

# Natural variation in the prolyl 4-hydroxylase gene *PtoP4H9* contributes to perennial stem growth in *Populus*

Liang Xiao <sup>1,2,3,†</sup> Yuanyuan Fang <sup>2,3,†</sup> He Zhang <sup>4,†</sup> Mingyang Quan <sup>2,3</sup> Jiakuan Zhou <sup>2,3</sup>  
Peng Li <sup>2,3</sup> Dan Wang <sup>2,3</sup> Li Ji <sup>5</sup> Pär K. Ingvarsson <sup>6</sup> Harry X. Wu <sup>7</sup> Yousry A. El-Kassaby <sup>8</sup>  
Qingzhang Du <sup>2,3,9</sup> and Deqiang Zhang <sup>1,2,3,\*</sup>

- 1 School of Landscape Architecture, Beijing University of Agriculture, Beijing 102206, China
- 2 National Engineering Research Center of Tree Breeding and Ecological Restoration, College of Biological Sciences and Technology, Beijing Forestry University, Beijing 100083, China
- 3 Key Laboratory of Genetics and Breeding in Forest Trees and Ornamental Plants, Ministry of Education, College of Biological Sciences and Technology, Beijing Forestry University, Beijing 100083, China
- 4 State Key Laboratory of Protein and Plant Gene Research, School of Advanced Agricultural Sciences, Peking University, Beijing 100871, China
- 5 MOE Engineering Research Center of Forestry Biomass Materials and Bioenergy, Beijing Forestry University, Beijing 100083, China
- 6 Linnean Center for Plant Biology, Department of Plant Biology, Swedish University of Agricultural Sciences, Box 7080, SE-750 07 Uppsala, Sweden
- 7 Umeå Plant Science Centre, Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Science, 90183 Umeå, Sweden
- 8 Department of Forest and Conservation Sciences, Faculty of Forestry, The University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada
- 9 State Key Laboratory of Tree Genetics and Breeding, College of Biological Sciences and Technology, Beijing Forestry University, Beijing, 100083, China

\*Author for correspondence: deqiang.zhang@bua.edu.cn, or deqiangzhang@bjfu.edu.cn

†These authors contributed equally to this work.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (<https://academic.oup.com/plcell/pages/General-Instructions>) is: Deqiang Zhang (Deqiang.Zhang@bua.edu.cn, or DeqiangZhang@bjfu.edu.cn)

## Abstract

Perennial trees must maintain stem growth throughout their entire lifespan to progressively increase in size as they age. The overarching question of the molecular mechanisms that govern stem perennial growth in trees remains largely unanswered. Here we deciphered the genetic architecture that underlies perennial growth trajectories using genome-wide association studies (GWAS) for measures of growth traits across years in a natural population of *Populus tomentosa*. By analyzing the stem growth trajectory, we identified *PtoP4H9*, encoding prolyl 4-hydroxylase 9, which is responsible for the natural variation in the growth rate of diameter at breast height (DBH) across years. Quantifying the dynamic genetic contribution of *PtoP4H9* loci to stem growth showed that *PtoP4H9* played a pivotal role in stem growth regulation. Spatiotemporal expression analysis showed that *PtoP4H9* was highly expressed in cambium tissues of poplars of various ages. Overexpression and knockdown of *PtoP4H9* revealed that it altered cell expansion to regulate cell wall modification and mechanical characteristics, thereby promoting stem growth in *Populus*. We showed that natural variation in *PtoP4H9* occurred in a BASIC PENTACYSSTEINE transcription factor *PtoBPC1*-binding promoter element controlling *PtoP4H9* expression. The geographic distribution of *PtoP4H9* allelic variation was consistent with the modes of selection among populations. Altogether, our study provides important genetic insights into dynamic stem growth in *Populus*, and we confirmed *PtoP4H9* as a potential useful marker for breeding or genetic engineering of poplars.

Received June 23, 2023. Accepted July 07, 2023. Advance access publication July 31, 2023

© The Author(s) 2023. Published by Oxford University Press on behalf of American Society of Plant Biologists.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)

Open Access

## IN A NUTSHELL

**Background:** Forest trees are long-lived, woody perennials, making them dramatically different from *Arabidopsis* and other herbaceous annual model plants. Poplar (*Populus*) offers the opportunity to study biological questions pertinent to perennial stem growth. The annual cycle of perennial growth is synchronized with the seasons, based on periodic environmental fluctuations. Thus, perennial growth requires sophisticated regulatory mechanisms to initiate growth in the meristems and facilitate annual growth. While the genetic basis of stem growth has been elucidated, the dynamic genetic architecture of perennial stem growth has remained largely unknown.

**Question:** Which genes are involved in the regulation of stem growth during the long lifespan of trees?

**Findings:** Based on the diameter at breast height (DBH) measured over 8 yr, we modeled the growth trajectories from 303 natural accessions of *Populus tomentosa* and identified a causal gene *PtoP4H9* responsible for natural variation of DBH. The allelic variants in *PtoP4H9* promoter altered the expression of *PtoP4H9* by affecting the binding affinities of the transcription factor PtoBPC1. Genetic analysis showed that *PtoP4H9* increased the modification abundance of O-arabinosides and decreased the mechanical rigidity of the cell wall to promote cell expansion and stem radial growth in *Populus*. More importantly, we found that *PtoP4H9* underwent strong natural selection during the local adaptation of *P. tomentosa*.

**Next steps:** Additional efforts will concentrate on the candidate genes responsible for a particular growth stage such as juvenile to adult phase change and senescence phase. Clarifying candidate gene function will enable us to understand the inter-annual variability of perennial stem growth and help predict forest productivity.

## Introduction

Perennial trees possess an incredible ability to maintain their growth over their entire life cycle over hundreds of years (Bhalerao and Fischer 2017). Perennial growth is represented by an increase in the height and diameter of the stem as the result of the formation of new cells and biomass accumulation (Rathgeber et al. 2016). The current understanding of perennial growth shows that tree stem cells can maintain the capacity for self-renewal during their long-life span, thereby maintaining the dynamic and continuous growth of stems (Johnsson and Fischer 2016; Hill and Hollender 2019). Stem cells can be found in the apical shoot and stem meristems contributing to stem elongation of the primary axes and radial thickening respectively (De Rybel et al. 2016; Lehmann and Hardtke 2016; Smet and De Rybel 2016). Meanwhile, the growth in perennial trees represents the outcome of the seasonal periodicity in the growth process, which is determined by regular seasonal environmental fluctuations associated with the annual temperature and daylength cycles (Vaganov et al. 2006). Perennial stem growth thus requires sophisticated regulatory networks to initiate growth in the meristems and facilitate the annual growth process (Fischer et al. 2019; Li et al. 2022).

