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

Michael F. Nagle[1#], Jialin Yuan[2], Damanpreet Kaur[2], Cathleen Ma[1], Ekaterina Peremyslova[1], Yuan Jiang[3], Christopher J. Willig[1], Greg S. Goralogia[1], Alexa Niño de Rivera[1], Megan McEldowney[1], Amanda Goddard[1], Anna Magnuson[1], Wellington Muchero[4], Li Fuxin[2], Steven H. Strauss[1@]

[1] Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR, USA

[2] Department of Electrical Engineering and Computer Science, Oregon State University, Corvallis, OR, USA

[3] Statistics Department, Oregon State University, Corvallis, OR, USA

[4] Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA

@ Correspondence to steve.strauss@oregonstate.edu

Co-correspondence to michael.nagle@oregonstate.edu

Abstract

Adventitious rooting is critical to the propagation, breeding, and genetic engineering or editing of trees. The capacity for plants to undergo these processes is highly heritable; however, the basis of its genetic variation is largely uncharacterized. To identify genetic regulators of these processes, we performed a genome-wide association study (GWAS) using 1,148 genotypes of *Populus trichocarpa*. GWAS are often limited by the abilities of researchers to collect precise phenotype data on a high-throughput scale; to help overcome this limitation, we developed a computer vision system to measure an array of traits related to adventitious root development in poplar, including temporal measures of lateral and basal root length and area. GWAS was performed using multiple methods and significance thresholds to handle non-normal phenotype statistics, and to gain statistical power. These analyses yielded a total of 277 unique associations, suggesting that genes that control rooting include regulators of hormone signaling, cell division and structure, and reactive oxygen species signaling. Genes related to other processes with known roles in root development, and numerous genes with uncharacterized functions and/or cryptic roles, were also identified. These candidates provide targets for functional analysis, including physiological and epistatic analyses, to better characterize the complex polygenic regulation of adventitious rooting.

Introduction

The species within the genus *Populus spp.* (poplar) are among the most rapidly-growing trees of the northern hemisphere (Dickmann & Stuart, 1983) and have outsized roles in natural ecosystems as keystone species (Brunner *et al.*, 2004; Kouki *et al.*, 2004; Bailey & Whitham, 2006; Kivinen *et al.*, 2020). They are also of major economic importance for agroforestry, and as sources of wood, fiber, and biofuel (Sun *et al.*, 2021; Fuertes *et al.*, 2021). In either context, the growth and asexual propagation of poplar relies on the rapid establishment, proliferation and maintenance of a robust root system for nutrient and water absorption. Elite hybrid clones of poplar are propagated in stool beds through a process that relies on the ability of cuttings to undergo adventitious rooting (Stanton *et al.*, 2019). In addition, asexual reproduction of poplar in nature commonly occurs via the process of "root sprouting" from existing roots or the root-shoot junction zone (Wiehle *et al.*, 2009). Moreover, biotic and abiotic stresses such as waterlogging and pest damage often interfere with above-ground tree health by damaging root systems (Brandt *et al.*, 2003; Štícha *et al.*, 2016), a stress that is certain to be exacerbated by climate

change (Overpeck & Udall, 2020; Gullino *et al.*, 2021). A deeper understanding of the genes that control rooting may provide new insights into means for improved propagation of recalcitrant genotypes and species, options for improvement of regeneration during genetic engineering/editing, and suggest new strategies for mitigating stress in managed and wild populations.

Adventitious rooting in poplar is a highly complex trait that is regulated by many factors, including plant age, genotype, and physiology; the many forms of plant stress; and environmental cues such as temperature, photoperiod, and nutrients. These act through phytohormone signaling cascades that lead to differentiation and development of root tissue. Overexpression and RNAi-mediated suppression of over two dozen genes involved in phytohormone synthesis and response have been reported to lead to increased or decreased adventitious root formation and/or root growth (reviewed by Bannoud & Bellini, 2021). These root-related traits have been shown to be genotype-dependent, with phenotypic variation across *Populus spp.* depending in large part on variable sequence and expression of phytohormone-related genes, especially those involved in auxin pathways or crosstalk between auxin and other phytohormones (Ribeiro *et al.*, 2016; Sun *et al.*, 2019). Genome-wide association studies (GWAS) provide opportunity for insight into how variation in root-related traits across genotypes results from variation in these phytohormone regulators and other genes. To date, GWAS of adventitious rooting performed in plants including rice (Ribeiro *et al.*, 2016) and *Populus* (Sun *et al.*, 2019) have contributed to an improved understanding of these gene-function relationships and others.

GWAS of root traits, both from adventitious roots and non-adventitious roots, commonly involve measurement of the lengths, diameters, or types of roots, among other statistics. As collection of these traits can prove laborious and time-consuming, a wide array of computer vision tools have been developed to extract these features from root images, with varying degrees of human intervention or automation (e.g., Arsenault *et al.*, 1995; Das *et al.*, 2015; Zhang *et al.*, 2018; Yasrab *et al.*, 2019; Zheng *et al.*, 2020). These methods have been applied to enable GWAS and other genetic analyses of rooting in common bean, cowpea (Burridge *et al.*, 2016), maize (Arsenault *et al.*, 1995; Zhang *et al.*, 2018) and hybrid poplar (Sun *et al.*, 2019), among others. Methods involving manual measurements with rulers or ImageJ have been applied to GWAS of rooting traits in plants including maize (Trachsel *et al.*, 2011), rice (Courtois *et al.*,

2013; Li *et al.*, 2017; Wang *et al.*, 2018; Xu *et al.*, 2020; Zhang *et al.*, 2020), wheat (Ayalew *et al.*, 2018) and hybrid poplar (Dash *et al.*, 2018). The production of more general and user-friendly root phenotyping platforms can help to extend automated methods to more species and laboratories, potentially assisting expansion of GWAS population size by reducing the amount of manual labor needed for phenotyping. With increased population size, GWAS gain greater statistical power and improved ability to detect significant effects of genes regulating traits, particularly those or smaller effect size, and those imparted by rare alleles (López-Cortegano & Caballero, 2019).

Here, we report insights into the genetic control of rooting obtained through GWAS. By use of a large resequenced and highly polymorphic population of wild *Populus trichocarpa* that shows very rapid decay of linkage disequilibrium—and a novel machine vision phenomic system and multiple GWAS pipelines—we were able to statistically detect large numbers of candidate genes. The potential for rare allele discovery was enhanced by our use of the SNP-Set Sequence Kernel Association Test, which upweights and combines effects from rare SNPs to increase power in their detection (Wu *et al.*, 2011). We report a total of 277 unique associations passing significance thresholds, including many involved in hormone signaling, cell division, and post-translational modification of proteins—in addition to many genes of unknown function.

Methods and Materials

Plant materials

We used a *P. trichocarpa* GWAS population that was recently expanded to include a total of 1,323 genotypes (Yates *et al.*, 2021); subsets of this population were used in previous GWAS projects (Tuskan *et al.*, 2018; Bdeir *et al.*, 2019; Weighill *et al.*, 2019; Chhetri *et al.*, 2020; Chen *et al.*, 2021; Nagle *et al.*, 2022). This population is comprised of variation in wild *P. trichocarpa* spanning regions of British Columbia, Washington, Oregon, Idaho and northern California. Clone banks for this GWAS population were produced in multiple locations, among which a replicate in a Corvallis, OR field location was utilized to obtain cuttings for this study. This study was performed using materials collected as described in previous work (Nagle *et al.*, 2022). In summary, dormant cuttings were collected in the winters of 2018, 2019 and 2020, then rooted up to a year later and grown in a greenhouse. Finally, fresh cuttings were collected and frozen for 2-4 weeks, then used for rooting assays.

Assay of rooting

Cuttings were placed in 50mL Falcon tubes with water and allowed to root. Beginning two weeks later, images were collected at weekly timepoints for four weeks. Prior to taking each image, plants were removed from water and placed on top of a surface with roots arranged to separate putative lateral and basal roots. To aid their recognition by our machine vision pipeline, basal roots were laid downward atop blue felt while lateral roots were laid to the side on gray felt. Each image also included a label and ruler. Plants were imaged from above using a Canon Rebel XSi DSLR camera attached to a mount and facing downward. Due to practical limitations in the number of cuttings that could be studied at once, the study was divided into eight "phases," each of which featured ~400 cuttings, including two replicates for each of ~200 given genotypes.

Computer vision pipeline

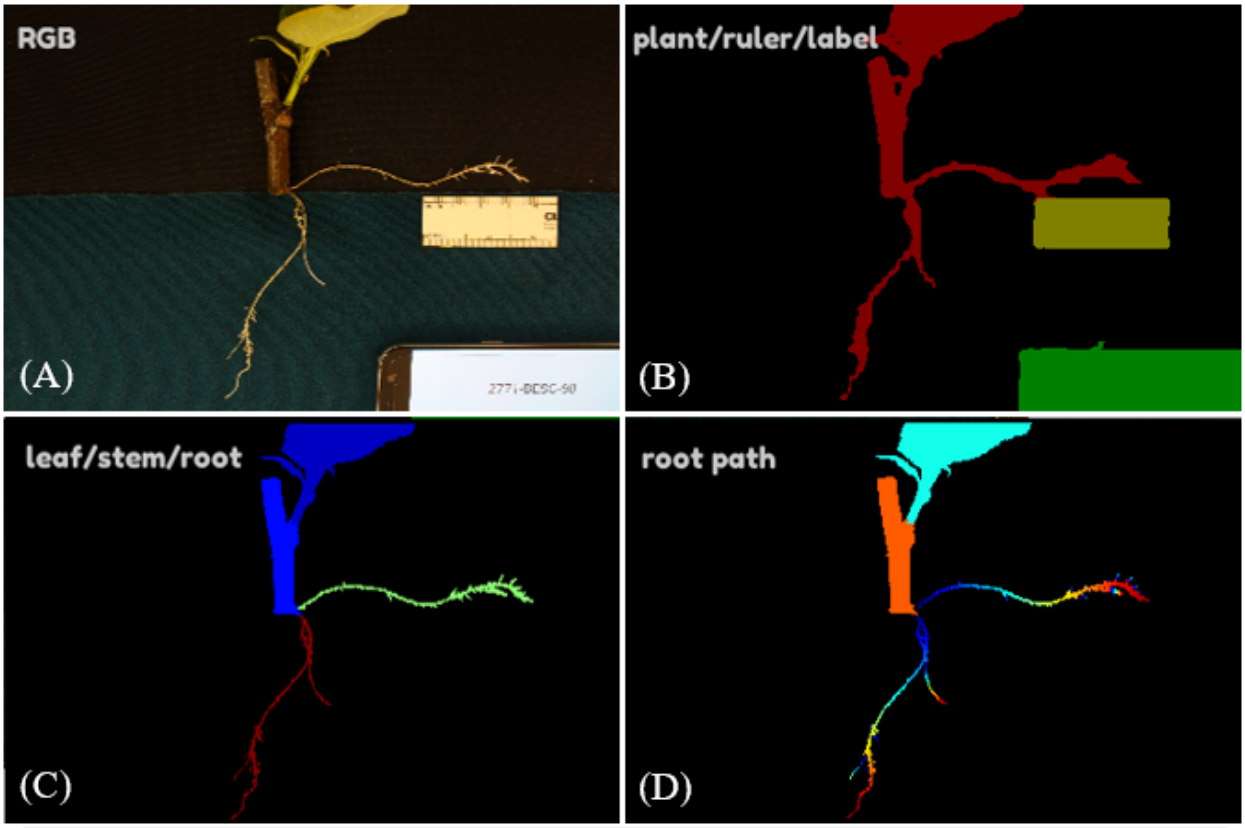
We adopted the DeepLab network (Chen *et al.*, 2018) with backbone ResNet50 (He *et al.*, 2016) as our segmentation model for its efficiency and accuracy. We trained two different segmentation models: the first was used to segment an image into background, plant, ruler, and label (Model 1); the second was used to segment the image into background, leaf, stem, and root (Model 2). Below, we introduce how we collected training labels to train the two networks and then used the two networks to measure biological traits of interest.

