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

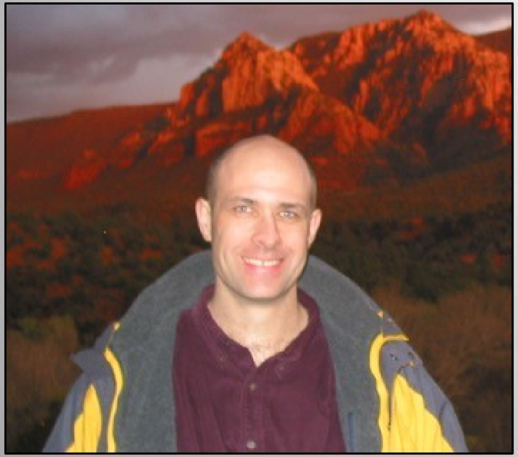
Steve Strauss

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

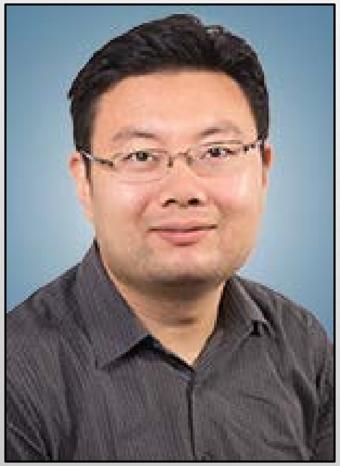
David Macaya Sanz

Jonathan Cumming

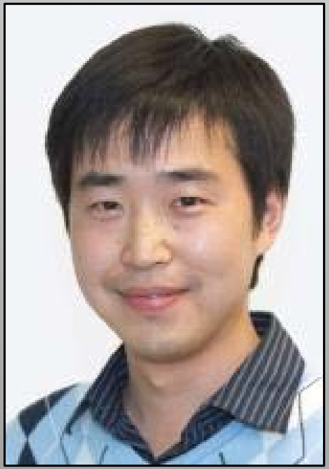
Christine Zawaski

Coauthors – OSU and ORNL

Wellington
Muchero,
ORNL



Fuxin Li



Yuan Jiang

Cathleen Ma

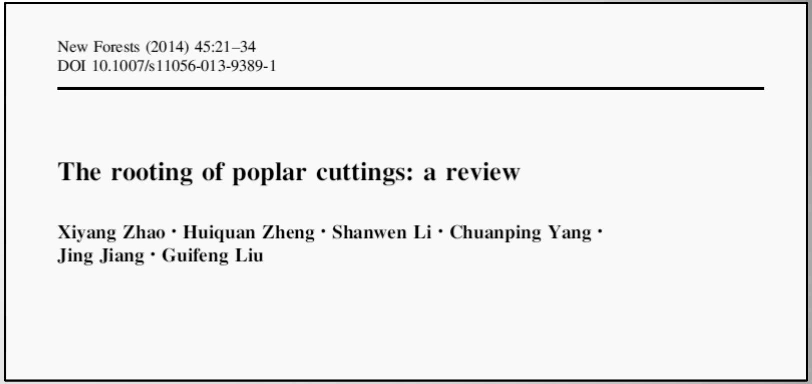
Anna Magnuson

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



New Forests (2014) 45:21–34
DOI 10.1007/s11056-013-9389-1

The rooting of poplar cuttings: a review

Xiyang Zhao · Huiquan Zheng · Shanwen Li · Chuanping Yang ·
Jing Jiang · Guifeng Liu

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

Wikipedia

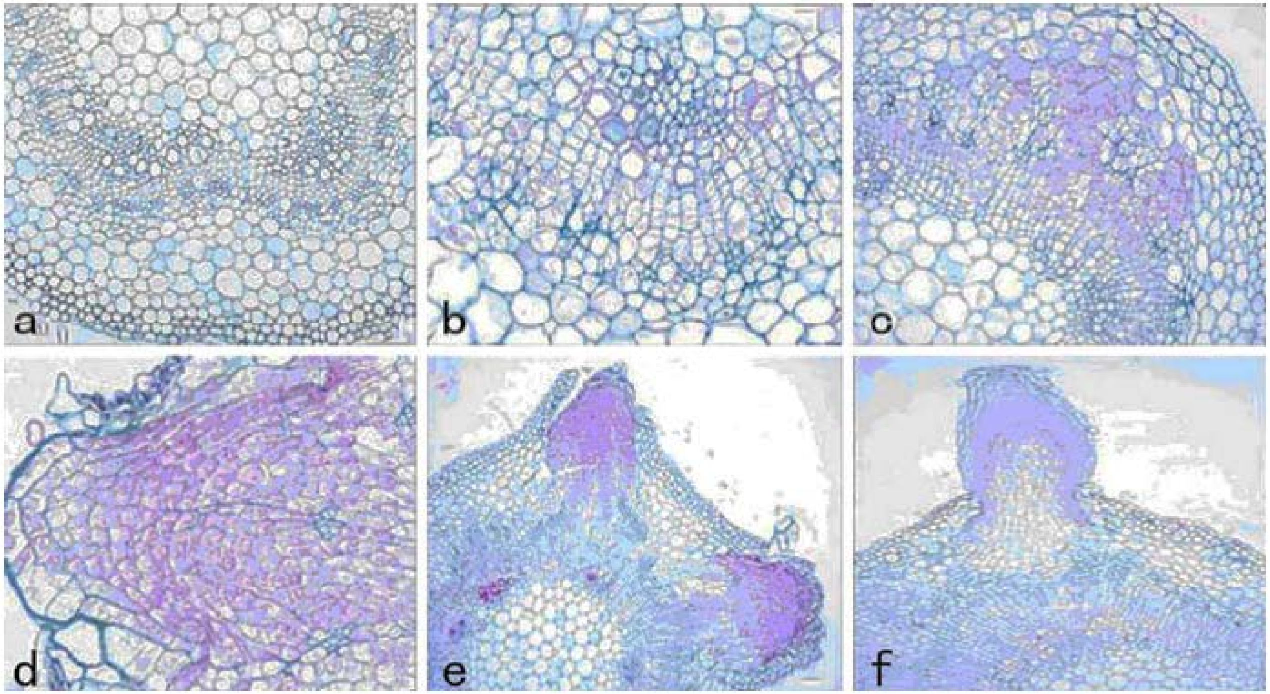


Fig. 1 Histology of adventitious root development in poplar. Ten micrometer transverse sections (**a**–**f**) and longitudinal sections were stained with fast *green* and *safranin red*. No new cell divisions were observed on Day 1 (**a**). Cell divisions were first observed on Day 2 (**b**) in the vascular cambium and surrounding tissue. Localized division occurred on Day 3 (**c**), and meristem organization occurred on Day 4 (**d**). Root primordia began to emerge on Day 5 (**e**), and root elongation and root hair formation occurred on Day 6 (**f**). (Color figure online)

