Identifying the Genomic Basis of Variation in Adventitious Rooting in *Populus*

Steve Strauss
Oregon State University / USA
Coauthors

• West Virginia University
  – Christine Zawaski – rooting data
  – David Macaya-Sanz – GWAS bioinformatics
  – Stephen P. DiFazio – GWAS lead
  – Jonathan R. Cumming – root physiology lead

• Oregon State University
  – Cathleen M and Brett Pierce – rooting data
  – Anna C. Magnuson – Phytozome analysis
  – Jialin Yuan – Machine vision
  – Yuan Jiang – Statistics
  – Fuxin Li – Machine vision

• Oak Ridge National Laboratory
  – Wellington Muchero – GWAS bioinformatics
Coauthors – WVU

Stephen DiFazio

David Macaya Sanz

Jonathan Cumming

Christine Zawaski
Coauthors – OSU and ORNL

- Wellington Muchero, ORNL
- Fuxin Li
- Cathleen Ma
- Anna Magnuson
- Yuan Jiang
- Jialin Yuan
Sources of grant support

• NSF Plant Genome Research Program (IOS # 1546900) to Oregon State University

• USDA National Institute for Food and Agriculture (2014-67013-21657) to West Virginia University
Adventitious rooting – new root formation

“In botany, adventitious refers to structures that develop in an unusual place.”

Wikipedia
Adventitious roots can develop from different cell types, affected by genotype, environment, epigenetics

- Usually develop near vascular tissues and vascular cambium, but location varies widely among tissues and species
- Also derived from callus that forms from wounded tissues at cut surfaces – like in many other forms of plant and animal regeneration processes
- Affected by plant age and tissue history, thus affected by epigenetic processes
- Rooting capacity varies widely among species and genotypes - an adaptive trait
- A model system for biological regeneration
Adventitious rooting is an important trait in forestry and horticulture

• Major means for cloning - thus exploiting genetic variation in heterozygous, outcrossing woody species
  – Inbreeding, inbred lines not an option
  – Captures all additive and non-additive genetic variance
  – High uniformity compared to sexual propagules

• Major propagation tool where grafting too costly for use
  – Common use of dormant cuttings (“sticks”)

• Also often problematic
  – Aging and physiological effects impart transient variation
  – *In vitro* behavior, as is cellular regeneration and genetic transformation for many plant species

• Even with auxin treatments, genetic variation in rootability is often high, making use impossible or reducing genetic base for breeding
Poplar stands from rooted cuttings
Rooting also of major adaptive value in *Populus* and other species

- Adapted to vegetative propagation in the wild
  - Shoots sprouting from roots – aspens
  - Stems producing new roots – cottonwoods

- Can see rooting of stems in nature
  - Spread of branches in river corridors
  - Stems fallen along rivers

- Thus an evolved trait under natural selection
  - Intraspecific and species level polymorphism
    - Many tree species will not root at all without artificial treatments
Broad goals of our studies

• Explore the use of GWAS to help understand adventitious rooting – focus today
• Identify candidate genes that affect the rate and mode of rooting
• Conduct validation experiments to provide insight into the physiological processes that control rooting
• Use genes and process insights to enable new methods to control rooting in horticulture and forestry
Approach

• Take advantage of a resequenced collection of wild black cottonwoods (*P. trichocarpa*) for genetic mapping
  – Department of Energy, Oak Ridge National Laboratory, Bioenergy Science Center resource

• Trees sampled from British Columbia to California

• Very low linkage disequilibrium in outcrossing, wild trees – markers should be “near” to causative genes/regulatory elements
Materials and measurements

- Experiments in Oregon (OSU, 537 genotypes) and West Virginia (WVU, 545 genotypes)
- Cuttings rooted in water and soil (OSU) or water and 0.5 mM CaNO$_3$ solution (WVU)
- At periodic intervals, root initiation and growth measurements taken
OSU measurement system
Machine vision analysis – OSU (n=423)
WVU measurement system
Phenotypes analyzed (visual scores, ImageJ, machine vision)

Examples

• Root score (number of major roots class)
• Longest root
• Root density (image)
• Days to root initiation
• Root growth rate
• Principal component scores

<table>
<thead>
<tr>
<th></th>
<th>PC1</th>
<th>PC2</th>
<th>PC3</th>
<th>PC4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Height</strong></td>
<td>0.44</td>
<td>-0.43</td>
<td>0.77</td>
<td>0.15</td>
</tr>
<tr>
<td><strong>Diameter</strong></td>
<td>0.65</td>
<td>0.19</td>
<td>-0.13</td>
<td>-0.73</td>
</tr>
<tr>
<td><strong>Ave. Stem Diameter</strong></td>
<td>0.61</td>
<td>0.25</td>
<td>-0.34</td>
<td>0.66</td>
</tr>
<tr>
<td><strong>Root Area</strong></td>
<td>0.10</td>
<td>-0.85</td>
<td>-0.52</td>
<td>-0.04</td>
</tr>
</tbody>
</table>
Numerous rooting traits derived