Because images were collected in various orientations, with the camera in either portrait or landscape mode, we first rotated images to a uniform orientation. Next, the background was segmented based on color thresholding. As the plant, ruler and label varied in color and were found in approximately similar positions from image to image, we first segmented these components based on their spatial positioning and colors using mean-shift segmentation and k-means clustering. Images successfully segmented as such were used to produce a training set with approximately 500 images used to train a deep model for segmentation of the remaining images. Inference was performed using this model, and correct examples were used to retrain the model with an expanded dataset, resulting in a final model trained with 2,239 examples (Deep Model 1). Afterward, to segment the plant into stem, root and leaf, we performed mean-shift segmentation and applied a location threshold to produce a training set of approximately 900 images. These were used to train a second deep model and the training set was again expanded

by running inference on new images and selecting correct results, resulting in a final training set with 3,496 training examples (Deep Model 2).

Following training of both deep segmentation models, they were applied for inference of the full dataset. All images were standardized to the same orientation, followed by deployment of both models. The final result of segmentation was separation of background, leaf, stem, root, ruler, and label for each image. Next, the segmentation results were further analyzed to produce statistics on biological traits of interest. Since the camera height varied across images, we computed the number of pixels per ruler width for each image to enable standardization via the actual size represented by a single pixel.

We proceeded to compute the lengths and area of roots in centimeters, as well as the diameters of stems. Root statistics were computed as follows. (1) First, we isolated the segment of root by distinct connections to the stem via connected components (2) The background was classified as top background and bottom background based on the color of the felt background below the stem, allowing each root to be classified as lateral or basal depending on the background. (3) For each basal and lateral root, we computed root length using a distance map and root area by counting pixels. Longest root length (LRL) and total root area were computed separately for basal and lateral roots in each image. RGB and false-color images of the segmentation process are shown in Fig. 1.



(E)	Total area (cm²)	Longest root length (cm)
Basal root	313.5	9.3
Lateral root	356.5	10.6

Figure 1. Workflow for phenotyping root traits: (A) RGB images were collected for plants with rulers and labels. (B) The first round of segmentation was performed to separate the plant, ruler and label. (C) A subsequent round of segmentation separated the roots by their type (basal or lateral, shown respectively in red and green). (D) Finally, the length of each root was measured from root tip to connection to stem. (E) An example of total area and longest root length computed over basal and lateral types of adventitious roots.

141 Data Error Checking Methods

142 Source images were compared to segmented images (e.g. Fig. 1) to identify cases where
143 data contained errors for the length, area, type or number of roots. Each image was scored
144 according to the error or errors observed (Table 1). Images were then sorted into folders

145 according to the types of data errors they contained using a spreadsheet and an R script.

146

Error ID	Description	Frequency among images
0	No significant error	81.3%
1	Undetected roots	1.3%
2	Incorrect number of roots	7.2%
3	Undetected basal root, but detected lateral root	0.2%
4	Incorrectly identified feature (e.g. basal root labeled as lateral)	4.6%
5	Root label shorter than root in image (truncation)	2.9%
6	Errant detection of root where no root exists	0.3%

Table 1. Descriptions of error types and their frequencies

147 Data Correction Methods

148 ImageJ version 1.53a software and the SmartRoot version 4.21 plugin were installed and used to
149 make all measurements. Source images were opened in ImageJ with the SmartRoot plugin then
150 converted to an 8-bit greyscale and inverted. The global scale was then set to 1 pixel/unit in
151 ImageJ and 2.54 DPI in SmartRoot to acquire units of pixels for all measurements. Roots that
152 required correction of length data were then traced using the SmartRoot Trace Root tool. To
153 calculate root areas, the image color threshold was adjusted until only the roots were highlighted,
154 then the area was selected using the Wand Tracing tool and measured in ImageJ. Root types
155 were corrected based on manual judgement of whether a root was a lateral root or basal root,
156 while recalling that these were placed on different areas of felt during imaging as previously
157 described.

158 Data preparation

159 Mean values of each trait were computed across replicates for each genotype and used for
160 downstream GWAS. Principal component analysis (PCA) was performed using `stats::princomp`
161 in R to produce PCs representing trends across traits and timepoints. PCA was performed over
162 three batches of traits: 1) area of basal or lateral roots at all timepoints; 2) LRL for basal or
163 lateral roots at all timepoints; and 3) all traits, including area and LRL of basal or lateral root at
164 all timepoints. Scree plots were consulted to determine the numbers of PCs, representing
165 significant variation for each batch, to be used for downstream GWAS.

Normality of traits was assessed in parallel to traits for our published GWAS of *in planta* regeneration in poplar, using the same methods (Nagle *et al.*, 2022). In short, we assessed normality of untransformed traits using Q-Q plots, histograms, Shapiro-Wilks tests and Pearson correlation coefficients with theoretical normal distributions, then applied necessary transformations including Box-Cox transformations, rank-based inverse normal transformations, removal of zero values and removal of outliers on a case-by-case basis for conservative but adequate transformation of each trait (Figure S2, Table S1-2).

Association mapping

Association mapping of all traits was performed in parallel with traits in our published GWAS of *in planta* regeneration, using methods and SNP sets detailed in this previous work (Nagle *et al.*, 2022). To summarize, four GWAS methods were used: 1) Genome-wide Efficient Mixed Model Association (GEMMA; Zhou & Stephens, 2012); 2) Generalized Mixed Model Association Test (GMMAT; Chen *et al.*, 2016); 3) Fixed and Random Model Circulating Probability Unification (FarmCPU; Liu *et al.*, 2016), specifically the implementation FarmCPUpp (Kusmec & Schnable, 2018); and 4) SNP-set (sequence) Kernel Association Test (SKAT; Ionita-Laza *et al.*, 2013) with the Multi-Threaded Monte Carlo SKAT (MTMC-SKAT) R extension we developed. Versions of trait data transformed toward normal distributions were analyzed with GEMMA and FarmCPU, while binarized traits were analyzed with GMMAT and untransformed traits with MTMC-SKAT using resampling to avoid violations of linear model assumptions for high-confidence associations. GEMMA and GMMAT were run using kinship matrices to adjust for population stratification, with a set of 13.2 million SNPs with a minor allele frequency (MAF) threshold of 1% and that are missing in no more than 10% of genotypes. FarmCPU uses a novel approach to adjust for population stratification while avoiding overcorrection and was run with a SNP set with MAF above 5%, missing rate below 10% and pruning based on linkage disequilibrium (LD). MTMC-SKAT was run using six principal components derived from SNP data to correct for population stratification and a set of 34.0 M SNPs with missing rate below 15%. MTMC-SKAT was deployed on the high-performance cluster COMET, made available through NSF XSEDE (Towns *et al.*, 2014). Stem diameter and phase were used as covariates for all GWAS methods.

To identify QTLs from results that are statistically significant, we computed multiple testing correction thresholds using the Bonferroni method (parameters: $\alpha = 0.05$, N tests equal to

197 *N* SNPs) and Benjamini-Hochberg false discovery rate ($\alpha = 0.10$). We further sought to identify
198 candidate genes that failed to meet significance according to either of these criteria, but were
199 represented by a peak of QTLs showing a pattern of LD decay suggestive of a causative
200 association. Toward this end, we applied an implementation of the Augmented Rank Truncation
201 (ART; Vsevolozhskaya *et al.*, 2019) over GEMMA and GMMAT results as we previously
202 described (Nagle *et al.*, 2022).

203 **Results**

204 Principal components describe complex patterns of root development

205 Significant patterns of root development over time and across root types (basal and
206 lateral) were summarized by PCA. For each of the three batches of traits used for PCA, the top
207 two PCs appear to explain significant portions of variance as indicated by Scree plots. PCA was
208 performed over three batches as previously described (Methods and Materials). The first two
209 batches, for root area or root length traits, each across basal and lateral roots and timepoints, both
210 produced a first PC representing lateral root development independent of basal root and a second
211 PC representing a lack of basal root development independent of lateral root (Fig. 2, Fig. S1).
212 For the third batch, including all area and length traits together, the first PC shows a trend of
213 overall root development and the second shows a preference for development of basal root rather
214 than lateral root (Fig. S2).