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

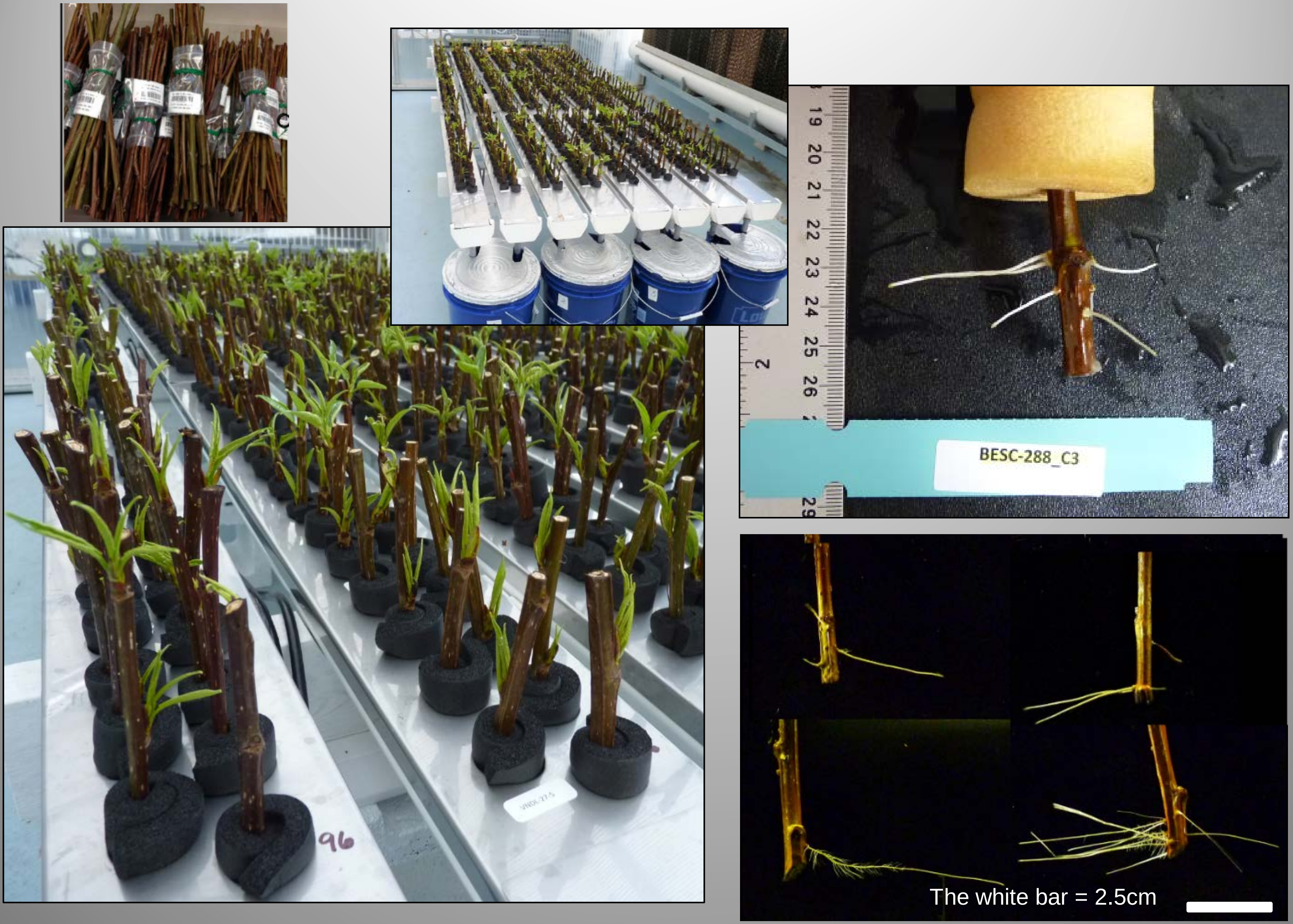
Materials and measurements

- 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

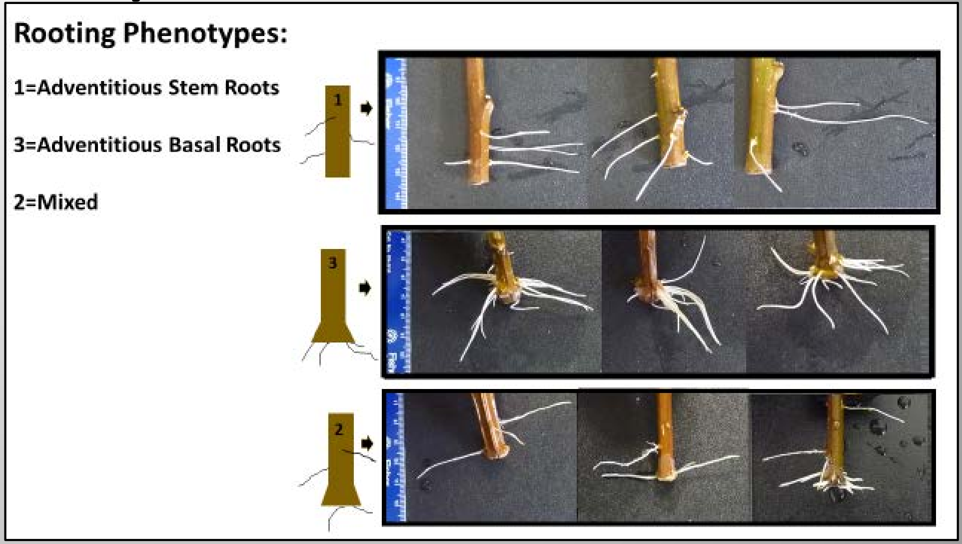
Machine vision analysis – OSU (n=423)

WVU measurement system



The white bar = 2.5cm

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

	PC1	PC2	PC3	PC4
Height	0.44	**-0.43**	0.77	0.15
Diameter	0.65	**0.19**	-0.13	-0.73
Ave. Stem Diameter	0.61	**0.25**	-0.34	0.66
Root Area	0.10	**-0.85**	-0.52	-0.04

Numerous rooting traits derived

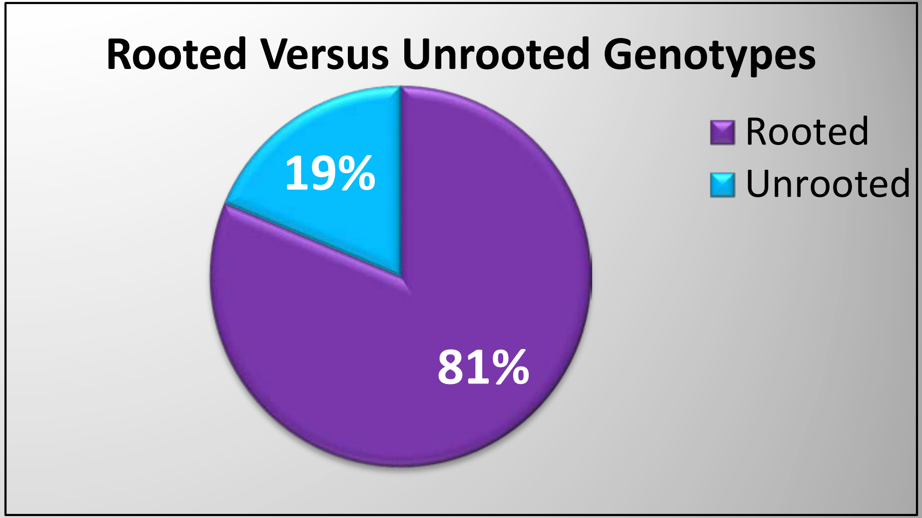
Traits	***Subset of tested individuals***	***Type of data***	***Processing of raw data***
Days to root initiation (WVU)	all	Integer	BLUPs
Root initiation score (WVU)	all	Discrete (score)	BLUPs
Longest root length (WVU)	all	Continuous	BLUPs
Total root length (WVU)	all	Continuous	BLUPs
Root density (WVU)	all	Continuous	BLUPs
Root growth rate (WVU)	all	Continuous	BLUPs
Average root length (WVU)	all	Continuous	BLUPs
Root number (WVU)	all	Integer	BLUPs
Mode of rooting score (WVU)	all	Discrete (score)	BLUPs
Basal rooting aptitude (WVU)	all	Binary	Raw data
Lateral rooting aptitude (WVU)	all	Binary	Raw data
Longest root length (OSU)	all	Continuous	BLUPs
Root score (OSU)	all	Discrete (score)	BLUPs
Root area (OSU)	all	Continuous	BLUPs
Root score in soil (OSU)	all	Discrete (score)	BLUPs
Longest root length (WVU)	only lateral-rooting genotypes	Continuous	BLUPs
Total root length (WVU)	only lateral-rooting genotypes	Continuous	BLUPs
Root growth rate (WVU)	only lateral-rooting genotypes	Continuous	BLUPs
Root density (WVU)	only lateral-rooting genotypes	Continuous	BLUPs
Average root length (WVU)	only lateral-rooting genotypes	Continuous	BLUPs
Longest root length (WVU)	only basal-rooting genotypes	Continuous	BLUPs
Total root length (WVU)	only basal-rooting genotypes	Continuous	BLUPs
Root growth rate (WVU)	only basal-rooting genotypes	Continuous	BLUPs
Root density (WVU)	only basal-rooting genotypes	Continuous	BLUPs
Average root length (WVU)	only basal-rooting genotypes	Continuous	BLUPs

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

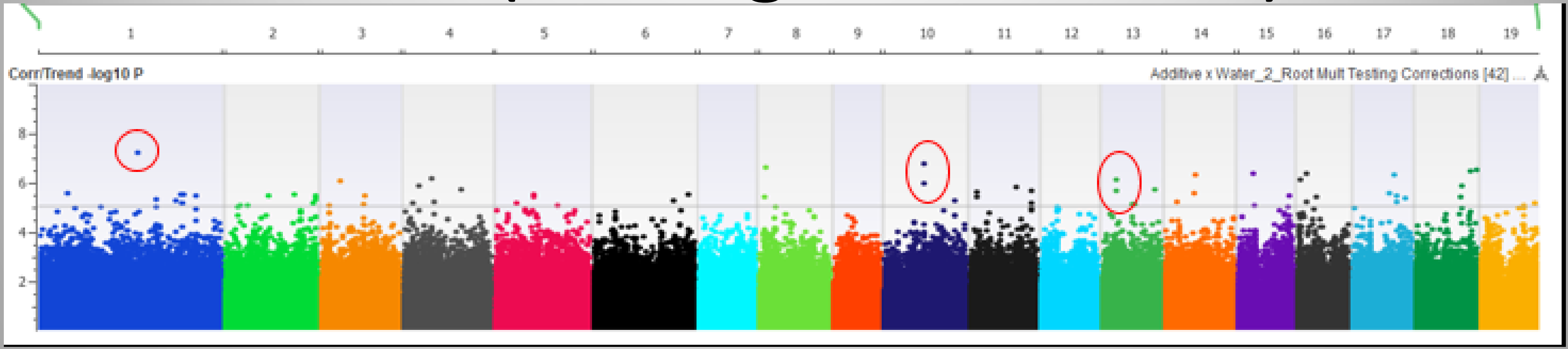
OSU	h^2
Root number score - Combined	0.13
Water	0.18
Soil	0.15