Perennial growth is a step-by-step process and growth rate varies over the seasonal cycle depending on exogenous and endogenous signals (Shim et al. 2014). Because of their long lifespans, perennial trees in boreal and temperate zones have evolved a unique trait that allows for the arrest of meristems in a scheduled growth cessation. Thus, these trees display seasonally synchronized growth resulting in an annually changing growth (Paul et al. 2014; Ding and Nilsson 2016;

Singh et al. 2019). Genomic and genetic approaches are yielding increasingly detailed insights into the regulation of diverse developmental events during the seasonal growth cycle in trees (Wang et al. 2011; Singh et al. 2017; Singh et al. 2018; Ding et al. 2021). However, the genetic regulation of stem perennial growth remains poorly understood, such as the changing effect of genes during perennial growth.

With large geographic ranges spanning wide environmental gradients, genetic variation in forest trees has been shaped by local adaptation in natural populations (Evans et al. 2014). Several studies have examined growth trends in natural populations and revealed that stem perennial growth traits showed widespread variation related to the geographic and climatic origin of a population (Reich and Oleksyn 2008; McLane et al. 2011). Also, the dynamic changes that underlie total phenotypic variance of growth are also under tight genetic control (Ye et al. 2015). Genome-wide association studies (GWAS) combined with detailed phenotypic measurements of natural populations are therefore necessary to determine which dynamic genetic mechanisms act to shape stem growth processes and ultimately deepen our understanding of perennial growth patterns in forest trees.

Most previous attempts at the genetic dissection of forest growth have focused on a limited number of growth stages of the stem, usually during the juvenile stage (Fahrenkrog et al. 2017; Resende et al. 2017; Bresadola et al. 2019). The limitation of dynamic phenotypic measurements has made our understanding of the genetic architecture of dynamic growth in perennials lag behind that of herbaceous plants. Research in herbaceous plants showed that plant growth is a dynamic process and that the set timing of measurement greatly influences the outcome of subsequent genetic mapping

(El-Lithy et al. 2004; Bac-Molenaar et al. 2015; Muraya et al. 2017; Zhang et al. 2017; Knoch et al. 2020). Thus, exploring the development processes associated with stem growth could significantly enhance our understanding of the underlying genetic mechanisms in perennial trees and provide important information on what factors determine trunk biomass (Sun et al. 2018).

Including growth time as an additional dimension by taking annual measurements of the growth phenotypes will enable us to detect candidate genes over time and identify their changeable genetic contributions throughout the entire growth process (Feng et al. 2021; Wei et al. 2022). This approach of adding additional dimensions has broadened our understanding of how genetic regulation varies across spatio-temporal gradients (Wu and Jiang 2021). Such knowledge is crucial for ongoing breeding efforts where targeting genetic variation that specifically incorporates dynamic genetic regulation will setup current breeding programs for sustained long-term genetic gain.

*Populus* has a roughly 10-year harvest rotation cycle for woody biomass (Kauter et al. 2003). The highest growth rate was observed between approximately 3 and 10 yr, which influenced the final harvest biomass (Han et al. 2020). We have modeled perennial growth trajectories by assessing longitudinal stem and radial growth using perennial measures taken over 8 yr (3 to 10 yr) in a population consisting of 303 natural accessions of *Populus tomentosa*. We dissected the genetic architecture of perennial growth and identified the genetic contributions of time-sequential genes during the growth process. We were interested in the candidate gene *PtoP4H9* which encodes a Prolyl 4 hydroxylase 9 significantly associated with growth trajectories of diameter at breast height (DBH). Natural variants in the promoter of *PtoP4H9* controlled the expression of *PtoP4H9* which subsequently regulated cell expansion of the stem. Furthermore, transgenic poplar with altered *PtoP4H9* expression exhibited changes in modification and mechanical characteristics of the cell wall without apparent differences in the cell wall component content in the stem xylem. These findings will provide beneficial information for genetic regulation underlying a long-term biological process and facilitate genetic improvement of stem growth in perennial trees.

## Results

### Phenotypic variation of stem perennial growth in 303 *P. tomentosa* accessions

To explore phenotypic variation and the genetic basis of the stem growth trajectories in a natural population of *P. tomentosa*, we measured the tree height (HEI) and DBH in 303 accessions over eight consecutive years (from 3 to 10 yr of age) (Supplemental Data Set 1; Fig. 1, A and C). Both DBH and HEI exhibited strong phenotypic correlations among different growth stages (Supplemental Data Set 2; Supplemental Fig. S1). This allowed us to detect candidate genes associated with the whole developmental process of the traits, instead