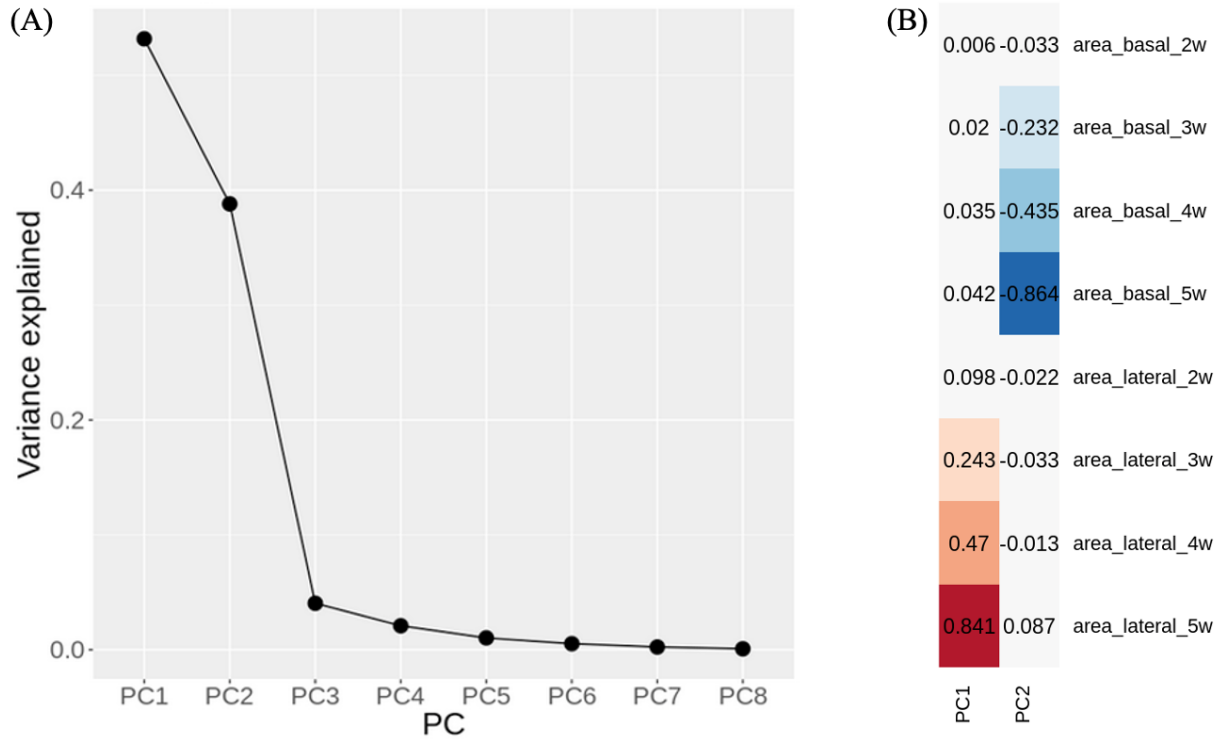


Figure 2. Results from PCA over all root area traits, across root type (basal or lateral) and all four timepoints of data collection: (A) Scree plot showing proportion of variance explained by each PC; (B) Loadings for top two PCs.

215

216 Use of multiple GWAS methods yielded numerous associations

217 We evaluated results to identify candidate QTLs and genes implicated for all 13 traits with h^2_{SNP}
218 above 0.10, three of which were evaluated with two different transformations due to ambiguity in
219 the ideal transformation (Table S1-S3). We term a "QTL peak" as any SNP or SNP window
220 associated with a given trait that is not within 30kb of any other SNP or SNP window with a
221 lower *p*-value for the associated trait, thus appearing as the peak position of a group of signals on
222 a Manhattan plot. Only MTMC-SKAT yielded QTL peaks passing the conventional Bonferroni
223 threshold under the assumption that each test (SNP window) is independent. MTMC-SKAT
224 using 3kb SNP windows yielded a total of 31 unique Bonferroni-significant associations across
225 all traits, 22 of which had a window center within 5kb of an annotated gene, as well as 164
226 unique associations passing the FDR ($\alpha = 0.10$) and/or the conservative Bonferroni threshold
227 (112 of which had window centers within 5kb of a gene). The apparent power of SKAT is in

contrast to GEMMA and GMMAT; GEMMA yielded only two associations passing the FDR (α = 0.10) threshold and none passing the conservative Bonferroni threshold, while GMMAT yielded none passing either. Statistical power for GEMMA and GMMAT was greatly increased by the use of ART to combine signals across windows of SNPs, which enabled detection of over 100 associations for GEMMA and eight for GMMAT (Fig. 3-4). Fig.5 provides an example of manual inspection of QTL peaks using integrative genomics viewer (IGV; Thorvaldsdóttir *et al.*, 2013). Associations of note that are addressed in the discussion section are presented in Table 2, with summary statistics for all associations in Table S4 and details presented in Table S5-8. No significant associations were found using FarmCPU.

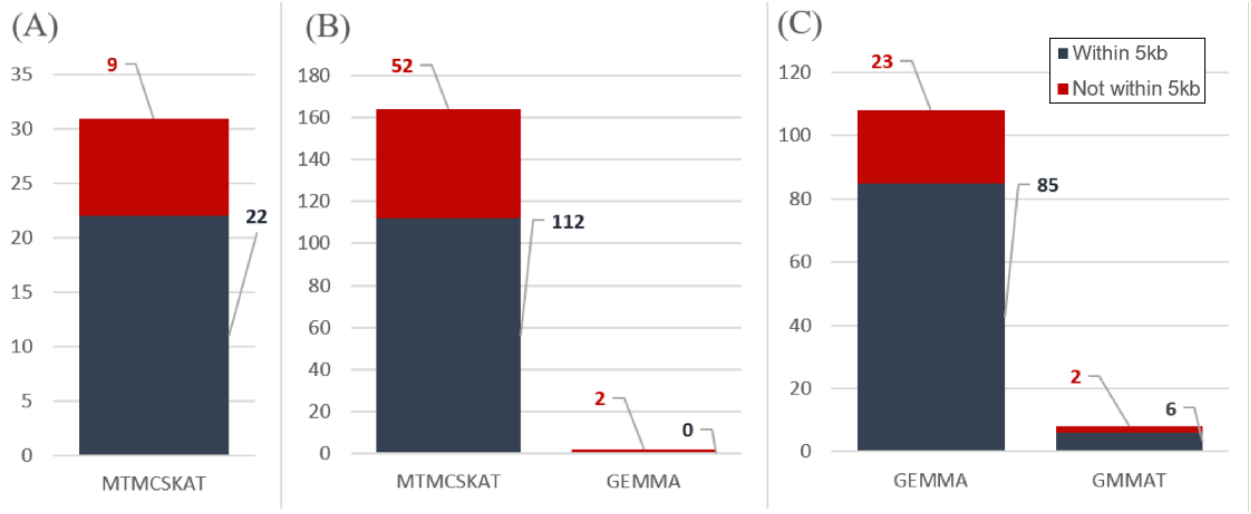


Figure 3. Barplots summarizing the numbers of associations from each GWAS method, with each type of significance threshold, as well as within a 5kb distance threshold to the nearest gene. QTL peaks were taken as the point with the lowest *p*-value at any given position near a significant SNP, where multiple points within the same peak may otherwise pass a given significance threshold. (A) QTL peaks passing the conservative Bonferroni threshold, given an assumption of independence of all SNP associations; (B) QTL peaks passing Benjamini-Hochberg threshold (FDR; α = 0.10) and/or the conservative Bonferroni Threshold. (C) QTL peaks passing ART-Bonferroni threshold (α = 0.05, N of # 1kb windows in genome); ART was only applied to the single-SNP GWAS methods, GEMMA and GMMAT.

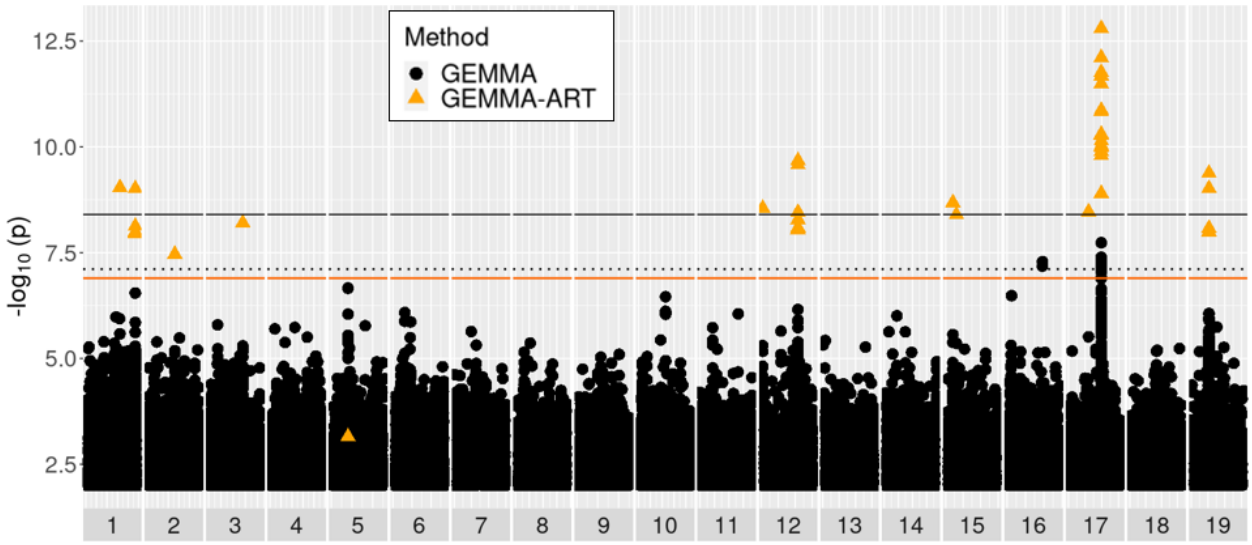


Figure 4. Manhattan plot of GEMMA results for the trait of longest lateral root length at week 3. Black and orange solid lines represent Bonferroni significance thresholds for GEMMA results with independent SNPs, and for ART applied to GEMMA over 1kb windows of SNPs. The black dotted line represents the significance threshold with a false discovery rate of 10% for independent SNP tests. Black circles represent tests of individual SNPs by GEMMA. Orange triangles represent 1kb windows tested by ART applied to GEMMA results.