WVU	h^2
Days to Root Initiation (DRI)	0.25
Longest Root Length (LRL)	0.17
Total Root Length (TRL)	0.15
Density Parameter (TRL/LRL)	0.12
Root Growth Rate (RGR)	0.17

Results – OSU/ORNL GWAS

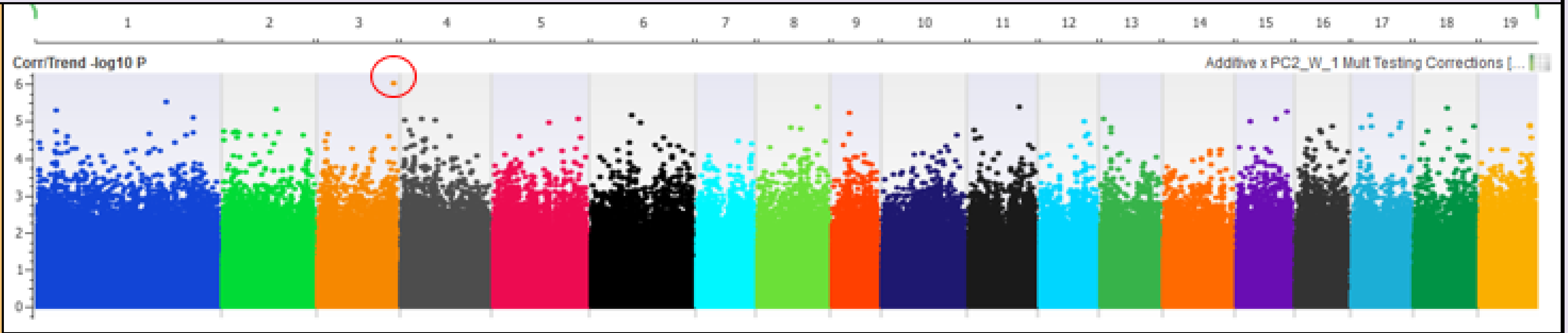
- Over 70 SNP loci were associated with one or more traits at non-conservative P-values to avoid Type II error (minimize false negatives)
  - Chose SNPs that were above a –log10P (LOD) score of 5-6 with at least two traits
- SNPs that passed this test were investigated further using Phytozome, *Populus 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)



Chromosome	Position	-log10 P-value	Gene	Notes/Description
1	27299510	7.20	Potri.001G264500.1	SNP in gene, UBX domain, Arabidopsis homolog PUX1 gene, loss of which is known for accelerated growth of roots
10	10735457	5.94	Potri.010G081300.1	Protein Argonaute 10, known for small RNA-directed gene slicing, expressed in roots
10	10808758	6.71	Potri.010G081900.1	Expression in roots, Arabidopsis homolog dyggve-melchior-clausen syndrome protein
13	3730879	6.09	Potri.013G051100.2	SNPs very close, both hit in gene. Zinc finger FYVE domain containing protein, high expression in root tips
13	3730867	5.61		

Examples of GWAS hits and nearby genes – OSU/ORNL (PC2, water, rooting vs. stem size)

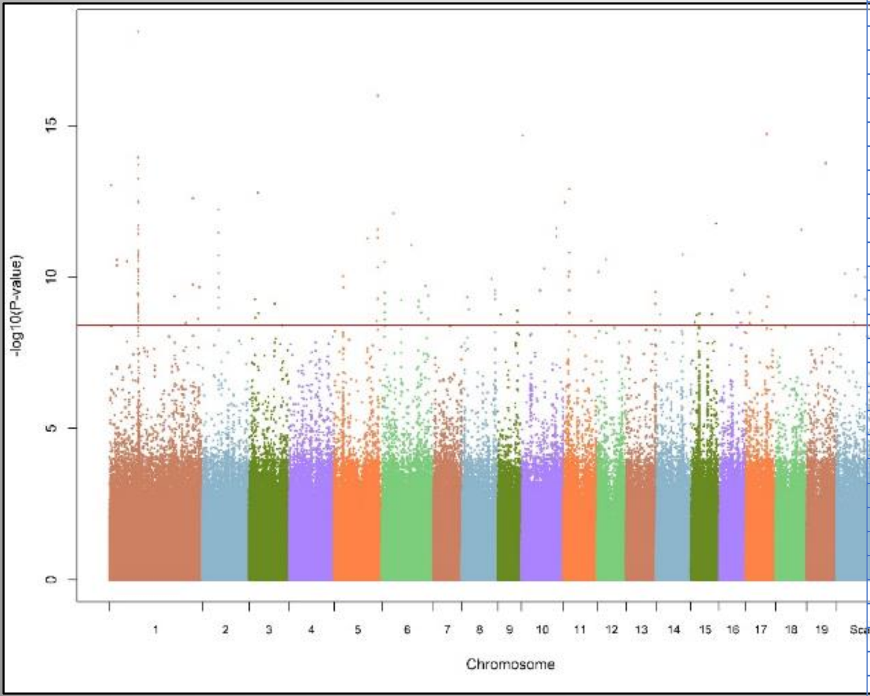


3	20611278	6.02	Potri.003G205000.1	No annotated domain but high root expression
			Potri.003G205100.1	Arabidopsis homolog ZTP29, involved in salt stress response and can be expressed in root tissue and mutations of gene affect widening of cells in the root elongation zone

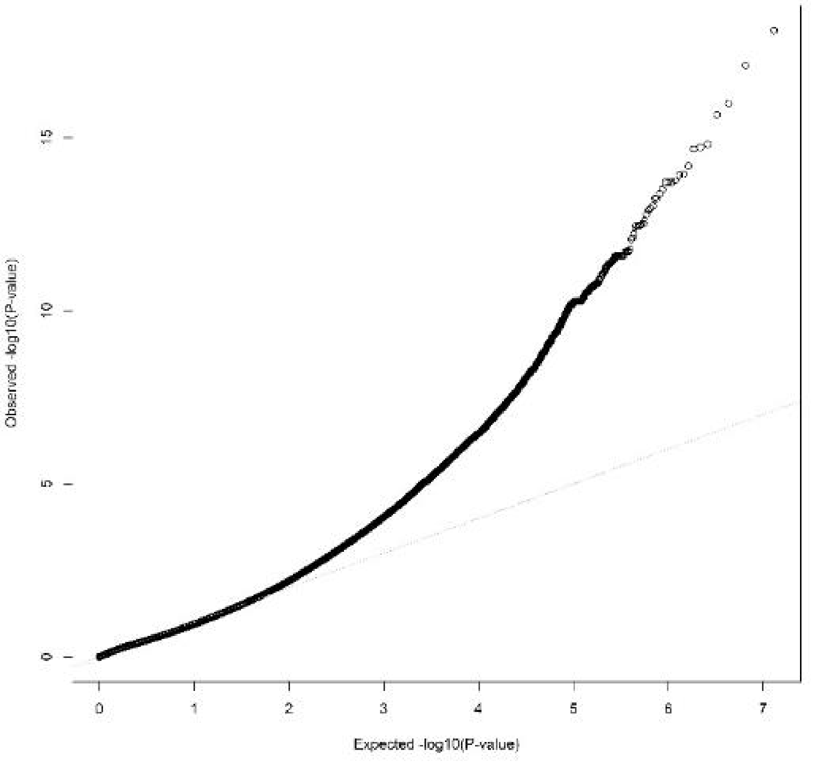
Results – WVU GWAS

- 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 <u>not</u> controlling for population structure, and subsequent permutation test
  - Population structure is also confounded with adaptation and natural selection