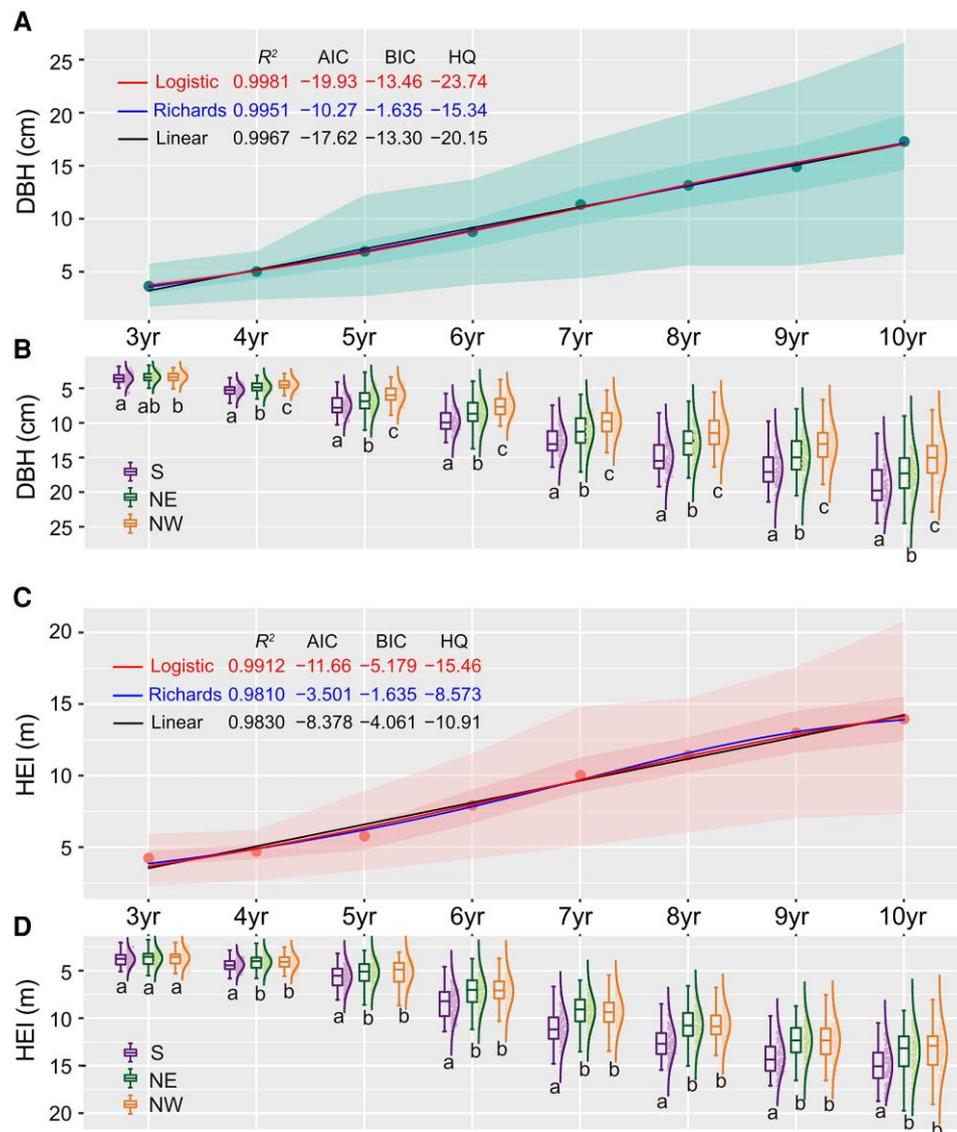
of being associated with any single measurement. Extensive phenotypic variation in stem growth was observed over eight time points in this natural population (Supplemental Fig. S2, A and B). The estimates of broad-sense heritability ( $h^2$ ) for growth traits displayed values greater than 0.5 although heritabilities of growth traits fluctuated with age (Supplemental Data Set 3).

Moreover, the traits showed significant differentiation among three geographic regions (Fig. 1, B and D,  $P < 0.01$ , ANOVA). The DBH and HEI in the southern population (S) were significantly higher than in the northeast (NE) and northwest (NW) populations. Only DBH showed differentiation between NE and NW populations. The distribution of the coefficient of variation in three geographical populations was similar to that observed for the entire population. We found that the phenotypic variation of S population was substantially different with NE and NW populations, which consistent with the phenotypic differentiation across three subpopulations (Supplemental Fig. S2, C and D).

### Growth trajectories of stem perennial growth in *Populus*

Integrating the phenotypic information over multiple years, we can track the growth trajectories of perennial stem in *Populus*. The temporal trends of the DBH and HEI against growth age were smooth and monotonic with visual inspection of the phenotypic trajectories (Fig. 1). Because of seasonal fluctuations and differences in the growth rates, the growth trajectories of perennial stem follow “S” shape approximately. Previous study has shown that nonlinear growth equations have been derived to capture the “S” shape of growth and accurately estimate growth parameters (Han et al. 2020; Gong et al. 2021).

To pick up an optimal equation for our growth data, we first modeled the average growth trajectory over 8 yr using two classical nonlinear growth equations (including Richards and logistic) and compare them to the linear regression model. The logistic equation gives a better goodness-of-fit to average growth trajectories than the other two equations in both DBH and HEI traits by comparing several statistics (Fig. 1, A and C). Consequently, we used logistic model to fit the growth trajectories of 303 accessions. And the growth of both DBH and HEI for each accession showed excellent goodness-of-fit to the logistic model ( $R^2 > 0.9746$ , Supplemental Data Set 4; Supplemental Fig. S3). The observed growth data are distributed randomly over the predicted values (Supplemental Fig. S4), suggesting that the logistic model possesses reasonably good statistical behavior in our data fitting. From the logistic model, we generated a parameter  $b$  to represent the stem growth rate across 8 yr. Because of the phenotypic differentiation of the DBH and HEI, we tested an association between three geographic regions in the parameters of stem growth. The geographic groups were significantly associated with estimates of DBH ( $P = 2.44 \times 10^{-6}$ , Kruskal–Wallis test) and HEI ( $P = 3.65 \times 10^{-4}$ , Kruskal–Wallis



**Figure 1.** Stem growth trajectories of *P. tomentosa* natural population. **A, C**) Observed phenotypic ranges of stem DBH (**A**) and HEI (**C**) of 303 *P. tomentosa* accessions for 3–10 yr. Light color ribbon shows the upper and lower quartiles of annual trait, and very light color ribbon shows traits extreme values (maximum and minimum values). The dots represent the average values of the annual traits. Average growth curves of DBH and HEI traits are fitted with logistic equation, Richards equation, and linear equation. Logistic curves give better goodness-of-fit to average growth trajectories in both DBH and HEI according to the Akaike information criterion (AIC), Bayesian information criterion (BIC), Hannan–Quinn criterion (HQC), and R-squared ( $R^2$ ). **B, D**) Box plots show the phenotypic differentiation of DBH (**B**) and HEI (**D**) during the growing periods across three geographical populations. In box plots, center line represents the median, box limits denote the upper and lower quartiles, whiskers indicate the 1.5 $\times$  interquartile range, and points are outliers. The letters a, b, and c indicate significant differences,  $P < 0.05$  (ANOVA LSD test).

test), indicating that the growth trajectories were distinct differences across three geographical population of *P. tomentosa*.

### Genetic architecture of stem perennial growth

The fitting growth trajectory reduced the dimensionality of the temporal data and represented significant characteristics in the trait development allowing us to detect putative genes at these critical junctures in stem perennial growth (Baison et al. 2019). To identify the genetic determinants of stem

growth, the parameter  $b$  from logistic model, which represents the growth rate, was used as an input for GWAS using EMMAX (Kang et al. 2010) along with 9.6 million single nucleotide polymorphisms (SNPs) in the 303 *P. tomentosa* accessions. This identified 14 significant marker-trait associations (MTAs), including 8 MTAs that conferred the growth rate of DBH and 6 MTAs that contributed to growth rate of HEI; there were no overlapping MTAs between DBH and HEI (Fig. 2A; Supplemental Fig. S5A; Supplemental Data Sets 5 to 7).