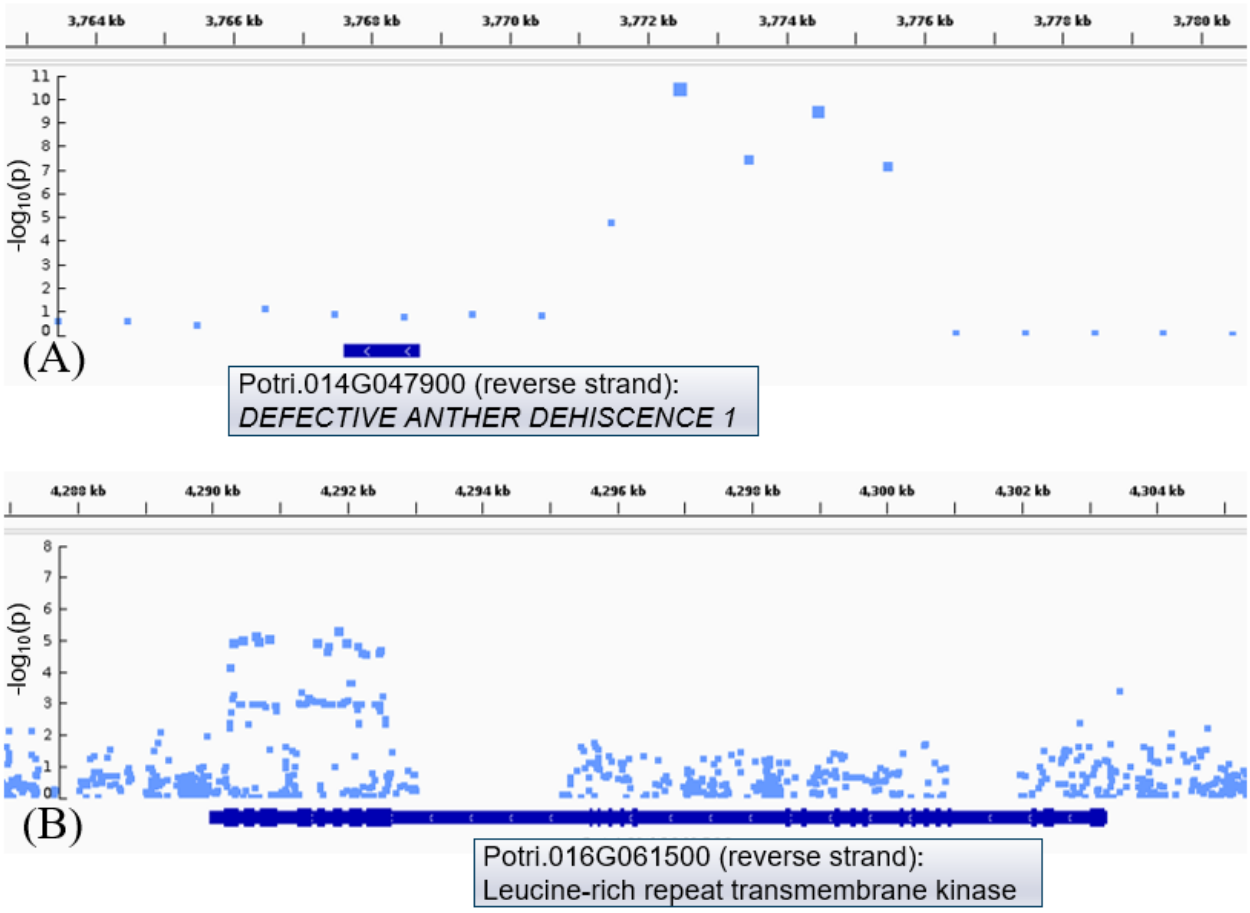


Figure 5. Close-up view of Manhattan plots aligned with *P. trichocarpa* genome annotation (v3.1) using IGV. (A) SKAT results (prior to resampling top associations with MTMC-SKAT) for LRL at week three display an association with a possible promoter region of a putative homolog of *DEFECTIVE ANTHER DEHISCENCE 1;* (B) GEMMA results (without GEMMA-ART displayed) for basal root area at week five display an association with a ~2kb region of exons and short introns of a putative leucine repeat rich transmembrane protein kinase. Exons are visualized as thickened portions on the gene track. Plots are displayed without MTMC-SKAT and GEMMA-ART *p*-values for simplicity; these statistics can be found in Table S6-7. Boxes with gene accession IDs, strand of gene and gene info were added manually to IGV plots.

Gene candidates							*Arabidopsis homologs*			
Threshold	**Trait name**	**Method**	**Transformation**	**Dist.**	**QTL Pos.**	**Accession ID**	**Description**	**Accession ID**	**Score**	**Similarity**
Bonf.	RA PC2	MTMC-SKAT	Untransformed	3,509	Intragenic, non-exonic	Potri. 001G149200	novel plant snare 11	AT 2G35190	403	92.3%
Bonf.	RA PC2; Basal RA growth wk. 2-5; Basal RA wk. 5	MTMC-SKAT	Untransformed	2,034	5'	Potri. 003G054300	lincRNA	NA	NA	NA
FDR (α = 0.1)	Overall PC2	MTMC-SKAT	Untransformed	22	3'	Potri. 004G210600	FASCICLIN-like arabinogalactan-protein 12	AT 5G60490	189	83.0%
Bonf.	RA PC1	MTMC-SKAT	Untransformed	52	5'	Potri. 005G141900	CYCLIN D3;2	AT 5G67260	361	74.9%
Bonf.	RA PC2; Basal RA growth wk. 2-5; Basal RA wk. 5	MTMC-SKAT	Untransformed	730; 270	5'; Intragenic, non-exonic	Potri. 006G161200	indoleacetic acid-induced protein 16	AT 3G04730	162	86.0%
Bonf.	Lateral LRL (wk. 3)	MTMC-SKAT	Untransformed	596	5'	Potri. 006G193400	lincRNA	NA	NA	NA
Bonf.	RA PC2; LRL PC2; Basal RA growth (wk. 2-5); Basal RA (wk. 5); Overall PC2	MTMC-SKAT	Untransformed	1,570; 2,570	Exonic	Potri. 007G056400	histidine kinase 1	AT 2G17820	1634	81.8%
Bonf.	RA PC1	MTMC-SKAT	Untransformed	10,480; 11,480	5'	Potri. 010G031800	Protein kinase family protein	AT 5G18700	1828	81.7%

							with ARM repeat domain			
FDR (α = 0.1)	LRL PC2	MTMC-SKAT	Untransformed	4,967	5'	Potri. 011G028200	cysteine-rich RLK (RECEPTOR-like protein kinase) 25	AT 4G05200	393	65.1%
Bonf.	LRL (wk. 3)	MTMC-SKAT	Untransformed	3,775; 5,775	5'	Potri. 014G047900	DEFECTIVE ANTHER DEHISCENCE 1	AT 2G44810	560	88.0%
Bonf.	RA PC1	MTMC-SKAT	Untransformed	2,296	5'	Potri. 015G105000	dual specificity protein phosphatase family protein	AT 5G23720	1189	81.8%
Bonf.	LRL (wk. 3)	MTMC-SKAT	Untransformed	1,181; 2,181	3'	Potri. 017G083100	lincRNA	NA	NA	NA
Bonf.	RA PC1	MTMC-SKAT	Untransformed	913; 87	3'; Exonic	Potri. 018G097000	Fasciclin-like arabinogalactan family protein	AT 3G46550	394	82.7%
Bonf.	RA PC1	MTMC-SKAT	Untransformed	1,437; 2,437	3'	Potri. 019G127200	Protein kinase family protein with leucine-rich repeat domain	AT 1G35710	893	68.9%
ART-Bonf.	Lateral LRL (wk. 3)	GEMMA	Box-Cox	1,165	5'	Potri. 001G436800	glutathione S-transferase TAU 25	AT 1G17180	308	80.7%
ART-Bonf.	RA PC2	GEMMA	Outliers removed, RB-INV	2,242	5'	Potri. 004G023100	Peroxidase superfamily protein	AT 3G01190; AT 5G15180	407; 403	75.4%; 75.8%
ART-Bonf.	RA growth (wk. 2-5)	GEMMA	Box-Cox	896	5'	Potri. 010G202000	actin-related protein 5	AT 3G12380	1113	87.3%
ART-Bonf.	Lateral LRL (wk. 3)	GEMMA	Box-Cox	1,753	5'	Potri. 012G081500	ataxia-telangiectasia mutated	AT 3G48190	350	58.8%

ART-Bonf.	Basal RA (wk. 5)	GEMMA	Box-Cox	0	Exonic	Potri. 016G061500	Leucine-rich repeat transmembrane protein kinase	AT 1G53440	622	62.7%
ART-Bonf.	RA growth (wk. 2-5)	GMMAT	Binarized trait	2,675	5'	Potri. 008G072700	ferritin 2; ferritin 4	AT 3G11050; AT 2G40300	373; 358	83.1%; 84.9%

Table 2. Twenty gene candidates with Arabidopsis homologs or encoding lincRNAs with putative role in biological processes of root development. For each candidate with an Arabidopsis homolog, relevant literature is summarized (Discussion). QTLs were identified from various traits related to root area (RA) and longest root length (LRL), for either basal or lateral ARs, or both types combined. For some associations found by MTMC-SKAT, two nearby SNP windows have equal *p*-values for association ($1*10^{-7}$) due to the practical limits of our permutation analysis. Remaining associations are presented in Tables S5-S8.

Discussion

As discussed in depth below, we reported a high number of putative associations from our GWAS pipeline. FarmCPU, however, gave no significant associations, and only a single candidate association in our prior study using this pipeline (Nagle *et al.*, 2022). We are unaware of any other reports of this GWAS method being used in poplar. We speculate that the lack of success with this method may stem from the high SNP density and rapid linkage disequilibrium (LD) decay in poplar, the unique approach FarmCPU employs in controlling for population stratification using a limited set of SNPs, and/or the LD-based pruning that was needed to avoid errors during this process for the SNP set we used.

Two putative modes of adventitious root regeneration in poplar

We observed that when ARs grew from the base of cuttings, the base was often enlarged, distorted, and disorganized in appearance, resembling the calli found in *in vitro* tissue cultures or at *in planta* wound sites (Fig. 6). We term these "basal ARs." Although we are aware of little research into this specific type of AR in poplar, research in *Populus balsamifera* (a closely related species interfertile with *P. trichocarpa*), suggested that AR development depends on calcium and pH, and that hard callus may inhibit root emergence (Cormack, 1965). More recently, basal ARs growing from callus have been studied in *Pinus* (Rasmussen *et al.*, 2009). These contrast with ARs growing from the sides of cuttings ("lateral ARs"), which appeared to grow directly from the stem without an intermediate callus stage (Fig. 6). Prior histological research in poplar indicated that lateral ARs appear to originate from secondary meristematic tissue in the cambium (Rigal *et al.*, 2012). We therefore set up our phenomics pipeline to measure lateral and basal ARs separately, assuming they are biologically distinct and likely originate from different progenitor cells. Their high degree of independence was supported by our PCA analysis, which by inspection of loadings appear to represent "ratios" between the two types of root (Fig. 2, Fig. S1-2). There was also a number of GWAS associations for one root type or the other, or for these PCs (discussed below).

Figure 6. Selected images collected during root phenotyping, with zoomed-in views showing callus or callus-like tissue at the base of stems from which basal ARs emerge. (A) Genotype BESC-153, (B) BESC-327, (C) GW-9914, and (D) BESC-337.

263

264 Phenomics workflow accelerated phenotyping but required human intervention

265 Our phenomics workflow using machine vision enabled us to extract root trait data from
266 a number of images that would otherwise have been infeasible for humans unassisted. The
267 training of models for this workflow did not require manual preparation of ground-truth semantic
268 labels by humans using an annotation interface, but rather utilized ground-truth labels prepared
269 by thresholding and cropping. Although this approach offers the advantage of accelerated
270 training dataset preparation, it lacks the ability for machine vision results to be compared to

human-produced ground truth labels using statistics such as intersection over union (IoU). To assess model performance, we inspected each image and grouped them according to notable errors that we observed. The most frequent errors were in root counting, found in at least 7.2% of images (Table 1). Root count statistics were not used in GWAS because of the relatively high error rate and because the summary statistics of aggregate root area provides a proxy for root system proliferation, while being more robust to errors. Another common type of error was of incorrectly applied labels, for example basal roots labeled as lateral roots or vice versa (Table 1). With our system, we attempted to facilitate the labeling of basal and lateral roots by manually placing roots of each type in areas with different background colors (Fig. 6). We are not aware of other root phenomics workflows that aim to distinguish basal and lateral ARs, although several distinguished root type for non-adventitious roots, such as for alfalfa (Xu *et al.*, 2022), Arabidopsis, wheat and *Brassica* (Yasrab *et al.*, 2019).