<table>
<thead>
<tr>
<th>Traits</th>
<th>Subset of tested individuals</th>
<th>Type of data</th>
<th>Processing of raw data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days to root initiation (WVU)</td>
<td>all</td>
<td>Integer</td>
<td>BLUPs</td>
</tr>
<tr>
<td>Root initiation score (WVU)</td>
<td>all</td>
<td>Discrete (score)</td>
<td>BLUPs</td>
</tr>
<tr>
<td>Longest root length (WVU)</td>
<td>all</td>
<td>Continuous</td>
<td>BLUPs</td>
</tr>
<tr>
<td>Total root length (WVU)</td>
<td>all</td>
<td>Continuous</td>
<td>BLUPs</td>
</tr>
<tr>
<td>Root density (WVU)</td>
<td>all</td>
<td>Continuous</td>
<td>BLUPs</td>
</tr>
<tr>
<td>Root growth rate (WVU)</td>
<td>all</td>
<td>Continuous</td>
<td>BLUPs</td>
</tr>
<tr>
<td>Average root length (WVU)</td>
<td>all</td>
<td>Continuous</td>
<td>BLUPs</td>
</tr>
<tr>
<td>Root number (WVU)</td>
<td>all</td>
<td>Integer</td>
<td>BLUPs</td>
</tr>
<tr>
<td>Mode of rooting score (WVU)</td>
<td>all</td>
<td>Discrete (score)</td>
<td>BLUPs</td>
</tr>
<tr>
<td>Basal rooting aptitude (WVU)</td>
<td>all</td>
<td>Binary</td>
<td>Raw data</td>
</tr>
<tr>
<td>Lateral rooting aptitude (WVU)</td>
<td>all</td>
<td>Binary</td>
<td>Raw data</td>
</tr>
<tr>
<td>Longest root length (OSU)</td>
<td>all</td>
<td>Continuous</td>
<td>BLUPs</td>
</tr>
<tr>
<td>Root score (OSU)</td>
<td>all</td>
<td>Discrete (score)</td>
<td>BLUPs</td>
</tr>
<tr>
<td>Root area (OSU)</td>
<td>all</td>
<td>Continuous</td>
<td>BLUPs</td>
</tr>
<tr>
<td>Root score in soil (OSU)</td>
<td>all</td>
<td>Discrete (score)</td>
<td>BLUPs</td>
</tr>
<tr>
<td>Longest root length (WVU)</td>
<td>only lateral-rooting genotypes</td>
<td>Continuous</td>
<td>BLUPs</td>
</tr>
<tr>
<td>Total root length (WVU)</td>
<td>only lateral-rooting genotypes</td>
<td>Continuous</td>
<td>BLUPs</td>
</tr>
<tr>
<td>Root growth rate (WVU)</td>
<td>only lateral-rooting genotypes</td>
<td>Continuous</td>
<td>BLUPs</td>
</tr>
<tr>
<td>Root density (WVU)</td>
<td>only lateral-rooting genotypes</td>
<td>Continuous</td>
<td>BLUPs</td>
</tr>
<tr>
<td>Average root length (WVU)</td>
<td>only lateral-rooting genotypes</td>
<td>Continuous</td>
<td>BLUPs</td>
</tr>
<tr>
<td>Longest root length (WVU)</td>
<td>only basal-rooting genotypes</td>
<td>Continuous</td>
<td>BLUPs</td>
</tr>
<tr>
<td>Total root length (WVU)</td>
<td>only basal-rooting genotypes</td>
<td>Continuous</td>
<td>BLUPs</td>
</tr>
<tr>
<td>Root growth rate (WVU)</td>
<td>only basal-rooting genotypes</td>
<td>Continuous</td>
<td>BLUPs</td>
</tr>
<tr>
<td>Root density (WVU)</td>
<td>only basal-rooting genotypes</td>
<td>Continuous</td>
<td>BLUPs</td>
</tr>
<tr>
<td>Average root length (WVU)</td>
<td>only basal-rooting genotypes</td>
<td>Continuous</td>
<td>BLUPs</td>
</tr>
</tbody>
</table>
Quantitative methods

- Linear mixed models were used to estimate heritabilities and BLUPs of the phenotypes taking into account experimental design (blocking, replicates).