Many significant associations with selected traits

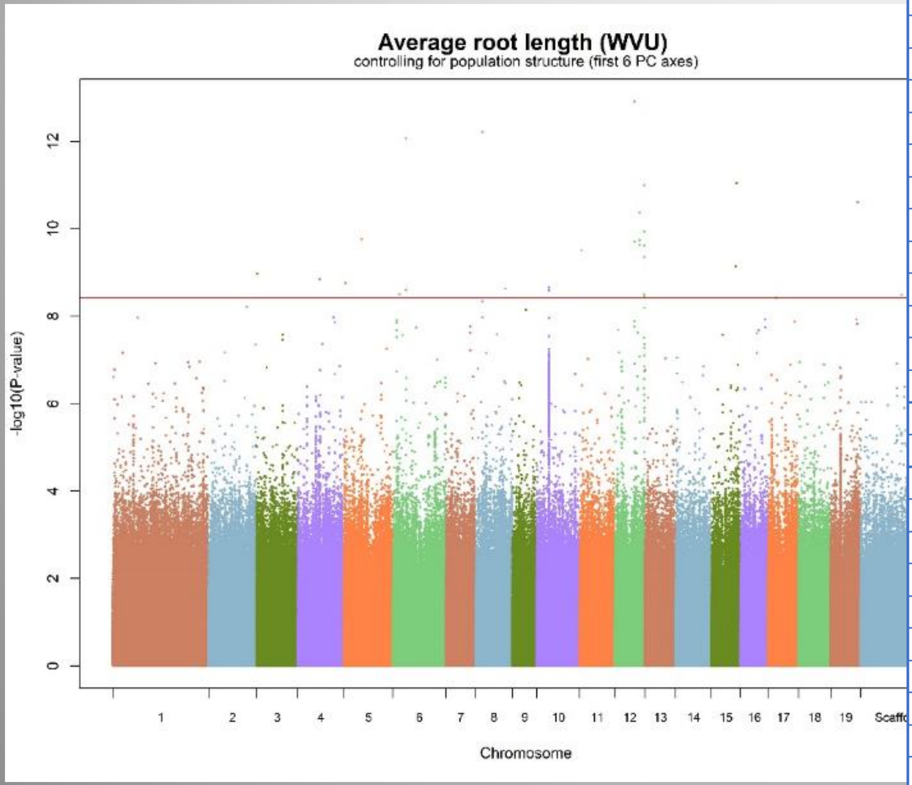


Gene	MinPval	Hits	ShortDef
Potri.001G180800	8.02E-19	9	Unknown Function
Potri.005G229900	8.15E-18	3	cytochrome P450, family 722, subfamily A, polypeptide 1
Potri.005G230000	1.04E-16	1	cytochrome P450, family 722, subfamily A, polypeptide 1
Potri.001G180600	2.19E-16	68	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family
Potri.017G096100	1.54E-15	3	Malectin/receptor-like protein kinase family protein
Potri.010G007500	2.15E-15	2	Kunitz family trypsin and protease inhibitor protein
Potri.005G230200	6.63E-15	1	inositol-pentakisphosphate 2-kinase 1
Potri.001G180000	1.13E-14	2	Pentatricopeptide repeat (PPR-like) superfamily protein
Potri.019G069900	1.73E-14	1	disease resistance protein (TIR-NBS-LRR class), putative
Potri.002G117500	3.14E-14	5	Transducin family protein / WD-40 repeat family protein
Potri.001G014000	4.26E-14	2	lipoyltransferase 2
Potri.011G040000	1.24E-13	9	phosphoinositide binding
Potri.003G037200	1.63E-13	2	Unknown Function
Potri.001G426800	2.51E-13	2	Unknown Function
Potri.001G181000	3.43E-13	8	F-box family protein
Potri.011G012900	3.48E-13	3	transmembrane receptors;ATP binding
Potri.002G117400	5.98E-13	2	endoplasmic reticulum-type calcium-transporting ATPase 3
Potri.006G084000	7.88E-13	1	ARF-GAP domain 6
Potri.T126700	8.87E-13	1	wall associated kinase-like 1
Potri.015G121300	1.70E-12	1	Unknown Function
Potri.010G197900	2.46E-12	3	Exocytosis family protein
Potri.001G180900			
Potri.005G229500			
Potri.018G117600			
Potri.002G117200			
Potri.007G018600			
Potri.006G228100			
Potri.008G218300			
Potri.019G046600			
Potri.005G169800			
Potri.013G156000			
Potri.009G051100			
Potri.005G071000			
Potri.006G162900			
Potri.017G011800			
Potri.008G218200			
Potri.005G229800			
Potri.008G051500			
Potri.014G176200			
Potri.001G054800			
Potri.012G050600			
Potri.006G027000			
NA			
Potri.001G118900			
Potri.T066000			
Potri.013G091800			
Potri.010G101300			
Potri.T045300			
Potri.T096200			
Potri.012G095400			

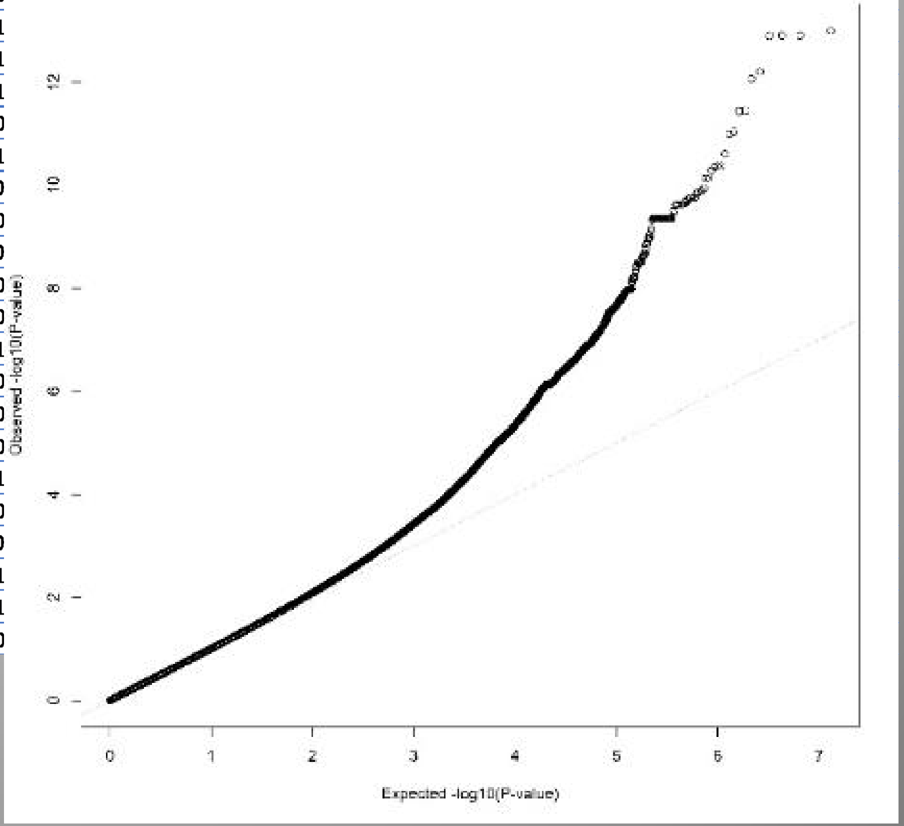


Total of 244 genes pass Bonferroni criterion for lateral rooting aptitude

A similar result with average root length



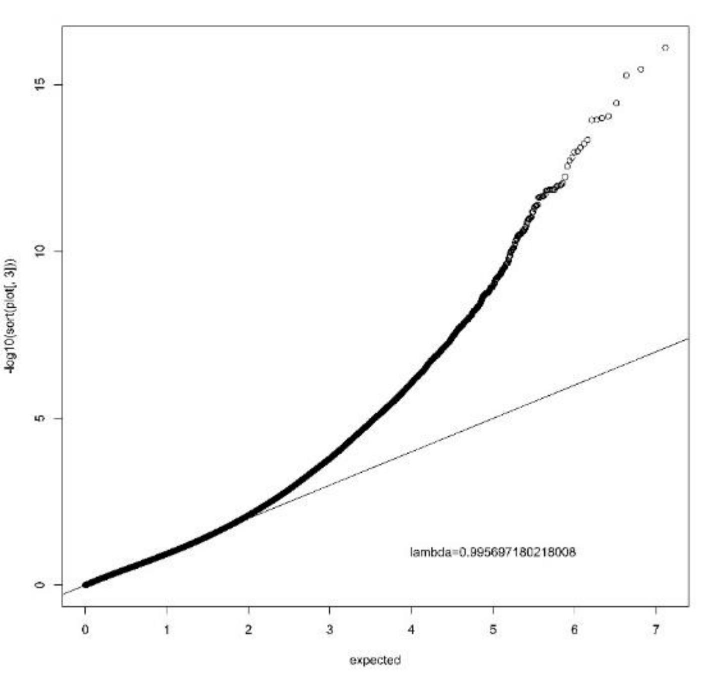
Gene	MinPval	Hits	ShortDef
Potri.012G142500	1.00E-13	26	Unknown Function
Potri.012G077100	1.24E-13	5	Unknown Function
Potri.008G061600	6.20E-13	2	Unknown Function
Potri.006G091000	8.60E-13	1	uracil phosphoribosyltransferase
NA	3.60E-12	2	Unknown Function
Potri.018G011200	3.67E-12	1	ubiquitin-specific protease 19
Potri.015G116100	9.05E-12	1	Protein of unknown function (DUF594)
Potri.019G115600	2.47E-11	1	Heavy metal transport/detoxification superfamily protein
Potri.012G106600	4.27E-11	10	magnesium-protoporphyrin IX methyltransferase
Potri.007G103500	4.42E-11	1	Unknown Function
Potri.015G123300	7.26E-11	1	Ribosomal protein L1p/L10e family
Potri.015G077300	7.28E-11	1	pantothenate kinase 2
Potri.012G077000	1.37E-10	1	Calmodulin binding protein-like
Potri.005G120800	1.73E-10	2	gamma carbonic anhydrase 1
Potri.01			
Potri.01			
Potri.01			
Potri.00			
Potri.01			
Potri.T0			
Potri.00			
Potri.T0			
Potri.00			
Potri.00			
Potri.01			
Potri.00			
Potri.00			
Potri.00			
Potri.01			
Potri.00			
Potri.T0			
Potri.01			
Potri.01			
Potri.00			