We attempted to correct errors using ImageJ with the SmartRoot plugin (Methods and Materials). This error correction was performed on approximately 18.7% of images. Considering our hybrid approach utilizing both machine vision and manual correction of errant labels, our method is comparable to RootReader (Clark *et al.*, 2013), a tool that similarly performs initial labeling of images and then allows for user correction. The scope of our work did not include the development of a user-friendly and generalizable root annotation method that can be practically applied by other labs with diverse imaging conditions and across diverse species of plants. However, several such tools have been recently developed, including RootNav 2.0, which has demonstrated an ability via transfer learning to generalize across different background types and species including maize, Arabidopsis and *Brassica* (Yasrab *et al.*, 2019).

Gene candidates represent diverse functional roles

Regulators of cell division and structure

D-type cyclins regulate the G1-to-S progression of the cell cycle and comprise a family of ten known proteins in Arabidopsis and 22 in poplar (Dong *et al.*, 2011). Potri.005G141900 encodes a homolog of CYCLIN D3;2 (CYCD3;2) or CYCLIN D3;3 (CYCD3;3), which are among at least five D-type cyclins known to have roles in root precursor cells in embryonic and/or mature tissues (De Veylder *et al.*, 1999; Nieuwland *et al.*, 2009; Collins *et al.*, 2012; Forzani *et al.*, 2014). A comparative transcriptomic study in *Populus* found this gene to be differentially expressed between two genotypes with contrasting rates of stem growth and

biomass accumulation (Han *et al.*, 2020). In Arabidopsis, concurrent knockout of all three members of the CYCD3 clade led to retarded seed development, while overexpression of CYCD3;1 led to premature, irregular and disorganized division of the hypophysis (Collins *et al.*, 2012). Transcription of CYCD3;1 and CYCD3;3 is negatively regulated by WUSCHEL-ASSOCIATED HOMEOBOX 5 (Forzani *et al.*, 2014), a poplar homolog of which enhances AR in poplar when overexpressed (Li *et al.*, 2018). Although we are unaware of any reports of mutant phenotypes resulting from CYCD3 overexpression or knockout in mature roots, several relevant root-related phenotypes have been reported for CYCD2 and CYCD4 members. CYCLIN D4;1 knockout was reported to reduce pericycle cell divisions as well as the number of lateral roots, while these phenotypes were rescued by exogenous auxin, perhaps due to auxin-responsiveness of D-type cyclins with overlapping functional roles (Nieuwland *et al.*, 2009). CYCD2;1 overexpression led to increased root apical meristem divisions and increased sensitivity to effects of exogenous auxin in promoting lateral root formation, while knockout led to reduced auxin sensitivity although no reduction in RAM divisions, possibly due to redundant homologs (Sanz *et al.*, 2011).

Potri.001G149200 encodes a homolog of NOVEL PLANT SNARE 11 (NPSN11), which is believed to interact with other SNARE proteins to provide energy needed for the fusion of membranes that give rise to the cell plate during cytokinesis (Zheng *et al.*, 2002). Knockout of NPSN11 alone yields no mutant phenotype in Arabidopsis, putatively due to redundancy with a similar SNARE protein. Defects in cytokinesis and embryo development are conferred by simultaneous knockout of NSPN11 and the functionally redundant SNAP33 (El Kasmi *et al.*, 2013).

Potri.004G210600 and Potri.018G097000 encode members of the FASICLIN-LIKE ARABINOGALACTAN (FLA) family, which consists of 21 members in Arabidopsis and has established roles in adhesion in cell walls, plasma membranes and extracellular matrices (reviewed by Zang *et al.*, 2015). Transcriptomic analysis of tension wood development in *Populus* provides support for a role of Potri.004G210600 in cell wall structure (Bygdell *et al.*, 2017). Knockout of Arabidopsis FLA4, also known as SALT OVERSENSITIVE 5, was reported to reduce root elongation, cell wall thickness and root tip swelling under salt stress (Shi *et al.*, 2003) and experiments with ethylene inhibitors suggest that FLA4 functions downstream of

ethylene signaling (Xu *et al.*, 2008). We are unaware of reports of root-related mutant phenotypes of other FLA members.

Cell expansion is influenced by bundling and rearrangements of actin filaments, regulated at least in part by auxin and cytokinin (Zhu & Geisler, 2015; Scheuring *et al.*, 2016; Arieti & Staiger, 2020). We report two gene candidates encoding putative actins or actin-like proteins, Potri.010G202000 and Potri.012G081500. In Arabidopsis, loss-of-function mutants of ACTIN 7 displayed reduced root length and reduced cell divisions in the proximal meristem (PM) of root tips as well as an increased number of transition zone (TZ) cells, while dual mutants of ACTIN 2 and ACTIN 8 presented a loss of root hairs and increases in TZ cells with a lack of PM effects (Kandasamy *et al.*, 2009; Takatsuka *et al.*, 2018). The variable effects of different auxins on root development, together with the lack of clear homology for these actin-related gene candidates, welcome physiological investigation of the mechanisms by which these gene candidates may affect root in poplar.

Regulators of hormone signaling

Potri.006G161200 encodes a member of the Aux/IAA F-Box protein family, believed to have 35 members in *Populus trichocarpa* (Kalluri *et al.*, 2007) and 29 in Arabidopsis (Overvoorde *et al.*, 2005), and appears to be a homolog of IAA16 or another member of the 29-gene IAA family. In Arabidopsis, Aux/IAA proteins have been well-characterized and are known to undergo auxin-dependent proteasomal degradation and to function via auxin-dependent protein-protein interactions that repress the transcriptional activity of various Auxin Response Factor (ARF) family members (reviewed by (Luo *et al.*, 2018)). A specific role for ARF members in regulating AR development is believed to function via the action of downstream *GRETCHEN HAGEN* family genes responsible for conjugating jasmonic acid (JA) into bioactive jasmonoyl-L-isoleucine (JA-Ile). Evidence for conservation of this pathway in *Populus* has been reported, with enhanced or delayed AR development respectively resulting from overexpression or knockdown of a homolog of TRANSPORT INHIBITOR RESPONSE 1, responsible for Aux/IAA proteasomal degradation and found to interact with a *Populus* homolog of IAA28 (Shu, et al. 2019).

Further evidence for a role of JA signaling in rooting of poplar is indicated by an association with Potri.014G047900, encoding a homolog of DEFECTIVE IN ANTHER DEHISCHENCE1 (DAD1), which catalyzes the first step of JA biosynthesis (Ishiguro *et al.*,

2001). Several possible roles for JA in adventitious rooting of *Populus* have been discussed in a recent review; in summary, these roles may include cross-talk with auxin signaling among others, and are evidenced to vary across genera (Bannoud & Bellini, 2021). Our previous GWAS of *in planta* regeneration in poplar support a major role for JA signaling in callus and shoot regeneration (Nagle *et al.*, 2022; and sources cited within), via pathways that are likely to be relevant to adventitious rooting considering the previously discussed emergence of basal ARs from callus.

Regulators of post-translational modifications

Among our candidate genes, we report a notable number of genes encoding putative catalysts of post-translational modifications (PTMs), including a histidine kinase (Potri.007G056400), serine/threonine kinases (Potri.011G028200, Potri.010G031800, Potri.019G127200 and Potri.016G061500), a serine/threonine phosphatase (Potri.015G105000) and a glutathione-*S*-transferase (Potri.001G436800). We are particularly unsure of the precise mechanisms by which these candidates affect root traits because their PTM activity may be highly nonspecific and a majority of plant genes are likely to undergo PTMs. Evidence has been found for PTMs of over 12,000 substrates in Arabidopsis (Xue *et al.*, 2022), including arabinogalactans (Schultz et al 2004) and microtubule proteins involved in cell structure and division (Parrotta *et al.*, 2014) as well as hormone signal regulators (reviewed by Hill, 2015). Specific interactions between PTM catalysts and other gene candidates can be interrogated via statistical tests for epistasis.

Arabidopsis mutants of certain PTM-related candidate homologs display phenotypes relevant to root development. Mutants of the serine/threonine phosphatase PROPYZAMIND-HYPERSENSITIVE 1 (homolog of Potri.015G105000) demonstrate embryo fatality (for null mutants) or microtubule defects leading to left-handed helical growth of roots in seedlings (for mutants with reduced phosphatase activity) (Naoi and Hashimoto, 2004). HISTIDINE KINASE 1 (homolog of Potri.007G056400) is believed to have a role in abscisic acid (ABA) signaling, indicated by increased sensitivity of seedlings to the effects of ABA in inhibiting germination (Tran *et al.*, 2007). In *Populus*, the putative histidine kinase Potri.007G056400 was found to be differentially expressed in *Populus* roots in response to boron deficiency (Su *et al.*, 2019) and the putative serine/threonine kinase Potri.015G105000 was previously identified in GWAS as an association with bud set and growth period (McKown *et al.*, 2014).

Noncoding RNAs

We sought to identify possible targets of putative ncRNAs that were found as associations in our GWAS, and indicated by the GreeNC pipeline to be probable ncRNAs (Di Marsico *et al.*, 2022). Potri.003G054300 shares significant homology with the predicted exon of Potri.006G260300, a gene that appears to encode a transmembrane protein but for which we were unable to identify a homolog in model species. We did not find any known protein-coding genes that Potri.006G193400 and Potri.017G083100 align well with. We note that noncoding RNAs may be involved in processes other than RNA interference of coding genes, such as in ribonucleoprotein complexes and chromatin modification (reviewed by Statello *et al.*, 2021). Whereas gene-silencing effects of ncRNAs can be predicted by sequence alignment, other roles may not be unraveled without wet-lab protocols (reviewed by Lucero *et al.*, 2021).