- OSU/Oak Ridge: GWAS using efficient mixed model association (EMMAX), accounting for kinship, was used with a panel of 8.2 million SNPs.

- WVU: Genome-wide efficient mixed model association (GEMMA), accounting for kinship, was used to correlate a panel of 13 million markers to phenotypic variation.
  - Simulations (permutation analysis) were conducted to better control for non-normality of data, imbalance in SNPs, and variable false discovery rates: David Maycaya-Sanz.
Results

• ~80% of genotypes rooted, 20% did not or to very limited degree

• Rooting traits had highly statistically significant variation, with heritabilities near 20% for most measures
Broad sense heritabilities were low in both experiments

<table>
<thead>
<tr>
<th>OSU</th>
<th>$h^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Root number score - Combined</td>
<td>0.13</td>
</tr>
<tr>
<td>Water</td>
<td>0.18</td>
</tr>
<tr>
<td>Soil</td>
<td>0.15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>WVU</th>
<th>$h^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days to Root Initiation (DRI)</td>
<td>0.25</td>
</tr>
<tr>
<td>Longest Root Length (LRL)</td>
<td>0.17</td>
</tr>
<tr>
<td>Total Root Length (TRL)</td>
<td>0.15</td>
</tr>
<tr>
<td>Density Parameter (TRL/LRL)</td>
<td>0.12</td>
</tr>
<tr>
<td>Root Growth Rate (RGR)</td>
<td>0.17</td>
</tr>
</tbody>
</table>
Results – OSU/ORNL GWAS

• Over 70 SNP loci were associated with one or more traits at non-conservative P-values to avoid Type II error (minimize false negatives)
  – Chose SNPs that were above a –log10P (LOD) score of 5-6 with at least two traits

• SNPs that passed this test were investigated further using Phytozome, *Populus trichocarpa* v 3.0 reference genome

• Several nearby genes with potential roles in rooting were identified
Examples of GWAS hits and nearby genes – OSU/ORNL (rooting score, water)

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Position</th>
<th>-log10 P-value</th>
<th>Gene</th>
<th>Notes/Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>27299510</td>
<td>7.20</td>
<td>Potri.001G264500.1</td>
<td>SNP in gene, UBX domain, Arabidopsis homolog PUX1 gene, loss of which is known for accelerated growth of roots</td>
</tr>
<tr>
<td>10</td>
<td>10735457</td>
<td>5.94</td>
<td>Potri.010G081300.1</td>
<td>Protein Argonaute 10, known for small RNA-directed gene slicing, expressed in roots</td>
</tr>
<tr>
<td>10</td>
<td>10808758</td>
<td>6.71</td>
<td>Potri.010G081900.1</td>
<td>Expression in roots, Arabidopsis homolog dyggve-melchior-clausen syndrome protein</td>
</tr>
<tr>
<td>13</td>
<td>3730879</td>
<td>6.09</td>
<td>Potri.013G051100.2</td>
<td>SNPs very close, both hit in gene. Zinc finger FYVE domain containing protein</td>
</tr>
<tr>
<td>13</td>
<td>3730867</td>
<td>5.61</td>
<td>Potri.013G051100.2</td>
<td></td>
</tr>
</tbody>
</table>
Examples of GWAS hits and nearby genes – OSU/ORNL (PC2, water, rooting vs. stem size)

<table>
<thead>
<tr>
<th>Rank</th>
<th>SNP ID</th>
<th>p-value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>20611278</td>
<td>6.02</td>
<td>Potri.003G205000.1: No annotated domain but high root expression</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Potri.003G205100.1: Arabidopsis homolog ZTP29, involved in salt stress response and can be expressed in root tissue and mutations of gene affect widening of cells in the root elongation zone</td>
</tr>
</tbody>
</table>
Results – WVU GWAS

• Goal of studying sources of statistical bias and minimizing Type I error (false positives)
• Ran GEMMA controlling for population structure; setting significance threshold with Bonferroni correction
• Also examined QQ plots as indication of normality and possible deviations from model assumptions
  – Binary traits that are very unbalanced (i.e., uneven counts of binary values) are especially susceptible to this effect
  – Continuous traits with severe departures from normal distribution can also be problematic
• Created permutation test to avoid statistical issues and provide much more conservative test
• Also ran GEMMA not controlling for population structure, and subsequent permutation test
  – Population structure is also confounded with adaptation and natural selection
Many significant associations with selected traits

Total of 244 genes pass Bonferroni criterion for lateral rooting aptitude
A similar result with average root length

~30 genes passed Bonferroni criterion
Badly behaved traits – binary traits with imbalanced distribution of effects

Simulated trait with a binomial distribution: $p=0.05$; $N=545$

Simulated trait with a binomial distribution: $p=0.50$; $N=545$
Badly behaved traits – Continuous traits with non-normal distribution

