at high concentrations of sucrose (from left to right, 1.5 to 3.0%, as graphed above).



Strong variation in shoot and callus regeneration as a function of light quality and quantity. Full-spectrum fluorescent lights (c) gave best shoot regeneration. A) R,G,B,FR. B) hi R,B. C) white, full spectrum (green/yellow, some blue, a little red, and almost no far-red) used in our tissue culture growth room. D) R, hi B. E) R,G,B,FR. F) hi R,B. G) white, full spectrum (green/yellow, some blue, a little red, and almost no far-red) used in our tissue culture growth room. H) R, hi B. Variation in light response among genotypes is also marked (top vs. bottom row).

In Vivo Callus and Shoot Induction

***In vivo* optimization experiments were performed on a subset of genotypes to determine a standard protocol for inducing callus and shoot growth on dormant cuttings.** Optimal conditions were those that would yield reliable growth while also allowing us to see maximum genetic variation between genotypes under a design that is efficiently reproducible for large GWAS experiments. Hormone concentration, environment, growing media, and moisture-retention methods were among the studied variables. We found that the most favorable and varied growth occurred when cuttings were grown in water under stable (± 5°) room temperature conditions and received a low concentration TDZ solution applied directly to the freshly cut stem tip. An open Eppendorf tube was then placed over the stem tip to maintain high moisture conditions around new growth.

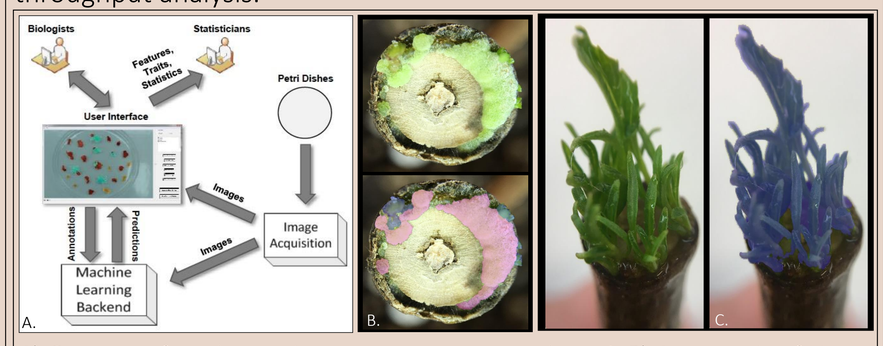
A) A cutting growing in H_20 being propped up to allow roots to grow unobstructed. B) This image shows the efficacy of moisture control on the stem tip with an Eppendorf tube capping it. C) 400 cuttings being grown in our Richardson Hall head house at initial planting and D) at three weeks of growth . E) Stem tips showing healthy shoots growing out of callus tissue after TDZ was applied to stem tip based on prior studies in which we were optimizing treatments and growth conditions.

Stem tips showing callus and shoot formation at week 3 after 0.5mg/L of TDZ was applied directly after cutting stems and inserting cuttings in H_20. Shown are different genotypes responding under the same treatment and conditions.

A scientist standing below a wild cottonwood growing in Oregon, similar to those used as germplasm for our studies.

Machine Vision and Graphical User Interface

The goal is to accurately recognize and quantify individual tissue types and sort them into pre-defined categories (e.g., roots, shoots, and calli), as well as to recognize more complex and unfamiliar physiological states (e.g., oxidant stressed, pigmented, necrotic, and modified organs). The annotation UI is currently using the GrabCut algorithm for image segmentation which relies on a user-defined bounding box in which an object will be analyzed based on color distribution of the target object. A deep learning algorithm, once optimized, will be integrated to further the UI's capabilities. The integration of the deep learning algorithm into the UI will also increase predictive power as well as drastically reduce processing time, therefore enhancing high-throughput analysis.



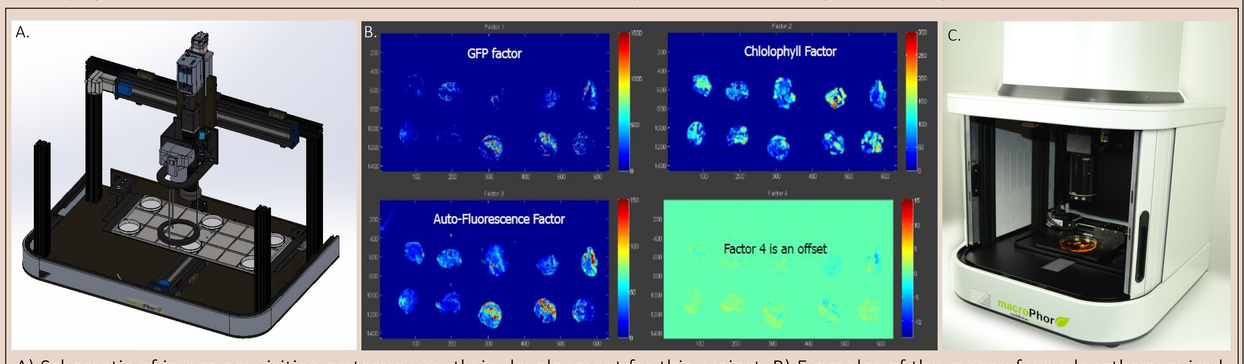
A) Overview of machine vision system components planned. B) An example of shoots being recognized at both the new regenerant (B) and full expansion (C) stages as shown with false-color. Callus is annotated in pink. Annotation was done using the GrabCut algorithm. The graphical user interface is currently in development.