Regulators of reactive oxygen species (ROS) signaling

ROS can affect or correspond to root development through multiple mechanisms. High levels of ROS are associated with various biotic and abiotic stressors in plants (Sharma *et al.*, 2019; Qamer *et al.*, 2021) and are well-established as a cause of DNA and tissue damage across eukaryotes (Arfin *et al.*, 2021). As a means of post-transcriptional regulation, ROS can catalyze activation of deactivation of transcription factors (Wu *et al.*, 2012; Kong *et al.*, 2018) and other developmental genes such as cell cycle regulators (Yi *et al.*, 2014). Additionally, the previously discussed roles of auxin in root signaling relate to ROS as the bioactive auxin IAA is produced by a peroxisome-mediated reaction involving the precursor IBA, producing nitric oxide as a byproduct (reviewed by Damodaran & Strader, 2019). Moreover, nitric oxide is involved in nitrosylation of proteins including the auxin receptor TRANSPORT INHIBITOR RESPONSE (TIR1), promoting its interaction with Aux/IAA proteins (Terrile *et al.*, 2012) such as the previously discussed IAA16. NO and ROS also have roles in mediating symbioses with mycorrhizae as well as pathogen defense (reviewed by Martínez-Medina *et al.*, 2019) although these roles are likely not relevant to our root assays in our laboratory using water rather than soil.

Oxidative stress is mitigated in part by ferritins, proteins that sequester Fe and thus prevent Fe from reacting with oxygen and producing oxygen radicals. Potri.008G072700 encodes a homolog of the four-member ferritin family in Arabidopsis, and is most closely related to *FERRITIN 2* (greater Smith-Waterman alignment score) and *FERRITIN 4* (greater residue similarity). The latter homolog has been studied in the context of root system architecture. While

increasing concentrations of Fe in media led to an increase in lateral root density, this effect was abolished in triple knockouts of FERRITIN 1, 2 and 4 (Reyt *et al.*, 2015).

In addition to catalyzing a PTM of endogenous proteins as previously discussed, glutathione *S*-transferases (GSTs) such as Potri.001G436800 are well-known to have roles in detoxification of xenobiotics such as herbicides, and have lesser-characterized roles in regulation of redox balance via glutathione, an antioxidant (reviewed by Hernández Estévez & Rodríguez Hernández, 2020). Transgenic studies have found Arabidopsis lines overexpressing various GST homologs to have increased tolerance to oxidative stress (Sharma *et al.*, 2014; Xu *et al.*, 2017) and enhanced root proliferation, although more research is needed to determine the specific mechanism or mechanisms by which GSTs regulate root development, whether via ROS, PTMs or other roles (Chen *et al.*, 2012).

Other enzymes have more direct roles in oxidative stress and signaling, such as peroxidases that catalyze redox reactions (Yoshida *et al.*, 2003). Potri.004G023100 is an example of a putative peroxidase and is closely related to Arabidopsis accessions AT3G01190 and AT5G15180, both of which are highly expressed in root apices (Klepikova *et al.*, 2016) but have not been characterized in mutant studies to our knowledge.

Agreement with previous GWAS of adventitious rooting in poplar

Very few of our gene candidates were also identified as possible regulators of adventitious root traits in published studies. In a prior GWAS in *Populus deltoides x simonii* that employed 434 genotypes and yield 224 QTLs, with multiple possible candidate genes being proposed for a given QTL, an uncharacterized gene believed to be a transcription factor (Potri.005G154200), as well as a putative xyloglucan endotransglucosylase/hydrolase (Potri.016G098600), were associated with rooting traits. They appeared as associations with lateral LRL at week three in our work, and with total number of roots in this prior work (Table S6). Potri.015G026500 encodes an uncharacterized putative phospholipase that is found as associated with lateral LRL at week 3 in our work, and with root volume and total root number in this prior work. Potri.016G098500 encodes a putative heme-binding protein we found to be associated with the first PC of LRL traits (across root types and timepoints) and that Sun et al. (2019), found associated with LRL (Table S7). Possible reasons for the relatively low level of overlap between these studies include differences between species of poplar, variation in GWAS

455 populations and statistical methods, as well as the use of different rooting assays (Sun *et al.*,
456 2019).

457 **Conclusion**

458 We performed GWAS to identify regulators of adventitious rooting capacity in 1,148
459 genotypes from a *P. trichocarpa* clone bank. To facilitate the collection of quantitative measures
460 of adventitious root development, we employed a phenotyping system tailored for our
461 adventitious rooting assay in poplar. The hundreds of gene candidates identified include
462 regulators of cell division and structure, hormone signaling, reactive oxygen species signaling
463 and post-translational modifications as well as many genes of miscellaneous or unknown
464 function. The distinct origins of basal and lateral roots were supported both by our multivariate
465 phenotype analysis and GWAS associations. As root development is a complex and polygenic
466 process, future research will benefit from investigation of interactions between genes such as the
467 candidates identified here, functional studies such as through mutagenesis, and differential
468 rooting responses to environmental treatments.

469 **Acknowledgements**

470 We thank the National Science Foundation Plant Genome Research Program for support
471 (IOS #1546900, Analysis of genes affecting plant regeneration and transformation in poplar),
472 and members of GREAT TREES Research Cooperative at OSU for its support of the Strauss
473 laboratory.

474 Support for the Poplar GWAS dataset is provided by the U.S. Department of Energy,
475 Office of Science Biological and Environmental Research (BER) via the Center for Bioenergy
476 Innovation (CBI) under Contract No. DE-PS02-06ER64304. The Poplar GWAS Project used
477 resources of the Oak Ridge Leadership Computing Facility and the Compute and Data
478 Environment for Science at Oak Ridge National Laboratory, which is supported by the Office of
479 Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725. We would
480 like to thank the efforts of personnel from the CBI in establishing the GWAS resource used for
481 this study.

482 This work used the COMET high-performance cluster at the San Diego Supercomputing
483 Center (University of California, San Diego) made available through the Extreme Science and

Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562.

Author Contributions

Strauss, Li, Jiang, and Muchero designed and directed the overall study, and obtained funding for its execution; Ma, Peremyslova, Magnuson, and Goddard designed and/or executed the phenotypic analyses; Nagle, Yuan, and Damanpreet created, adapted, and executed the machine vision, computation, and data analysis pipelines; Niño de Rivera assisted with inspecting results in IGV. Nagle wrote the manuscript with editing from Strauss, and all others contributed further edits and revisions.

Data Availability

Raw data and code used for this project is available upon request to the authors. MTMC-SKAT is available on GitHub (https://github.com/naglemi/mtmcskat).

References

Arfin S, Jha NK, Jha SK, Kesari KK, Ruokolainen J, Roychoudhury S, Rathi B, Kumar D. 2021. Oxidative Stress in Cancer Cell Metabolism. *Antioxidants* **10**: 642.

Arieti RS, Staiger CJ. 2020. Auxin-induced actin cytoskeleton rearrangements require AUX1. *The New Phytologist* **226**: 441–459.

Arsenault J-L, Poulcur S, Messier C, Guay R. 1995. WinRHlZO™, a Root-measuring System with a Unique Overlap Correction Method. *HortScience* **30**: 906D – 906.

Ayalew H, Liu H, Börner A, Kobiljski B, Liu C, Yan G. 2018. Genome-Wide Association Mapping of Major Root Length QTLs Under PEG Induced Water Stress in Wheat. *Frontiers in Plant Science* **9**: 1759.

Bailey JK, Whitham TG. 2006. Interactions between cottonwood and beavers positively affect sawfly abundance. *Ecological Entomology* **31**: 294–297.

Bannoud F, Bellini C. 2021. Adventitious Rooting in Populus Species: Update and Perspectives. *Frontiers in Plant Science* **12**: 918.

Bdeir R, Muchero W, Yordanov Y, Tuskan GA, Busov V, Gailing O. 2019. Genome-wide association studies of bark texture in Populus trichocarpa. *Tree Genetics & Genomes* **15**: 14.

Brandt JP, Cerezke HF, Mallett KI, Volney WJA, Weber JD. **2003**. Factors affecting trembling aspen (Populus tremuloides Michx.) health in the boreal forest of Alberta, Saskatchewan, and Manitoba, Canada. *Forest Ecology and Management* **178**: 287–300.

Brunner AM, Busov VB, Strauss SH. **2004**. Poplar genome sequence: functional genomics in an ecologically dominant plant species. *Trends in Plant Science* **9**: 49–56.

Burridge J, Jochua CN, Bucksch A, Lynch JP. **2016**. Legume shovelomics: High—Throughput phenotyping of common bean (Phaseolus vulgaris L.) and cowpea (Vigna unguiculata subsp, unguiculata) root architecture in the field. *Field Crops Research* **192**: 21–32.

Bygdell J, Srivastava V, Obudulu O, Srivastava MK, Nilsson R, Sundberg B, Trygg J, Mellerowicz EJ, Wingsle G. **2017**. Protein expression in tension wood formation monitored at high tissue resolution in Populus. *Journal of Experimental Botany* **68**: 3405–3417.

Chen J-H, Jiang H-W, Hsieh E-J, Chen H-Y, Chien C-T, Hsieh H-L, Lin T-P. **2012**. Drought and Salt Stress Tolerance of an Arabidopsis Glutathione S-Transferase U17 Knockout Mutant Are Attributed to the Combined Effect of Glutathione and Abscisic Acid. *Plant Physiology* **158**: 340–351.

Chen L-C, Papandreou G, Kokkinos I, Murphy K, Yuille AL. **2018**. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **40**: 834–848.

Chen H, Wang C, Conomos MP, Stilp AM, Li Z, Sofer T, Szpiro AA, Chen W, Brehm JM, Celedón JC, *et al.* 2016. Control for Population Structure and Relatedness for Binary Traits in Genetic Association Studies via Logistic Mixed Models. *The American Journal of Human Genetics* **98**: 653–666.

Chen Y, Wu H, Yang W, Zhao W, Tong C. **2021**. Multivariate linear mixed model enhanced the power of identifying genome-wide association to poplar tree heights in a randomized complete block design. *G3 Genes|Genomes|Genetics* **11**.

Chhetri HB, Furches A, Macaya-Sanz D, Walker AR, Kainer D, Jones P, Harman-Ware AE, Tschaplinski TJ, Jacobson D, Tuskan GA, *et al.* 2020. Genome-Wide Association Study of Wood Anatomical and Morphological Traits in Populus trichocarpa. *Frontiers in Plant Science* **11**: 1391.

Clark RT, Famoso AN, Zhao K, Shaff JE, Craft EJ, Bustamante CD, Mccouch SR, Aneshansley DJ, Kochian LV. **2013**. High-throughput two-dimensional root system phenotyping platform facilitates genetic analysis of root growth and development. *Plant, Cell & Environment* **36**: 454–466.

Collins C, Dewitte W, Murray JAH. **2012**. D-type cyclins control cell division and developmental rate during Arabidopsis seed development. *Journal of Experimental Botany* **63**: 3571–3586.

Cormack RGH. **1965**. The effects of calcium ions and pH on the development of callus tissue on stem cuttings of basalm poplar. *Canadian Journal of Botany* **43**: 75–83.

Courtois B, Audebert A, Dardou A, Roques S, Herrera TG-, Droc G, Frouin J, Rouan L, Gozé E, Kilian A, *et al.* 2013. Genome-Wide Association Mapping of Root Traits in a Japonica Rice Panel. *PLOS ONE* **8**: e78037.