BLUPs from average root length

Simulated normal distribution
Lateral rooting aptitude is binary and highly imbalanced – unreliable
QQ plot with randomizations

QQ plot for phenotype AdvApt and covF with 1000 randomizations

- Randomized phenotype (grey dots)
- Top 5 percentile of randomizations –log P-vals
- Top 10 percentile of randomizations –log P-vals
- Actual phenotype (red dots)
- Mean of randomizations –log P-vals (solid)
A few associations for some traits remain significant after permutation analysis (root number)

<table>
<thead>
<tr>
<th>Chr</th>
<th>Marker</th>
<th>Position</th>
<th>AF</th>
<th>Pval</th>
<th>Distance</th>
<th>Gene_ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chr18</td>
<td>Chr18_10186245</td>
<td>10186245</td>
<td>0.990</td>
<td><strong>1.910189e-14</strong></td>
<td>1703</td>
<td>Potri.018G076400</td>
</tr>
</tbody>
</table>
Genome browser view

Significant SNP position

ARF-GAP GTPase

copper-binding family protein

VIN3-like protein
VIN3-like protein is considered part of root gene expression network

- Similar to VIN3 (vernalization insensitive)
- Plant homeodomain finger, chromatin remodeling protein

**Gene** Potri.018G076500

**Gene Info**

- **Organism**: Populus trichocarpa
- **Transcript Name**: Potri.018G076500.1 (primary)
- **Other transcripts**: Potri.018G076500.3 Potri.018G076500.2 Potri.018G076500.4
- **Location**: Chr18:10212068..10217311 forward
- **Alias**: POPTR_0018s06200 POPTR_0018s06200.v2.2
- **Description**: (1 of 3) PTHR21736.SF17 - VIN3-LIKE PROTEIN 1
- **Gene Atlas Desc**: Coexpressed with genes in roots specific coexpression subnetwork
- **Links**: B M
ARF-GAP GTPase is root expressed, with a zinc finger domain

<table>
<thead>
<tr>
<th>ID</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTHR23180</td>
<td>PANTHER</td>
<td>CENTAURIN/ARF</td>
</tr>
<tr>
<td>PTHR23180_S67200</td>
<td>PANTHER</td>
<td>ARF-GAP DOMAIN AND FG REPEAT-CONTAINING PROTEIN 1</td>
</tr>
<tr>
<td>PF01412</td>
<td>PFAM</td>
<td>Putative GTPase activating protein for Arf</td>
</tr>
<tr>
<td>KOG0702</td>
<td>KEGG</td>
<td>Predicted GTPase-activating protein</td>
</tr>
<tr>
<td>GO:00050996</td>
<td>GO</td>
<td>Binds to and increases the activity of a GTPase, an enzyme that catalyzes the hydrolysis of GTP</td>
</tr>
</tbody>
</table>
Some relevant observations about the gene and its homologs

“Arf GTPase activating protein… a small family… important for the regulation of the ADP ribosylation factor ARF… essential for the maintenance of normal Golgi morphology… required for budding and fission of membranes….

Closest Arabidopsis homolog is strongly expressed in root tips…

Arf GTPase activating proteins are involved with many root functions, membrane curvature of vesicles, including directionality and root hair development

It also interacts with phospholipase D, which is a plasma membrane signaling enzyme strongly involved in root development, and antagonists of it are known to stunt root growth in Arabidopsis
Key messages

• Though strongly affected by heredity, adventitious rooting is a complex trait with weak heritability
  – Many rooting traits, difficult to phenotype
  – Responsive to variation in environment and physiology of cuttings (variation in size, age, bud proximity, shoot flushing)

• Numerous associations of SNPs with rooting traits discovered, some with plausible physiological roles

• However, statistics of associations are complex – much potential for Type I and Type II errors
Key messages

• Genome biology to identify the best candidates is complex!

• More precise phenotypes, larger sample sizes, independent GWAS results, in depth genome analysis, and physiological validations are needed for confirmation and biological interpretation.
Thanks to coauthors for lots of help

• West Virginia University
  – Christine Zawaski – rooting data
  – David Macaya-Sanz – GWAS bioinformatics
  – Stephen P. DiFazio – GWAS lead
  – Jonathan R. Cumming – root physiology lead

• Oregon State University
  – Cathleen Ma – rooting data
  – Anna C. Magnuson – Phytozome analysis
  – Jialin Yuan – Machine vision
  – Yuan Jiang – Statistics
  – Fuxin Li – Machine vision

• Oak Ridge National Laboratory
  – Wellington Muchero – GWAS bioinformatics