Image Acquisition System

The ability to obtain large amounts of quantitative, precise *in vitro* data that reflect the complex developmental biology of regenerating tissues is challenging, but critical for GWAS of regeneration traits. We therefore set out to design a new phenomics system that will yield precise and physiologically sensitive measurements.

We contracted with Middleton Spectral Vision (Middleton, WI) to design and building a high-throughput plant imaging scanner for the needs of this project. The system will be equipped with one high sensitivity fluorescence/VNIR hyperspectral camera as well as one high resolution RGB digital camera. Both of these cameras will be on a motorized track which can be programmed to move into any position in the grid that the user defines to capture the image of the Petri dishes. This system will be able to house a custom designed tray holding up to 21 Petri dishes enabling high-throughput imaging. The tray is being engineered to enable rapid transfer and return to our growth chambers. The images files will then be subject to machine vision analyses to identify and quantify different developmental states (see above).

Below we show schematic of the instrument as now planned, scanning Petri dishes (A), and test runs of GFP fluorescence in transformed poplar tissues scanned at a variety of hyperspectral wavelengths to maximize GFP, chlorophyll, and innate fluorescence (B) as determined with a parent instrument produced by Middleton, the MacroPhor.



A) Schematic of image acquisition system currently in development for this project. B) Examples of the range of wavelengths acquired by the imaging systems hyperspectral camera. Quantification of the spatial locations of transient expression of GFP is one example of a major benefit of this imaging system. C) The fluorescent images above were acquired by the macroPhor hyperspectral imaging system which a modified version of this system for high-throughput imaging is currently being designed and built for this project.

Genome Wide Association Study

As part of plant propagation, we set up a GWAS study of *in vivo* rooting. Rooting and plant size phenotypes were scored both manually and using machine vision analysis (panel B below), and Principle Components Analysis used to combine these datasets. Efficient mixed model association (EMMA), accounting for kinship, was used to correlate a panel of 8.2 million SNPs to phenotypic variation data and associated Principal Component scores. This population has a panel of 29 million SNPs representing a marker every 17-base across the genome and rapidly decaying linkage disequilibrium that falls below 0.2 within 3Kb.

Over 70 SNP loci were associated with one or more traits at a false discovery rate (FDR) threshold of 0.05. Significance thresholds were set at negative log p-values of 6 for PC score runs and 5.5 for raw data. Significant SNP loci were investigated further using Phytozome, *Populus trichocarpa* v3.0 reference genome. An example of two significant SNPs and nearby genes are given below (panel A).

1	23802203	6.04	Potri.001G226600.1	Sugar transporter ERD6-like 14-related, high expression in roots
10	14966452	6.47	Potri.010G136100.1	Close to genes, first gene listed is a basic helix-loop-helix family protein, similar to bHLH transcription factor GBOF-1 with high root expression, Arabidopsis homolog regulates growth
			Potri.010G135900.1	WPP domain-interacting tail-anchored protein differentially downregulated in root in urea vs. standard, Arabidopsis homolog WIT2, association in root tip cells

A) SNPs of interest are circled in red and their chromosome, position, LOD value, nearby gene(s) that were annotated, and functional annotations for those genes are shown. B) Segment masks produced by color-based and morphology segmentation from preliminary rooting experiment. Root area and mean stem diameters were computed from the segmentations and used for PCA and GWAS.

Summary

- Preliminary *in vitro* optimization studies showed that optimal shoot formation resulted from lower nitrogen salt concentrations in media, higher sucrose concentration , and full spectrum and high light intensities.
- Preliminary *in vivo* optimization studies showed that optimal callus formation followed by shoot formation resulted from being grown in H_20, stem tips capped with Eppendorf tube to maintain moisture, temperature maintained at ± 5° room temperature, and for shoot formation TDZ being applied directly after cutting stem tips and planting.
- Currently in development is a web-based annotation interface that will allow us to quickly annotate various organ types via machine vision which will incorporate a deep learning algorithm.
- A GWAS study of *in vivo* rooting identified dozens of significant SNPs, and a GWAS study of callus and shoot formation from cut stems is underway.

Acknowledgments

We thank the National Science Foundation Plant Genome Research Program for support (IOS # 1546900, Analysis of genes affecting plant regeneration and transformation in poplar). We also thank the members of the Tree Biosafety and Genomics Research Cooperative at OSU for its long term support of our transformation and regeneration studies in poplar. We thank Middleton Spectral Vision for their collaboration and phenomic system imagery. We also thank Dr. Kevin Folta (professor and chairman of the horticultural sciences department at the University of Florida) for collaboration on our light spectra studies.