We then estimated the linkage disequilibrium (LD) blocks surrounding each significant SNPs and identified a list of 14 candidate genes for which there may be allelic variation contributing to associated phenotypic variation (Supplemental Data Set 5). A number of the candidate genes could participate in functional roles in plant growth and/or secondary growth in *Populus* or other species based on the published literature (Supplemental Data Set 6). For instance, we found a gene surrounding the lead SNP of growth rate of HEI that encodes a MATE Transporter BIG EMBRYO1 (*PtoBIGE1B*) (Supplemental Fig. S5, B and C). A previous study reported that *BIGE1B* regulated the rate of lateral organ initiation in *Arabidopsis* (*Arabidopsis thaliana*) (Suzuki et al. 2015).

To further determine whether the *PtoBIGE1B* is associated with stem perennial growth, we monitored the spatio-temporal expression pattern of *PtoBIGE1B*. *PtoBIGE1B* was only expressed in the cambium and developing xylem tissues over different aged poplar (Supplemental Fig. S5D), supporting the functional role of *PtoBIGE1B* in modulating the perennial growth of the stem. *PtoSAUR36*, encoded a SMALL AUXIN UP RNA36, is associated with growth rate of HEI (Supplemental Data Set 5). *SAUR36* has been shown to modulate adventitious root development through the auxin pathway in *Populus* (Liu et al. 2022a). Except these genes, we also identified several candidate genes regulating stem growth. For example, the first ranking signal of growth rate of DBH is located in the *PtoP4H9*, encoding a Prolyl 4 hydroxylase 9 (Fig. 2A). *PtoP4H9* is a member of P4Hs family involving in cell wall modification in *Arabidopsis* root and leaves. We hypothesized that *PtoP4H9* was a key gene regulating stem growth in perennial *Populus*.

Based on growth trajectories analysis, we tested the association between SNPs and the growth rate. Treating the stem perennial growth as a functional curve increased the statistical power for identifying MTAs. The GWAS of estimations from function curve focused on SNPs associated with the whole growth trajectories of the trait which increased the statistical power for identifying MTAs. But it neglected the genetic contributions of SNPs at a particular growth stage (Li and Sillanpää 2015). Based on the analysis of time-point traits separately, we traced the phenotypic variation explained (PVE) for 8 yr to quantify the dynamic genetic contribution of MTAs to stem growth. Only two significant SNPs were found to be in the pivotal dominant position of regulation in the stem growth with high PVE value in time-point traits (PVE > 5%) (Supplemental Fig. S6). *PtoBIGE1B* locus explained 6.11% to 7.48% of the phenotypic variation of HEI across 7 to 9 yr in the population (Supplemental Fig. S5, E to G), while *PtoP4H9* explained the high phenotypic variations of DBH for seven continuous years (Supplemental Fig. S6A).

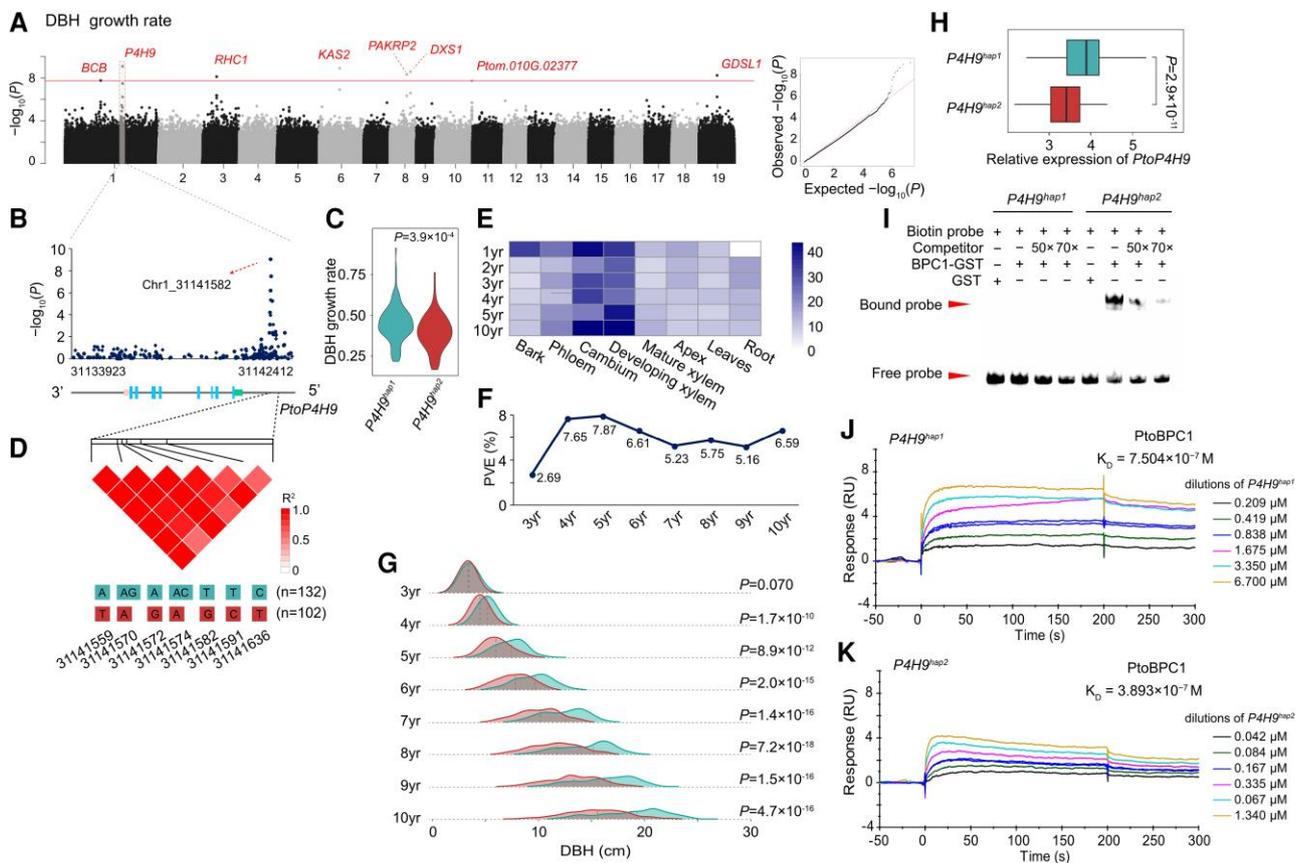
**Natural variations in *PtoP4H9* confer perennial growth of DBH by altering the expression of *PtoP4H9***  
We have identified an array of candidate genes that are associated with stem growth traits in *Populus*. Here we confined