Damodaran S, Strader LC. **2019**. Indole 3-Butyric Acid Metabolism and Transport in Arabidopsis thaliana. *Frontiers in Plant Science* **10**.

Das A, Schneider H, Burridge J, Ascanio AKM, Wojciechowski T, Topp CN, Lynch JP, Weitz JS, Bucksch A. **2015**. Digital imaging of root traits (DIRT): a high-throughput computing and collaboration platform for field-based root phenomics. *Plant Methods* **11**: 51.

Dash M, Yordanov YS, Georgieva T, Wei H, Busov V. **2018**. Gene network analysis of poplar root transcriptome in response to drought stress identifies a PtaJAZ3PtaRAP2.6-centered hierarchical network. *PLOS ONE* **13**: e0208560.

De Veylder L, de Almeida Engler J, Burssens S, Manevski A, Lescure B, Van Montagu M, Engler G, Inzé D. **1999**. A new D-type cyclin of Arabidopsis thaliana expressed during lateral root primordia formation. *Planta* **208**: 453–462.

Dickmann DI, Stuart KW. **1983**. The culture of poplars in eastern North America. *The culture of poplars in eastern North America.*: 118–134.

Di Marsico M, Paytuvi Gallart A, Sanseverino W, Aiese Cigliano R. **2022**. GreeNC 2.0: a comprehensive database of plant long non-coding RNAs. *Nucleic Acids Research* **50**: D1442–D1447.

Dong Q, Zhao Y, Jiang H, He H, Zhu S, Cheng B, Xiang Y. **2011**. Genome-wide identification and characterization of the cyclin gene family in Populus trichocarpa. *Plant Cell, Tissue and Organ Culture (PCTOC)* **107**: 55.

El Kasmi F, Krause C, Hiller U, Stierhof Y-D, Mayer U, Conner L, Kong L, Reichardt I, Sanderfoot AA, Jürgens G. **2013**. SNARE complexes of different composition jointly mediate membrane fusion in Arabidopsis cytokinesis. *Molecular Biology of the Cell* **24**: 1593–1601.

Forzani C, Aichinger E, Sornay E, Willemsen V, Laux T, Dewitte W, Murray JAH. **2014**. WOX5 suppresses CYCLIN D activity to establish quiescence at the center of the root stem cell niche. *Current biology: CB* **24**: 1939–1944.

Fuertes A, Oliveira N, Cañellas I, Sixto H, Rodríguez-Soalleiro R. **2021**. An economic overview of Populus spp. in Short Rotation Coppice systems under Mediterranean conditions: An assessment tool for decision-making. *Renewable and Sustainable Energy Reviews* **151**: 111577.

Gullino ML, Albajes R, Al Jboory I, Angelotti F, Chakraborty S, Garrett KA, Hurley BP, Juroszek P, Makkouk K, Pan X, *et al.* 2021. *Summary for policymakers of the scientific review*

of the impact of climate change on plant pests: a global challenge to prevent and mitigate plant pest risks in agriculture, forestry and ecosystems. Rome: FAO: IPPC, 2021.

Han X, An Y, Zhou Y, Liu C, Yin W, Xia X. 2020. Comparative transcriptome analyses define genes and gene modules differing between two Populus genotypes with contrasting stem growth rates. *Biotechnology for Biofuels* **13**: 139.

He K, Zhang X, Ren S, Sun J. 2016. Deep Residual Learning for Image Recognition. In: 770–778.

Hernández Estévez I, Rodríguez Hernández M. 2020. "Plant Glutathione S-transferases: An overview". *Plant Gene* **23**: 100233.

Hill K. 2015. Post-translational modifications of hormone-responsive transcription factors: the next level of regulation. *Journal of Experimental Botany* **66**: 4933–4945.

Ionita-Laza I, Lee S, Makarov V, Buxbaum JD, Lin X. 2013. Sequence Kernel Association Tests for the Combined Effect of Rare and Common Variants. *American Journal of Human Genetics* **92**: 841–853.

Ishiguro S, Kawai-Oda A, Ueda J, Nishida I, Okada K. 2001. The DEFECTIVE IN ANTHER DEHISCENCE1 Gene Encodes a Novel Phospholipase A1 Catalyzing the Initial Step of Jasmonic Acid Biosynthesis, Which Synchronizes Pollen Maturation, Anther Dehiscence, and Flower Opening in Arabidopsis. *The Plant Cell* **13**: 2191–2209.

Kalluri UC, DiFazio SP, Brunner AM, Tuskan GA. 2007. Genome-wide analysis of Aux/IAA and ARF gene families in Populus trichocarpa. *BMC Plant Biology* **7**: 59.

Kandasamy MK, McKinney EC, Meagher RB. 2009. A single vegetative actin isovariant overexpressed under the control of multiple regulatory sequences is sufficient for normal Arabidopsis development. *The Plant Cell* **21**: 701–718.

Kivinen S, Koivisto E, Keski-Saari S, Poikolainen L, Tanhuanpää T, Kuzmin A, Viinikka A, Heikkinen RK, Pykälä J, Virkkala R, *et al.* 2020. A keystone species, European aspen (Populus tremula L.), in boreal forests: Ecological role, knowledge needs and mapping using remote sensing. *Forest Ecology and Management* **462**: 118008.

Klepikova AV, Kasianov AS, Gerasimov ES, Logacheva MD, Penin AA. 2016. A high resolution map of the Arabidopsis thaliana developmental transcriptome based on RNA-seq profiling. *The Plant Journal* **88**: 1058–1070.

Kong X, Tian H, Yu Q, Zhang F, Wang R, Gao S, Xu W, Liu J, Shani E, Fu C, *et al.* 2018. PHB3 Maintains Root Stem Cell Niche Identity through ROS-Responsive AP2/ERF Transcription Factors in Arabidopsis. *Cell Reports* **22**: 1350–1363.

Kouki J, Arnold K, Martikainen P. 2004. Long-term persistence of aspen – a key host for many threatened species – is endangered in old-growth conservation areas in Finland. *Journal for Nature Conservation* **12**: 41–52.

Kusmec A, Schnable PS. **2018**. FarmCPUpp: Efficient large-scale genomewide association studies. *Plant Direct* **2**: e00053.

Li X, Guo Z, Lv Y, Cen X, Ding X, Wu H, Li X, Huang J, Xiong L. **2017**. Genetic control of the root system in rice under normal and drought stress conditions by genome-wide association study. *PLOS Genetics* **13**: e1006889.

Li J, Zhang J, Jia H, Liu B, Sun P, Hu J, Wang L, Lu M. **2018**. The WUSCHEL-related homeobox 5a (PtoWOX5a) is involved in adventitious root development in poplar. *Tree Physiology* **38**: 139–153.

Liu X, Huang M, Fan B, Buckler ES, Zhang Z. **2016**. Iterative Usage of Fixed and Random Effect Models for Powerful and Efficient Genome-Wide Association Studies. *PLOS Genetics* **12**: e1005767.

López-Cortegano E, Caballero A. **2019**. Inferring the Nature of Missing Heritability in Human Traits Using Data from the GWAS Catalog. *Genetics* **212**: 891–904.

Lucero L, Ferrero L, Fonouni-Farde C, Ariel F. **2021**. Functional classification of plant long noncoding RNAs: a transcript is known by the company it keeps. *New Phytologist* **229**: 1251–1260.

Luo J, Zhou J-J, Zhang J-Z. **2018**. Aux/IAA Gene Family in Plants: Molecular Structure, Regulation, and Function. *International Journal of Molecular Sciences* **19**: 259.

Martínez-Medina A, Pescador L, Terrón-Camero LC, Pozo MJ, Romero-Puertas MC. **2019**. Nitric oxide in plant–fungal interactions. *Journal of Experimental Botany* **70**: 4489–4503.

McKown AD, Klápště J, Guy RD, Geraldes A, Porth I, Hannemann J, Friedmann M, Muchero W, Tuskan GA, Ehlting J, *et al.* 2014. Genome-wide association implicates numerous genes underlying ecological trait variation in natural populations of Populus trichocarpa. *New Phytologist* **203**: 535–553.

Nagle MF, Yuan J, Kaur D, Ma C, Peremyslova E, Jiang Y, Willig CJ, Goralogia GS, Niño de Rivera A, McEldowney M, *et al.* 2022. GWAS identifies candidate regulators of in planta regeneration in Populus trichocarpa. *bioRxiv*: 2022.06.08.495082.

Nieuwland J, Maughan S, Dewitte W, Scofield S, Sanz L, Murray JAH. **2009**. The D-type cyclin CYCD4;1 modulates lateral root density in Arabidopsis by affecting the basal meristem region. *Proceedings of the National Academy of Sciences* **106**: 22528–22533.

Overpeck JT, Udall B. **2020**. Climate change and the aridification of North America. *Proceedings of the National Academy of Sciences* **117**: 11856–11858.

Overvoorde PJ, Okushima Y, Alonso JM, Chan A, Chang C, Ecker JR, Hughes B, Liu A, Onodera C, Quach H, *et al.* 2005. Functional Genomic Analysis of the AUXIN/INDOLE-3-ACETIC ACID Gene Family Members in Arabidopsis thaliana. *The Plant Cell* **17**: 3282–3300.

Parrotta L, Cresti M, Cai G. **2014**. Accumulation and post-translational modifications of plant tubulins. *Plant Biology* **16**: 521–527.

Qamer Z, Chaudhary MT, Du X, Hinze L, Azhar MT. **2021**. Review of oxidative stress and antioxidative defense mechanisms in Gossypium hirsutum L. in response to extreme abiotic conditions. *Journal of Cotton Research* **4**: 9.

Rasmussen A, Smith TE, Hunt MA. **2009**. Cellular stages of root formation, root system quality and survival of Pinuselliottii var. elliottii × P. caribaea var. hondurensis cuttings in different temperature environments. *New Forests* **38**: 285.

Reyt G, Boudouf S, Boucherez J, Gaymard F, Briat J-F. **2015**. Iron- and Ferritin-Dependent Reactive Oxygen Species Distribution: Impact on Arabidopsis Root System Architecture. *Molecular Plant* **8**: 439–453.

Ribeiro CL, Silva CM, Drost DR, Novaes E, Novaes CRDB, Dervinis C, Kirst M. **2016**. Integration of genetic, genomic and transcriptomic information identifies putative regulators of adventitious root formation in Populus. *BMC Plant Biology* **16**: 66.

Rigal A, Yordanov YS, Perrone I, Karlberg A, Tisserant E, Bellini C, Busov VB, Martin F, Kohler A, Bhalerao R, *et al.* 2012. The *AINTEGUMENTA LIKE1* Homeotic Transcription Factor *PtAIL1* Controls the Formation of Adventitious Root Primordia in Poplar. *Plant Physiology* **160**: 1996–2006.