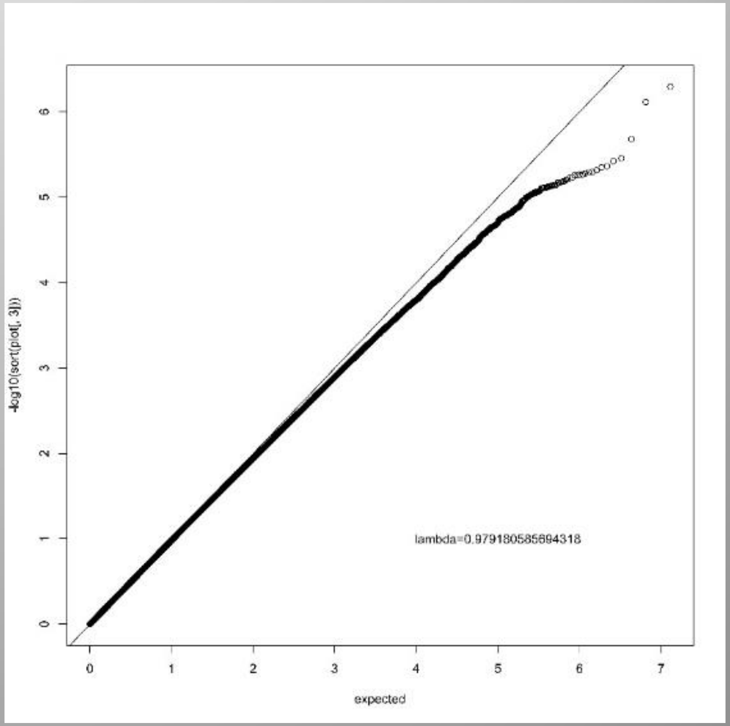
~30 genes passed Bonferroni criterion

Badly behaved traits – binary traits with imbalanced distribution of effects

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

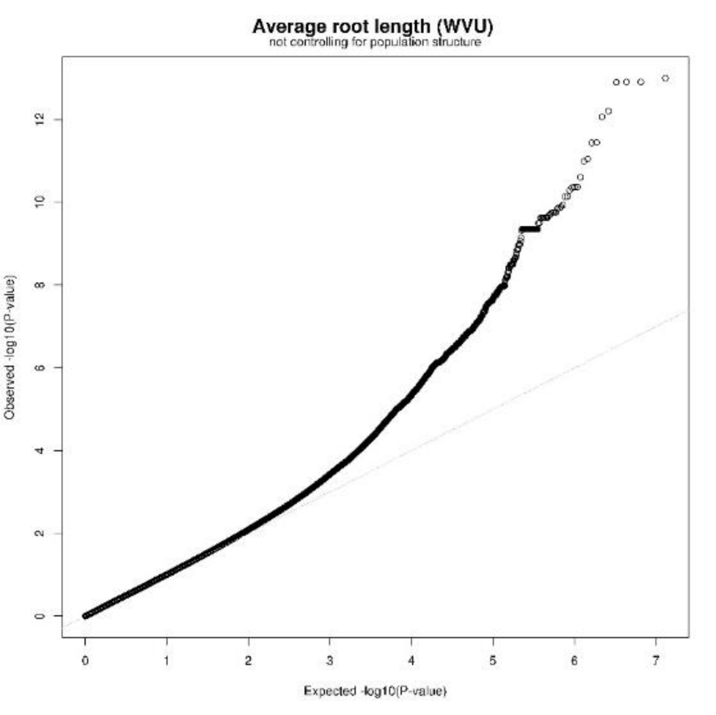


Simulated trait with a binomial distribution: p=0.50; N=545

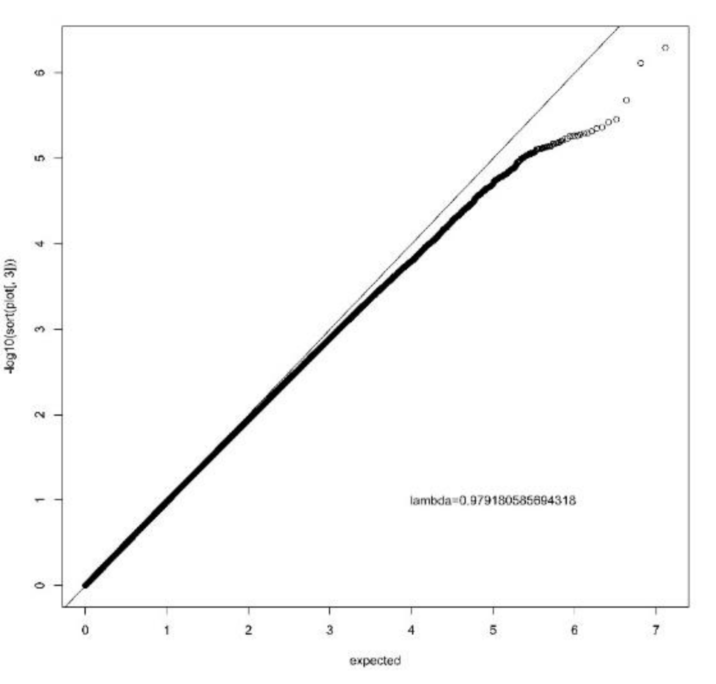


Badly behaved traits – Continuous traits with non-normal distribution