our attention to *PtoP4H9* which is the first ranking signal of growth rate of DBH (Fig. 2, A to D) and gives high PVE values for DBH over the growth period (Supplemental Fig. S6B). The peak SNP, Chr1\_31141582 (G/T), is located in the promoter of *PtoP4H9* (Fig. 2B). Analysis of the spatiotemporal expression pattern of *PtoP4H9* showed that expression was relatively high in cambium and developing xylem tissues over all the growth stages (Fig. 2E). We then tested the association between *PtoP4H9* loci with time-point traits independently, which found that *PtoP4H9* locus contributes the variation of DBH across seven consecutive years (from 4 to 10 yr) (Fig. 2, F and G; Supplemental Fig. S7). This suggested that *PtoP4H9* is a key and consistent genetic factor contributing to stem radial growth.

Incorporating several insertions and deletions (InDels) with the SNPs in the genomic region of *PtoP4H9* highlighted a major haplotype (hap) block in the *PtoP4H9* promoter that appeared to be responsible for phenotypic variations in DBH across perennial growth (Fig. 2D). The haplotype block consisted of five SNPs and two InDels, which subdivided the association mapping population into two haplotypes, named *PtoP4H9*<sup>hap1</sup> and *PtoP4H9*<sup>hap2</sup> (Fig. 2D). The 132 accessions that carry the *PtoP4H9*<sup>hap1</sup> haplotype are associated with a faster growth rate and bigger DBH. Conversely, the 102 accessions carrying the *PtoP4H9*<sup>hap2</sup> haplotype are associated with a slower growth rate and smaller DBH. Sequence analysis showed that the haplotype block in the *PtoP4H9*<sup>hap2</sup> promoter contained the binding sequence, GAGA, for the transcription factor BASIC PENTACYSSTEINE1 (*PtoBPC1*), while *PtoP4H9*<sup>hap1</sup> haplotype lacked this sequence feature. We therefore hypothesized that this haplotype could be responsible for the differences in *PtoP4H9* expression levels between the two different alleles.

RT-qPCR analysis of the cambium from the 234 accessions revealed that the expression level of *PtoP4H9* was significantly different between *PtoP4H9*<sup>hap1</sup> and *PtoP4H9*<sup>hap2</sup> (Fig. 2H), suggesting that relative *PtoP4H9* expression is lower in *PtoP4H9*<sup>hap2</sup> harboring the GAGA promoter element. Accessions that carried the *PtoP4H9*<sup>hap1</sup> and *PtoP4H9*<sup>hap2</sup> exhibited similar *PtoP4H9* temporal expression patterns, but the expression levels of *PtoP4H9*<sup>hap1</sup> were higher than that of *PtoP4H9*<sup>hap2</sup> (Supplemental Fig. S8). The temporal expression pattern of *PtoBPC1* demonstrated that *PtoP4H9* was negatively co-expressed with *PtoBPC1* in cambium tissues, which confirmed the negative role of *PtoBPC1* in regulating the expression of *PtoP4H9* (Supplemental Fig. S9).

Next, to verify whether *PtoBPC1* regulates *PtoP4H9* expression, we investigated the ability of *PtoBPC1* to bind to the different *PtoP4H9* alleles. Electrophoretic mobility shift assays (EMSA) showed that *PtoBPC1* preferentially binds to *PtoP4H9*<sup>hap2</sup> rather than *PtoP4H9*<sup>hap1</sup> (Fig. 2I). Surface plasmon resonance (SPR) also supported that *PtoBPC1* had low affinity with *PtoP4H9*<sup>hap1</sup> compared to *PtoP4H9*<sup>hap2</sup> (Fig. 2, J and K), suggesting that the allelic variants in the haplotype regulated *PtoP4H9* expression by affecting the binding of the transcription factor *PtoBPC1*.



**Figure 2.** Natural variation in *PtoP4H9* promoter contributes to DBH growth. **A)** Manhattan plot shows GWAS results of DBH growth rate ( $b$ , parameter derived from logistic model). The line marks the GWAS threshold ( $P = 1.84 \times 10^{-8}$ ). The candidate genes are marked with words. Quantile-quantile plot for DBH growth rate is shown in the right panel. The full gene names are shown in [Supplemental Data Set 6](#). **B)** Local Manhattan plot of *PtoP4H9* genomic regions on chromosome 1. Diamonds denote InDels and dots represent SNPs. *PtoP4H9* gene model is shown at the bottom of the Manhattan plot. **C)** Comparison of growth rate of DBH group by two alleles of *PtoP4H9*. **D)** Linkage disequilibrium heat map (top) and *PtoP4H9* haplotypes (bottom) among *P. tomentosa* natural variation.  $R^2$  represents the LD R-squad between two SNPs. **E)** Expression patterns from eight tissues (bark, phloem, cambium, developing xylem, mature xylem, apex, leaves, and root) for different ages (1 to 5 and 10 yr old) are shown in the heatmap. **F)** Line chart shows the PVE by signal SNP Chr1\_31141582. **G)** Ridge map shows the genotypic effect of *PtoP4H9* haplotypes contributing to DBH at 8 yr. **H)** Comparison of *PtoP4H9* expression between *PtoP4H9*<sup>hap1</sup> and *PtoP4H9*<sup>hap2</sup>. Gene expression level is determined among the cambium tissue of 234 *P. tomentosa* accessions. In box plots, center line represents the median, box limits denote the upper and lower quartiles, whiskers indicate the 1.5 $\times$  interquartile range, and points are outliers. **I)** Gel-shift analysis of PtoBPC1 binding to *PtoP4H9*<sup>hap1</sup> and *PtoP4H9*<sup>hap2</sup> promoters. EMSAs of a Glutathione S-transferase-labeled probe with PtoBPC1 protein. DNA probes representing *PtoP4H9*<sup>hap1</sup> and *PtoP4H9*<sup>hap2</sup> promoters (probe sequences are shown in [Supplemental Table S1](#)). Protein/DNA complexes are indicated by arrows. **J–K)** SPR binding profiles of PtoBPC1 proteins to promoter of *PtoP4H9*<sup>hap1</sup> (**J**) and *PtoP4H9*<sup>hap2</sup> (**K**). RU, response units; KD, dissociation constant. The significance of differences in **C**, **G**, and **H** is derived with a two-tailed  $t$  test.