Sanz L, Dewitte W, Forzani C, Patell F, Nieuwland J, Wen B, Quelhas P, De Jager S, Titmus C, Campilho A, *et al.* 2011. The Arabidopsis D-Type Cyclin CYCD2;1 and the Inhibitor ICK2/KRP2 Modulate Auxin-Induced Lateral Root Formation. *The Plant Cell* **23**: 641–660.

Scheuring D, Löfke C, Krüger F, Kittelmann M, Eisa A, Hughes L, Smith RS, Hawes C, Schumacher K, Kleine-Vehn J. **2016**. Actin-dependent vacuolar occupancy of the cell determines auxin-induced growth repression. *Proceedings of the National Academy of Sciences* **113**: 452–457.

Sharma P, Jha AB, Dubey RS. **2019**. Oxidative Stress and Antioxidative Defense System in Plants Growing under Abiotic Stresses. In: Handbook of Plant and Crop Stress, Fourth Edition. CRC Press.

Sharma R, Sahoo A, Devendran R, Jain M. **2014**. Over-Expression of a Rice Tau Class Glutathione S-Transferase Gene Improves Tolerance to Salinity and Oxidative Stresses in Arabidopsis. *PLOS ONE* **9**: e92900.

Shi H, Kim Y, Guo Y, Stevenson B, Zhu J-K. **2003**. The Arabidopsis SOS5 Locus Encodes a Putative Cell Surface Adhesion Protein and Is Required for Normal Cell Expansion. *The Plant Cell* **15**: 19–32.

Stanton BJ, Haiby K, Gantz C, Espinoza J, Shuren RA. **2019**. The Economics of Rapid Multiplication of Hybrid Poplar Biomass Varieties. *Forests* **10**: 446.

Statello L, Guo C-J, Chen L-L, Huarte M. **2021**. Gene regulation by long non-coding RNAs and its biological functions. *Nature Reviews Molecular Cell Biology* **22**: 96–118.

Štícha V, Macků J, Nuhlíček O. **2016**. Effect of permanent waterlogging on the growth of poplar clones MAX 4, MAX 5 (J-104, J-105) (Populus maximowiczii A. Henry × P. nigra Linnaeus) and evaluation of wood moisture content in different stem parts - Short Communication. *Journal of Forest Science* **62 (2016)**: 186–190.

Su W-L, Liu N, Mei L, Luo J, Zhu Y-J, Liang Z. **2019**. Global Transcriptomic Profile Analysis of Genes Involved in Lignin Biosynthesis and Accumulation Induced by Boron Deficiency in Poplar Roots. *Biomolecules* **9**: 156.

Sun P, Jia H, Zhang Y, Li J, Lu M, Hu J. **2019**. Deciphering Genetic Architecture of Adventitious Root and Related Shoot Traits in Populus Using QTL Mapping and RNA-Seq Data. *International Journal of Molecular Sciences* **20**: 6114.

Sun J, Jiao W, Wang Q, Wang T, Yang H, Jin J, Feng H, Guo J, Feng L, Xu X, *et al.* 2021. Potential habitat and productivity loss of Populus deltoides industrial forest plantations due to global warming. *Forest Ecology and Management* **496**: 119474.

Takatsuka H, Higaki T, Umeda M. **2018**. Actin Reorganization Triggers Rapid Cell Elongation in Roots. *Plant Physiology* **178**: 1130–1141.

Terrile MC, París R, Calderón-Villalobos LIA, Iglesias MJ, Lamattina L, Estelle M, Casalongué CA. **2012**. Nitric oxide influences auxin signaling through S-nitrosylation of the Arabidopsis TRANSPORT INHIBITOR RESPONSE 1 auxin receptor. *The Plant Journal: For Cell and Molecular Biology* **70**: 492–500.

Thorvaldsdóttir H, Robinson JT, Mesirov JP. **2013**. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics* **14**: 178–192.

Towns J, Cockerill T, Dahan M, Foster I, Gaither K, Grimshaw A, Hazlewood V, Lathrop S, Lifka D, Peterson GD, *et al.* 2014. XSEDE: Accelerating Scientific Discovery. *Computing in Science & Engineering* **16**: 62–74.

Trachsel S, Kaeppler SM, Brown KM, Lynch JP. **2011**. Shovelomics: high throughput phenotyping of maize (Zea mays L.) root architecture in the field. *Plant and Soil* **341**: 75–87.

Tran L-SP, Urao T, Qin F, Maruyama K, Kakimoto T, Shinozaki K, Yamaguchi-Shinozaki K. **2007**. Functional analysis of AHK1/ATHK1 and cytokinin receptor histidine kinases in response to abscisic acid, drought, and salt stress in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 20623–20628.

Tuskan GA, Mewalal R, Gunter LE, Palla KJ, Carter K, Jacobson DA, Jones PC, Garcia BJ, Weighill DA, Hyatt PD, *et al.* 2018. Defining the genetic components of callus formation: A GWAS approach. *PLoS ONE* **13**: e0202519.

Vsevolozhskaya OA, Hu F, Zaykin DV. 2019. Detecting Weak Signals by Combining Small P-Values in Genetic Association Studies. *Frontiers in Genetics* **10**: 1051.

Wang F, Longkumer T, Catausan SC, Calumpang CLF, Tarun JA, Cattin-Ortola J, Ishizaki T, Pariasca Tanaka J, Rose T, Wissuwa M, *et al.* 2018. Genome-wide association and gene validation studies for early root vigour to improve direct seeding of rice. *Plant, Cell & Environment* **41**: 2731–2743.

Weighill D, Tschaplinski TJ, Tuskan GA, Jacobson D. 2019. Data Integration in Poplar: 'Omics Layers and Integration Strategies. *Frontiers in Genetics* **10**: 874.

Wiehle M, Eusemann P, Thevs N, Schnittler M. 2009. Root suckering patterns in Populus euphratica (Euphrates poplar, Salicaceae). *Trees* **23**: 991–1001.

Wu A, Allu AD, Garapati P, Siddiqui H, Dortay H, Zanor M-I, Asensi-Fabado MA, Munné-Bosch S, Antonio C, Tohge T, *et al.* 2012. JUNGBRUNNEN1, a Reactive Oxygen Species–Responsive NAC Transcription Factor, Regulates Longevity in Arabidopsis. *The Plant Cell* **24**: 482–506.

Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. 2011. Rare-Variant Association Testing for Sequencing Data with the Sequence Kernel Association Test. *American Journal of Human Genetics* **89**: 82–93.

Xu S-L, Rahman A, Baskin TI, Kieber JJ. 2008. Two Leucine-Rich Repeat Receptor Kinases Mediate Signaling, Linking Cell Wall Biosynthesis and ACC Synthase in Arabidopsis. *The Plant Cell* **20**: 3065–3079.

Xu J, Tian Y-S, Xing X-J, Xu Z-S, Zhu B, Fu X-Y, Peng R-H, Yao Q-H. 2017. Enhancement of phenol stress tolerance in transgenic Arabidopsis plants overexpressing glutathione S-transferase. *Plant Growth Regulation* **82**: 37–45.

Xu X, Ye J, Yang Y, Zhang M, Xu Q, Feng Y, Yuan X, Yu H, Wang Y, Yang Y, *et al.* 2020. Genome-Wide Association Study of Rice Rooting Ability at the Seedling Stage. *Rice* **13**: 59.

Xu Z, York LM, Seethepalli A, Bucciarelli B, Cheng H, Samac DA. 2022. Objective Phenotyping of Root System Architecture Using Image Augmentation and Machine Learning in Alfalfa (Medicago sativa L.). *Plant Phenomics* **2022**.

Xue H, Zhang Q, Wang P, Cao B, Jia C, Cheng B, Shi Y, Guo W-F, Wang Z, Liu Z-X, *et al.* 2022. qPTMplants: an integrative database of quantitative post-translational modifications in plants. *Nucleic Acids Research* **50**: D1491–D1499.

Yasrab R, Atkinson JA, Wells DM, French AP, Pridmore TP, Pound MP. 2019. RootNav 2.0: Deep learning for automatic navigation of complex plant root architectures. *GigaScience* **8**: giz123.

Yates TB, Feng K, Zhang J, Singan V, Jawdy SS, Ranjan P, Abraham PE, Barry K, Lipzen A, Pan C, *et al.* 2021. The Ancient Salicoid Genome Duplication Event: A Platform for

Reconstruction of De Novo Gene Evolution in Populus trichocarpa. *Genome Biology and Evolution* **13**: evab198.

Yi D, Alvim Kamei CL, Cools T, Vanderauwera S, Takahashi N, Okushima Y, Eekhout T, Yoshiyama KO, Larkin J, Van den Daele H, *et al.* 2014. The Arabidopsis SIAMESE-RELATED cyclin-dependent kinase inhibitors SMR5 and SMR7 regulate the DNA damage checkpoint in response to reactive oxygen species. *The Plant Cell* **26**: 296–309.

Yoshida K, Kaothien P, Matsui T, Kawaoka A, Shinmyo A. **2003**. Molecular biology and application of plant peroxidase genes. *Applied Microbiology and Biotechnology* **60**: 665–670.

Zang L, Zheng T, Su X. **2015**. Advances in research of fasciclin-like arabinogalactan proteins (FLAs) in plants. *Plant Omics* **8**: 190–194.

Zhang H, San ML, Jang S-G, Lee J-H, Kim N-E, Lee A-R, Park S-Y, Cao F-Y, Chin J-H, Kwon S-W. **2020**. Genome-Wide Association Study of Root System Development at Seedling Stage in Rice. *Genes* **11**: 1395.

Zhang Z, Zhang X, Lin Z, Wang J, Xu M, Lai J, Yu J, Lin Z. **2018**. The genetic architecture of nodal root number in maize. *The Plant Journal* **93**: 1032–1044.

Zheng H, Bednarek SY, Sanderfoot AA, Alonso J, Ecker JR, Raikhel NV. **2002**. NPSN11 is a cell plate-associated SNARE protein that interacts with the syntaxin KNOLLE. *Plant Physiology* **129**: 530–539.

Zheng Z, Hey S, Jubery T, Liu H, Yang Y, Coffey L, Miao C, Sigmon B, Schnable JC, Hochholdinger F, *et al.* 2020. Shared Genetic Control of Root System Architecture between Zea mays and Sorghum bicolor. *Plant Physiology* **182**: 977–991.

Zhou X, Stephens M. **2012**. Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics* **44**: 821–824.

Zhu J, Geisler M. **2015**. Keeping it all together: auxin-actin crosstalk in plant development. *Journal of Experimental Botany* **66**: 4983–4998.