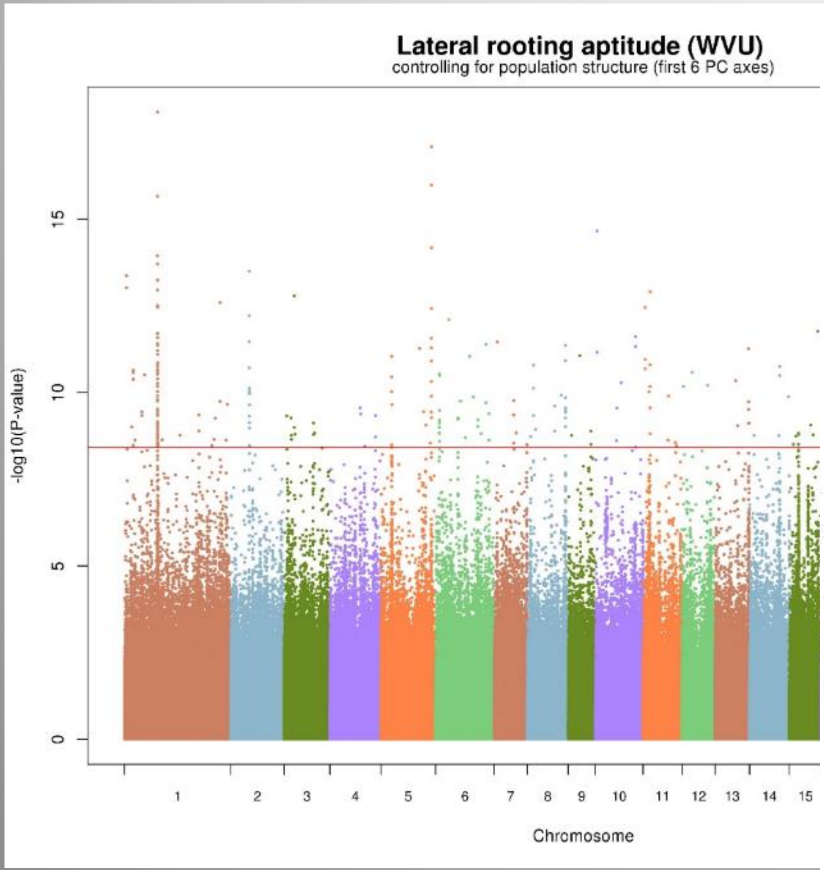
BLUPs from average root length

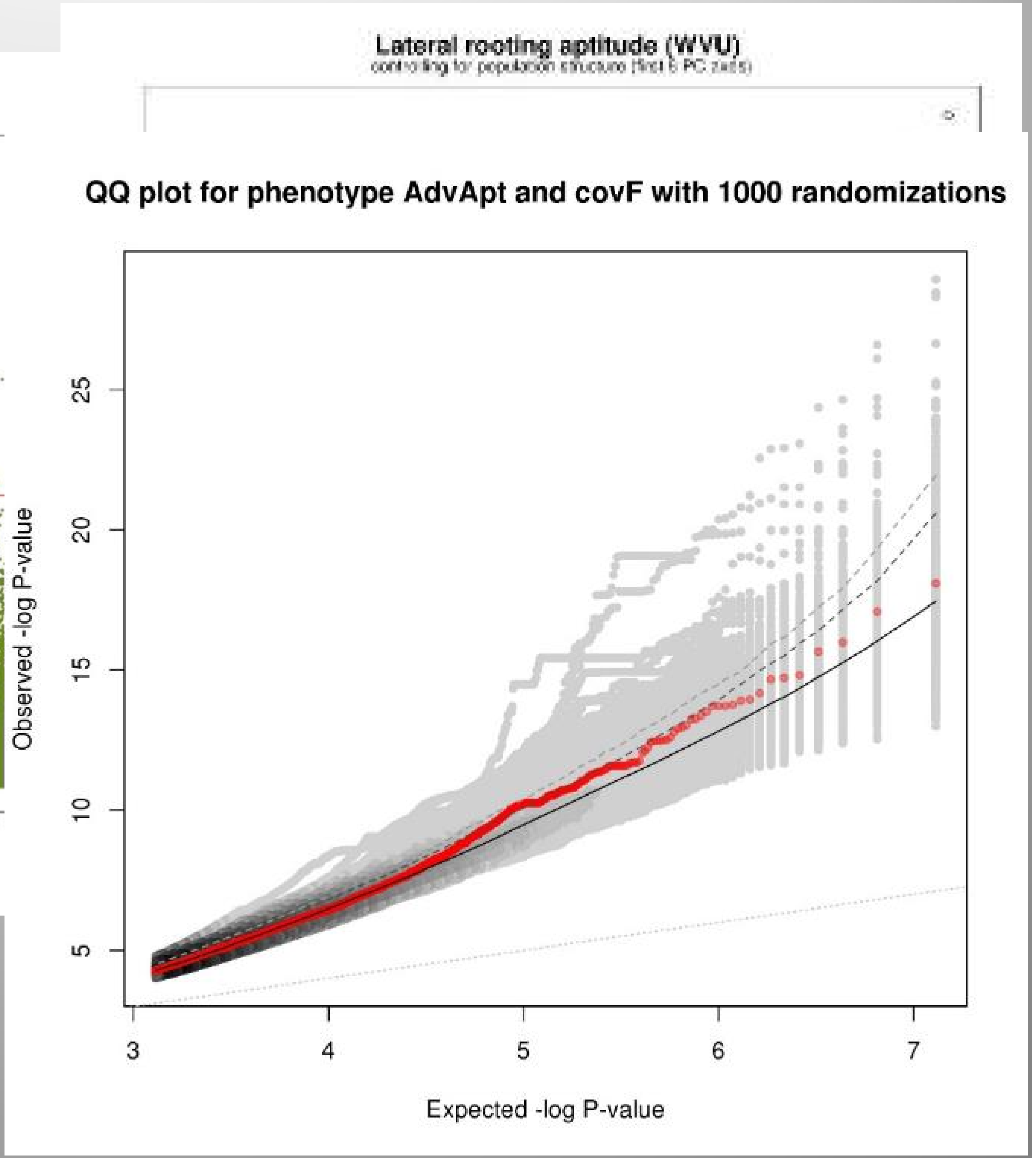


Simulated normal distribution

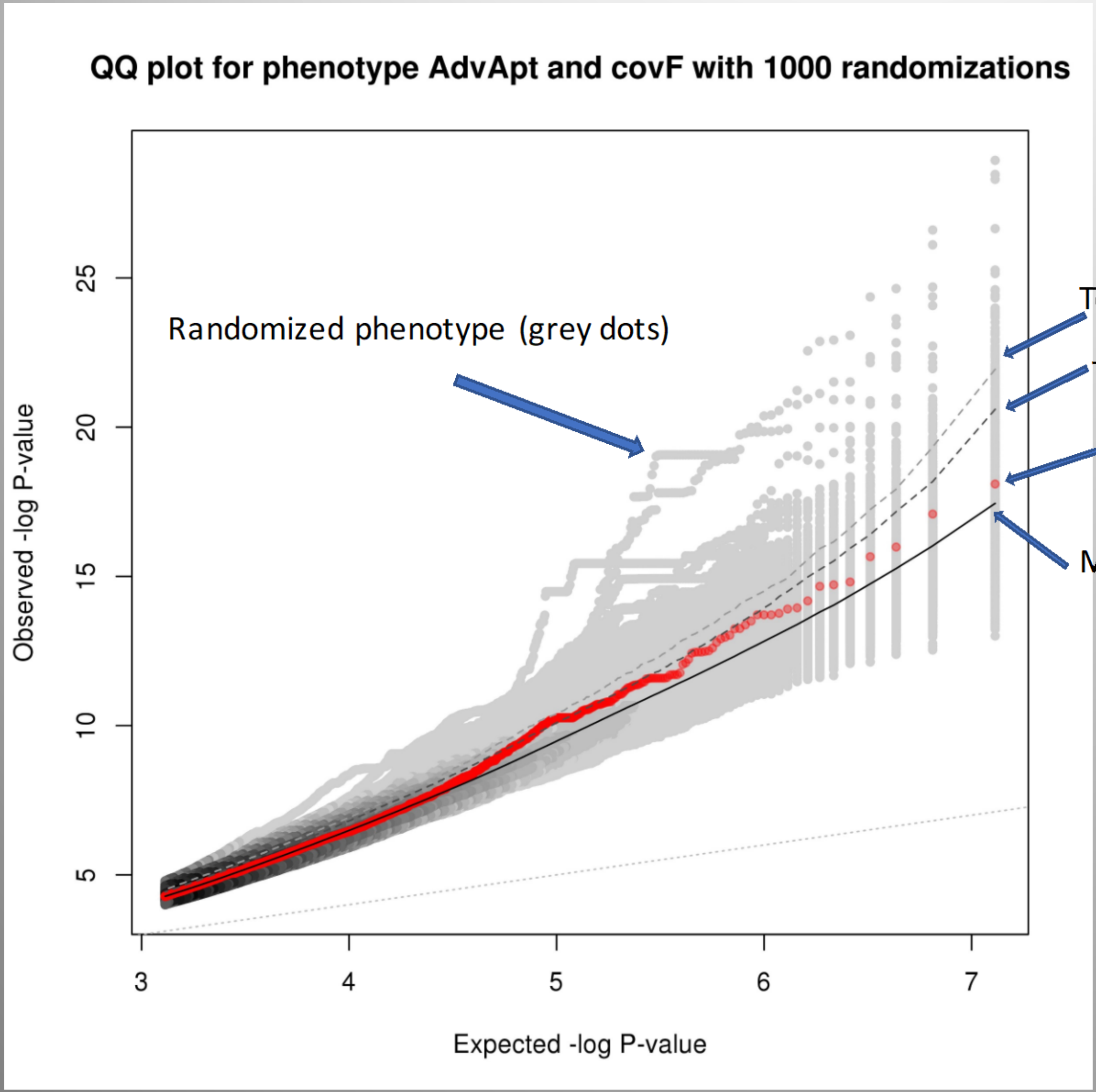


Lateral rooting aptitude is binary and highly imbalanced – unreliable





QQ plot with randomizations



Randomized phenotype (grey dots)

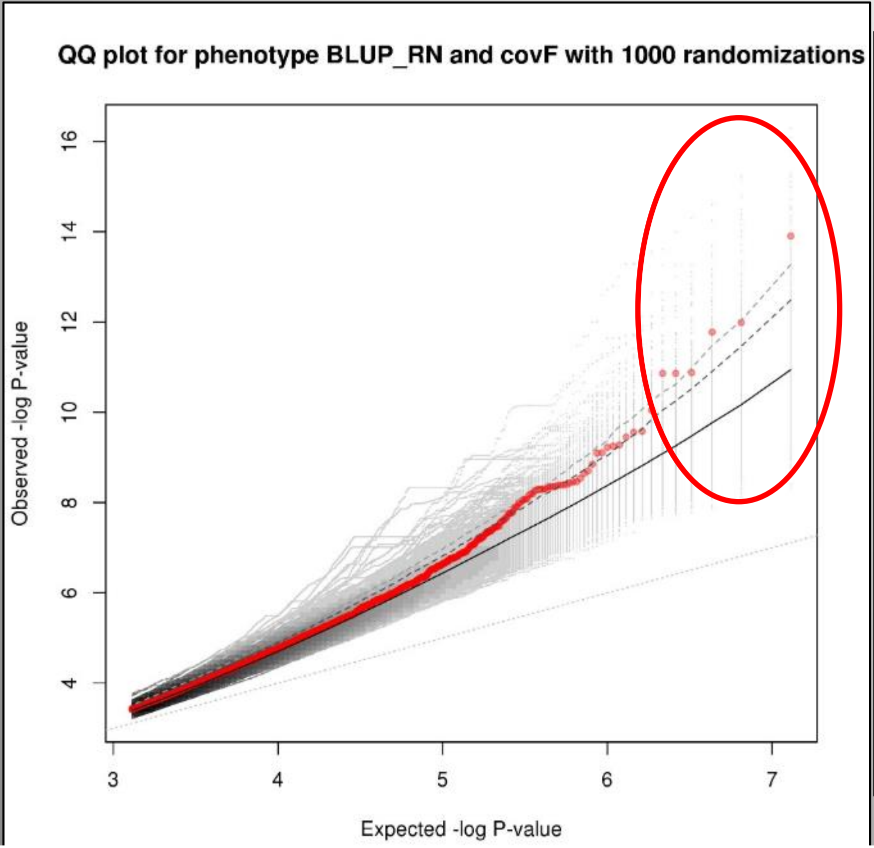
Top 5 percentile of randomizations –log P-vals