### *PtoP4H9* regulates stem growth via promoting cell expansion

To investigate the morphological variation between different *PtoP4H9* genotypes in more detail, we performed paraffin sectioning which demonstrated the stem xylem fiber cells were larger in *PtoP4H9*<sup>hap1</sup> than in *PtoP4H9*<sup>hap2</sup> ([Supplemental Fig. S10](#)). To further validate this, we generated transgenic poplar lines overexpressing *PtoP4H9* (designated *P4H9-OE*) and knocking down *PtoP4H9* expression (designated *p4h9-KD*). The stem diameter and height were significantly reduced in all six independent *p4h9-KD* lines and dramatically increased in all six independent *P4H9-OE* lines compared to the wild type (WT) ([Fig. 3, A to D](#); [Supplemental Fig. S11, A to D](#)).

Stem growth was driven by cell expansion and/or cell proliferation and therefore we examined the cell layers and cell sizes in the stem xylem. The size of xylem cells was significantly altered in the transgenic individuals while the layers remained relatively constant ([Fig. 3, E to G](#); [Supplemental Fig. S11, E to G](#)). These results indicated that variation in stem diameter that is caused by expression level differences in *PtoP4H9* is mainly driven by changes in cell size.

### *PtoP4H9* has putative effect on the cell wall modification

*PtoP4H9* encoded a Prolyl 4 hydroxylases (P4Hs) that catalyzed proline hydroxylation of cell wall hydroxyproline-rich

glycoproteins (HRGPs). A phylogenetic tree of the P4Hs family was generated using MEGA7 and the results indicated P4Hs members showed the conserved syntenic relationships between *A. thaliana* and *P. tomentosa* (Supplemental Fig. S12). We found that most of the members (7 out of 13) showed high expression level in the cambium using the Aspwood database from *Populus tremula* (Supplemental Fig. S13; Sundell et al. 2017). On the basis of previous studies of P4Hs in Arabidopsis, we proposed that *PtoP4H9* was likely to be involved in post-translational modifications of HRGPs.

We found that a *PtoP4H9*-GFP fusion protein co-localized with Phosphate Transporter PHT4; 6-mRFP (a Golgi-localized protein) in *Nicotiana tabacum* leaf cells (Fig. 4A; Hassler et al. 2012), indicating *PtoP4H9* was located in the Golgi and consistent with the role of *PtoP4H9* in the conversion of proline (Pro) to hydroxyproline (Hyp) in HRGPs. We next detected the Hyp content in fiber cell walls in the stem and found that the Hyp content was significantly increased in *P4H9-OE* stem and significantly decreased in *p4h9-KD* compared to the WT (Fig. 4B; Supplemental Fig. S11H), which was consistent with the expression of *PtoP4H9*. The same conclusion was also supported by a co-expression analysis. We identified 530 genes that were co-expressed with *PtoP4H9* in cambium tissues from differently aged *P. tomentosa*. Gene ontology (GO) analysis revealed that these 530 genes were significantly enriched in protein N-linked glycosylation process (Supplemental Data Set 8; Supplemental Fig. S14). These results implied that *PtoP4H9* was involved in the cell wall modification to modulate cell wall assembly.

The glycosidic linkage analysis provided information on average composition rather than a specific sequence and/or distribution of the extension (EXT) oligoarabinosides that could vary between hyp-arabinosides (Velasquez et al. 2011). Therefore, we analyzed the detailed EXT glycopeptides (Hyp-O-glycans) from the xylem of *PtoP4H9* transgenic plants using MALDI mass spectrometer (MALDI-MS). The *p4h9-KD* plants exhibited decreased modifications of hyp-arabinose chains [Hyp-(Ara)<sub>n</sub>] compared to WT, while *P4H9-OE* plants displayed increased modifications of hyp-arabinose chains [Hyp-(Ara)<sub>n</sub>] (Fig. 4C; Supplemental Fig. S11I). We reasoned that alteration of *PtoP4H9* expression destabilized the modification of O-arabinosides on EXTs affecting cell wall structure. Consequently, we investigated changes in surface structures and mechanical characteristics of the xylem cell wall induced by *PtoP4H9* using atomic force microscopy (AFM).

In accordance with the transmission electron microscopy (TEM) results (Fig. 4D), the 3D contour mapped by the AFM revealed that the thickness of fiber cell walls in the stem xylem did not change significantly in *PtoP4H9* transgenic poplars (Fig. 4E; Supplemental Fig. S11J). However, the elastic modulus of *P4H9-OE* stem xylem cell walls was significantly lower than in the WT while *p4h9-KD* was significantly higher than WT (Fig. 4F; Supplemental Fig. S11K). The major components of the secondary cell wall include lignin, cellulose, and hemicellulose, and these were all present in

similar level between *P4H9-OE* and *p4h9-KD* plants as compared to WT (Supplemental Fig. S15). Collectively, the data showed that *PtoP4H9* was responsible for cell expansion by destabilizing the hyp modification of O-arabinosides to alter the wall rigidity of the stem xylem fiber cell wall without any obvious effect on cell wall chemical composition.

### *PtoP4H9* is a target of natural selection during local adaptation

Dynamic stem growth, as an adaptive trait, was related to the geographic and climatic origin of the populations as we found that the dynamic traits showed significant differences between three geographic regions (Fig. 1). Given the important role of *PtoP4H9* in affecting dynamic traits, it was also conceivable that *PtoP4H9* played an important role in explaining natural variation in DBH. To assess this, we scanned genomic selective sweep region across three geographic populations that evaluated by ratios of nucleotide diversity ( $\pi$ ), population-differentiation statistic (*F<sub>st</sub>*), selection statistic (Tajima's *D*), and composite likelihood ratio (CLR) (Xiao et al. 2021).

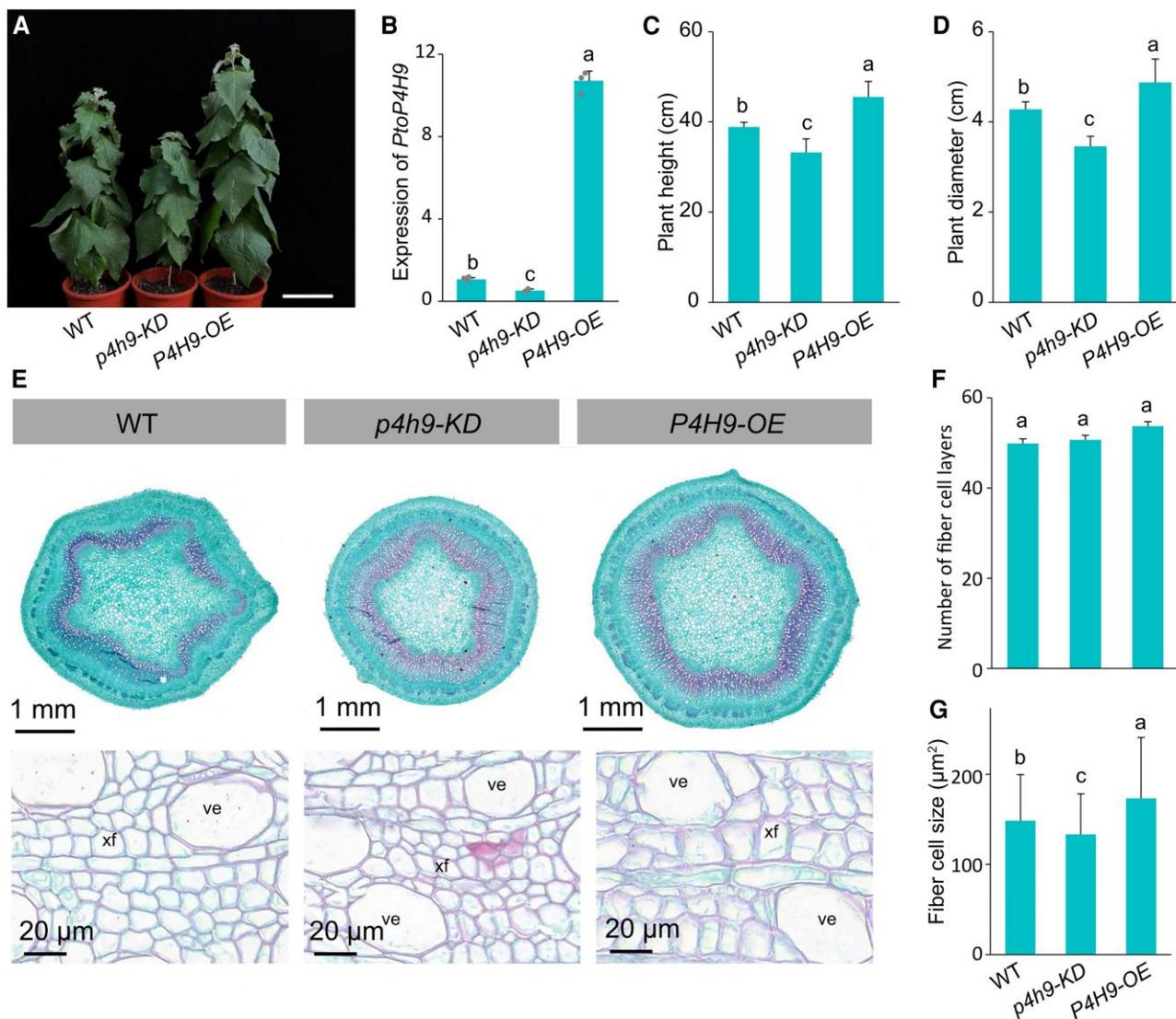
We found that *PtoP4H9* was located within selective sweep region on Chr1 (Fig. 5, A to D). We thus postulated that *PtoP4H9* may be a target of natural selection during local adaptation. To verify this, we performed an analysis of the geographic distribution of two haplotypes of *PtoP4H9* in natural population of *P. tomentosa*. We determined the *PtoP4H9*<sup>hap1</sup> and *PtoP4H9*<sup>hap2</sup> alleles in 231 *P. tomentosa* accessions. The frequency of the *PtoP4H9*<sup>hap1</sup> and *PtoP4H9*<sup>hap2</sup> alleles in the accessions from the southern region was 96.70% and 3.30% respectively (Fig. 5E). Meanwhile, the frequency of the *PtoP4H9*<sup>hap2</sup> allele gradually increased in the northeast (46.88%) and northwest regions (93.42%), which consistent with the phenotypic differentiation of stem trunks. This may be a signature of adaptation to local climates.

There was an obvious climatic gradient across the distribution area of *P. tomentosa* (Supplemental Fig. S16). The northwest and northeast regions characterized by low temperatures and limited annual rainfall compared to areas in the southern region. Among those regions, the climatic conditions were more severe in north regions. The expression levels of *PtoP4H9* also decreased in the two northern regions compared to the south region (Fig. 5F), suggesting that long-term natural selection in response to local temperature and moisture regimes have resulted in reduced stem diameter growth rate through the altered expression of *PtoP4H9*. Collectively, our results showed that the promoter alleles of *PtoP4H9* were associated with stem growth and have been the target of natural selection in *P. tomentosa*.

## Discussion

### Dynamic traits facilitate the dissection of the stem perennial growth process in *Populus*

Many aspects of tree biology are common to annual herbs but there are some unique facets of tree anatomy and