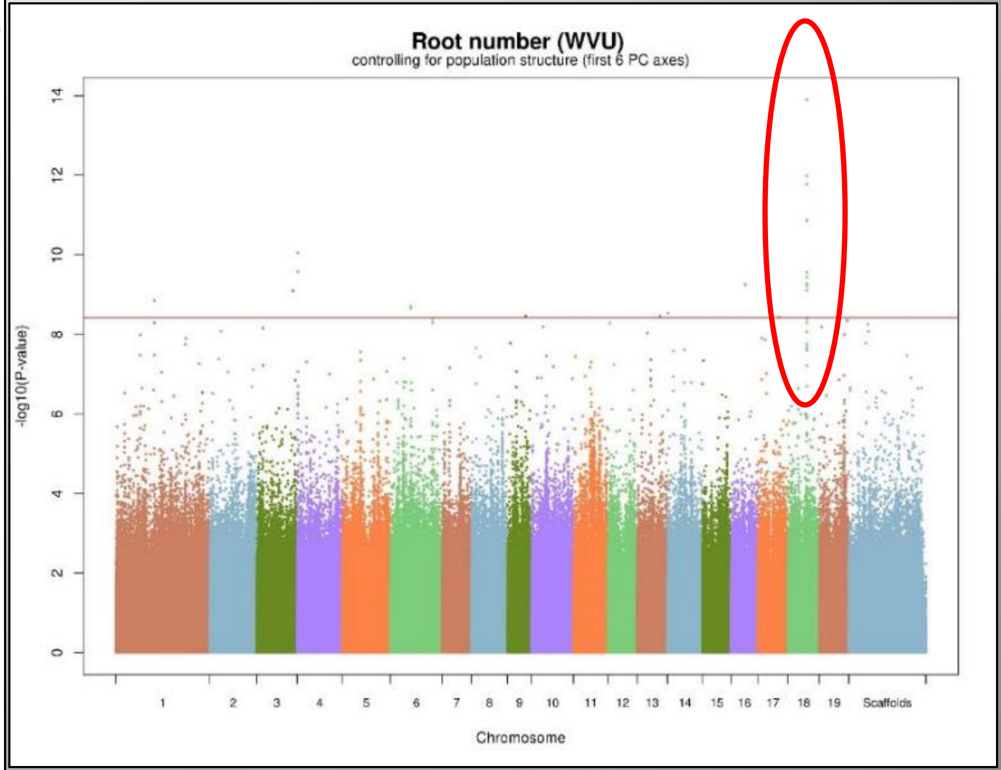
Top 10 percentile of randomizations –log P-vals

Actual phenotype (red dots)

Mean of randomizations –log P-vals (solid)

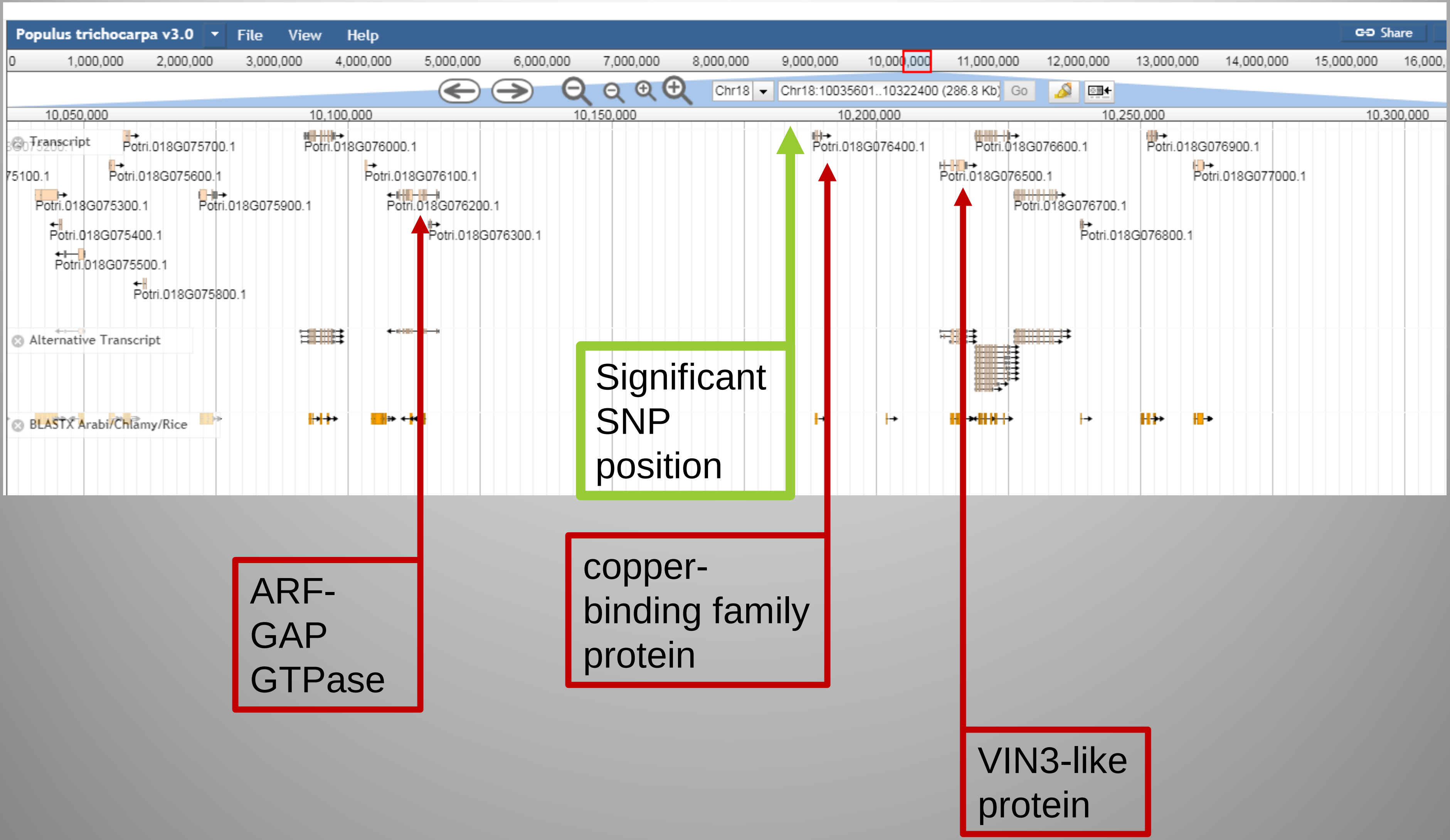
A few associations for some traits remain significant after permutation analysis (root number)





Chr	Marker	Position	AF	Pval	Distance	Gene_ID
Chr18	Chr18_10186245	10186245	0.990	**1.910189e-14**	1703	Potri.018G076400

Genome browser view



Significant SNP position

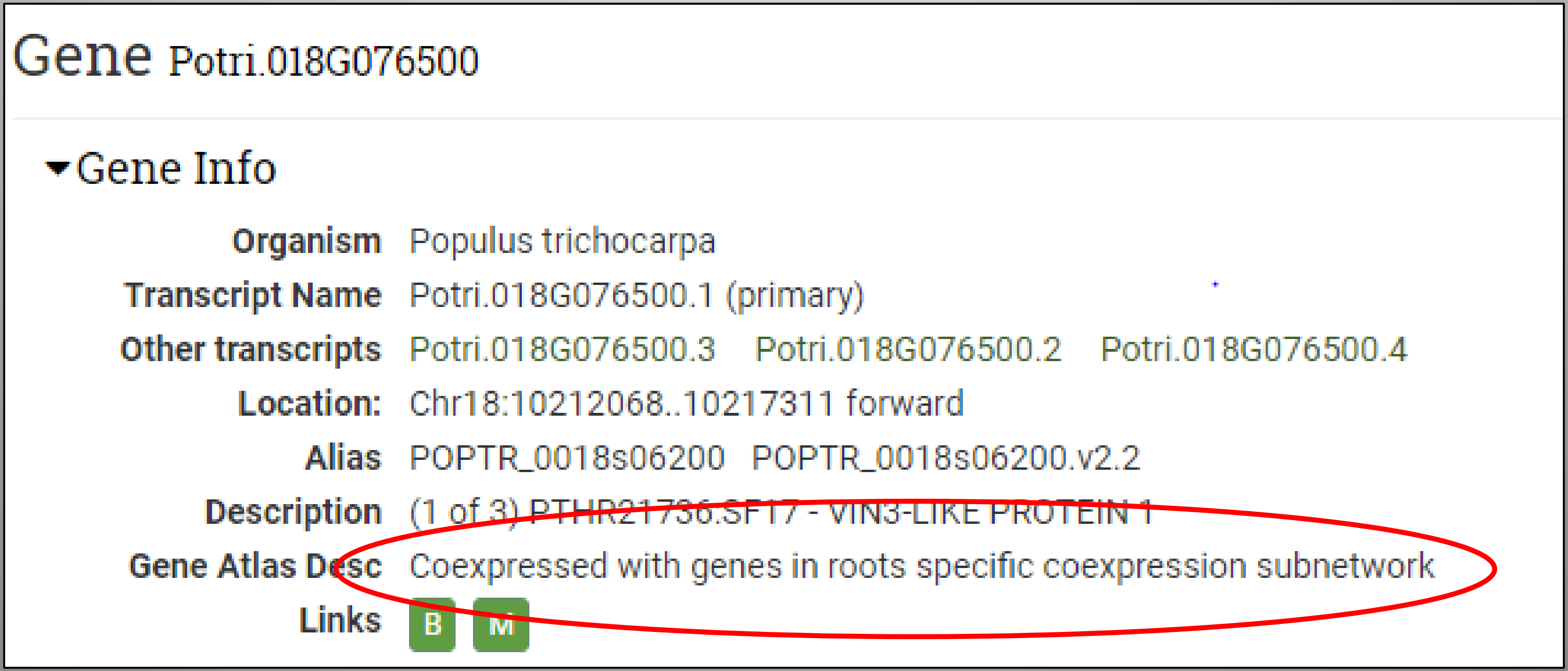
copper-binding family protein

ARF-GAP GTPase

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

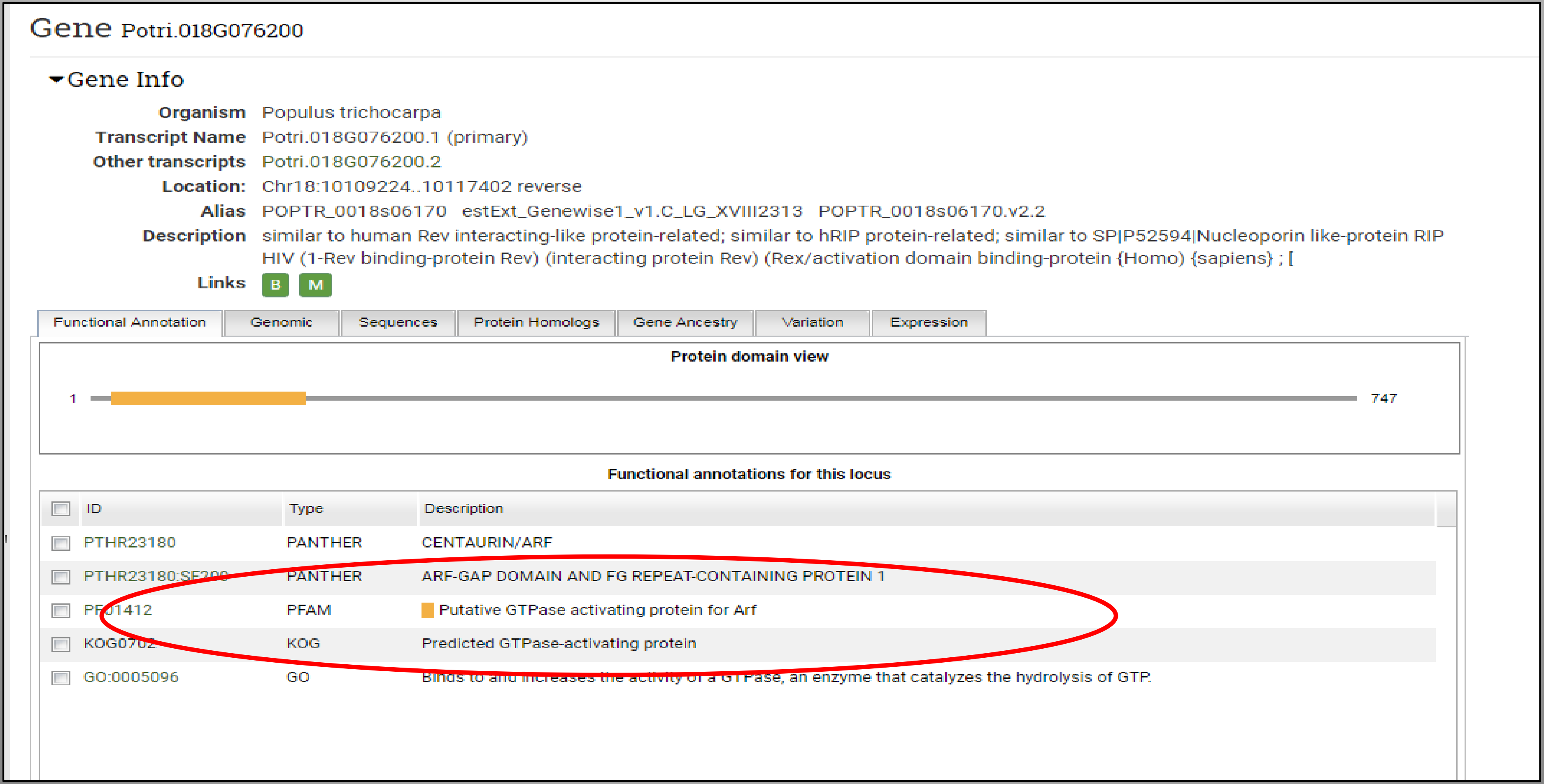


Gene Potri.018G076500

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



Gene Potri.018G076200

Gene Info

Organism	Populus trichocarpa
Transcript Name	Potri.018G076200.1 (primary)
Other transcripts	Potri.018G076200.2
Location:	Chr18:10109224..10117402 reverse
Alias	POPTR_0018s06170 estExt_Genewise1_v1.C_LG_XVIII2313 POPTR_0018s06170.v2.2
Description	similar to human Rev interacting-like protein-related; similar to hRIP protein-related; similar to SP|P52594|Nucleoporin like-protein RIP HIV (1-Rev binding-protein Rev) (interacting protein Rev) (Rex/activation domain binding-protein {Homo) {sapiens} ; [
Links	B M

Functional annotations for this locus

☐	ID	Type	Description
☐	PTHR23180	PANTHER	CENTAURIN/ARF
☐	PTHR23180:SF200	PANTHER	ARF-GAP DOMAIN AND FG REPEAT-CONTAINING PROTEIN 1
☐	PF01412	PFAM	🟧 Putative GTPase activating protein for Arf
☐	KOG0702	KOG	Predicted GTPase-activating protein
☐	GO:0005096	GO	Binds to and increases the activity of a GTPase, an enzyme that catalyzes the hydrolysis of GTP.

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

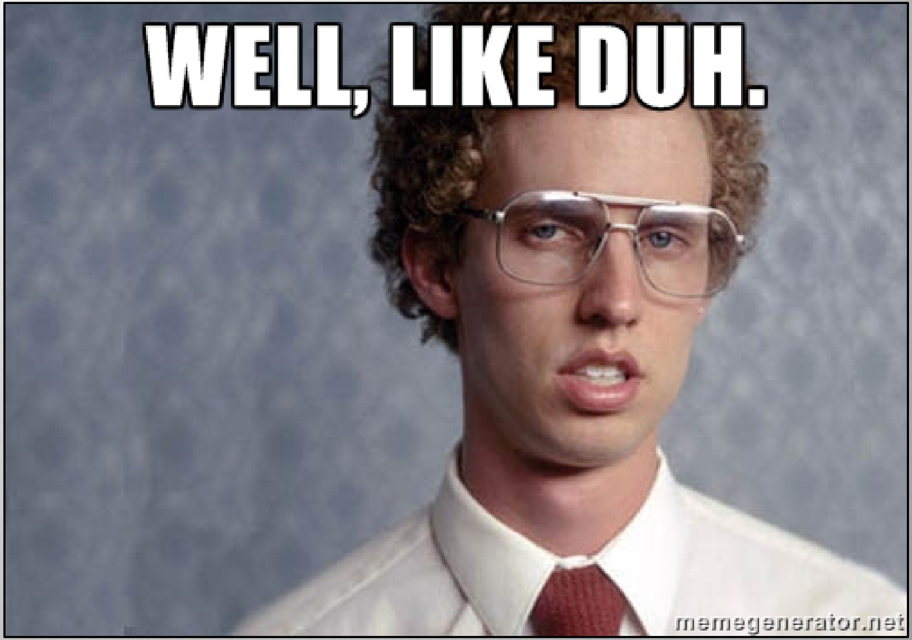
It also interacts with phosphlipase D, which is a plasma membrane signaling enzyme strongly involved in root development, and antagonists of it are known to stunt root growth in Arabidopsis

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