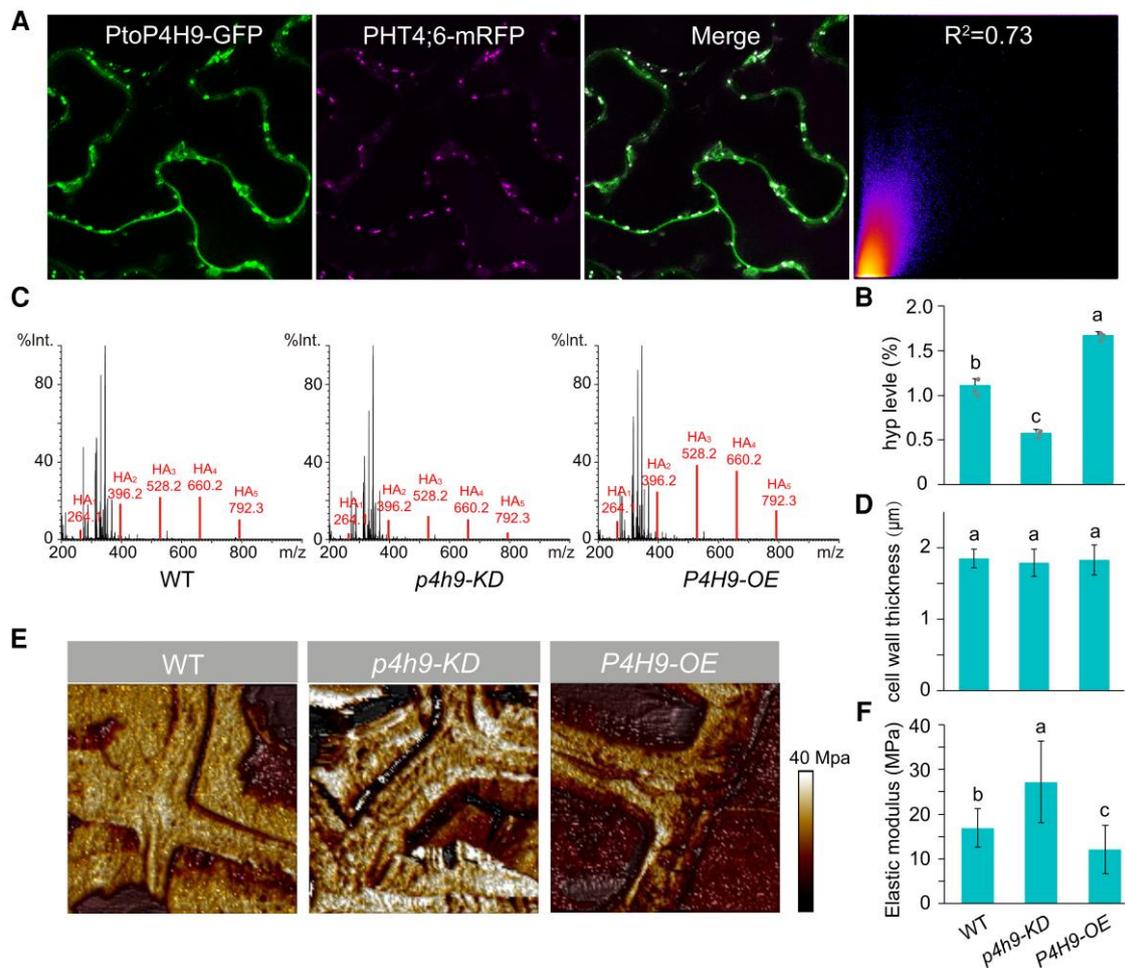


**Figure 3.** *PtoP4H9* promote cell expansion to improve stem radial growth. **A)** Phenotype of representative greenhouse-grown individuals 6 wk after transfer to soil. Scale bar = 10 cm. WT; *p4h9-KD*: *PtoP4H9* knocked down line; *P4H9-OE*: *PtoP4H9* overexpressed line. **B)** Expression of *PtoP4H9* in transgenic *Populus*. Data are means  $\pm$  SD from three technical replicates. The points represented the data of three replicates. **C–D)** Quantitative measurement of plant height (**C**) and diameter (**D**). Data are means  $\pm$  SD from 12 individual plants. **E)** Stem cross section of transgenic poplar stained with safranin O/fast green staining. Scale bars represent 1 mm (upper panels) and 20  $\mu\text{m}$  (lower panels). **F–G)** Quantitative measurement of xylem cell layers (**F**) and sizes (**G**). Data are means  $\pm$  SD from individual cells of three individual plants. The significance of difference in (**C**), (**D**), and (**G**) is derived with two-tailed *t* test ( $*P < 0.01$ ,  $n = 3$  for **B**,  $n = 12$  for **C** and **D**,  $n = 10$  for **F**, and  $n > 100$  for **G**). xf = xylem fiber; ve = vessel element. The phenotype analysis of the representative line of *p4h9-KD* (#1) and *P4H9-OE* (#10) is from six independent transgenic lines respectively. Results from other independent lines are shown in Supplemental Fig. S11A to C.

physiology that must specifically be investigated in trees (Bradshaw et al. 2000). Perennial stem growth is perhaps the most obvious unique character. The repeated cycling between active and dormant meristems makes annual stem growth capacity closely related to seasonal environmental fluctuations in the perennial trees (Rohde and Bhalerao 2007). Stem growth in perennial *Populus* depends on both radial and longitudinal growth which can be measured by two unambiguously quantitative traits, DBH and HEI. Annual measurements of DBH and HEI therefore supply important

information on how stem growth is affected by environmental variation and especially climate variability.

As a dynamic trait, stem growth undergoes consecutive annual growth process. However, the genetic dissection of stem growth has predominantly been performed using a single data point, representing the value of the trait at a given developmental stage. Using a single time-point overlooks many of the factors that define the process of formation and development for many important traits. It is worth noting that the functional importance of most genes was not

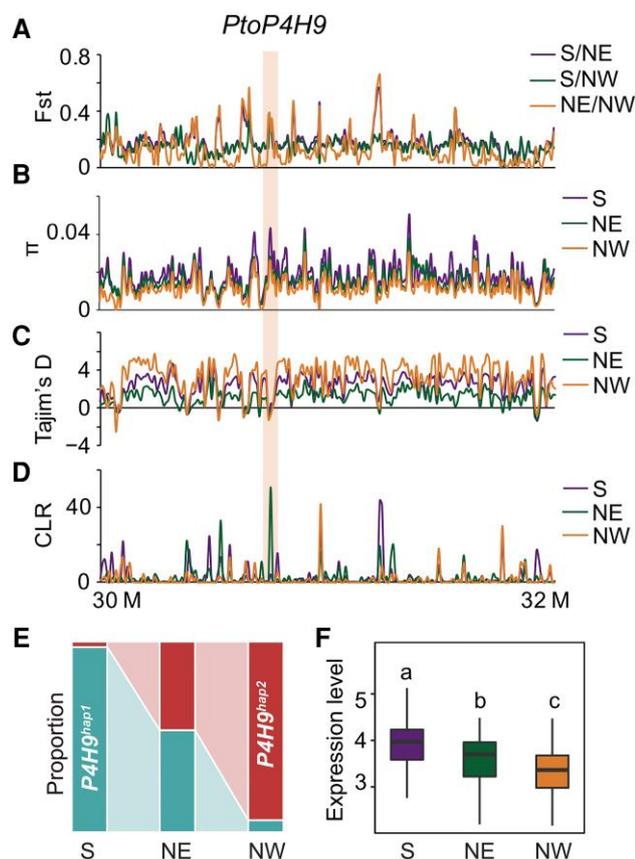


**Figure 4.** *PtoP4H9* has altered the cell wall architecture. **A**) Co-localization of *PtoP4H9*-GFP with *PHT4*;6-mRFP transiently expressed in *N. tabacum* leaf epidermal cells. Scatter plot shows correlation of the GFP and mRFP fluorescence intensity with an  $R^2$  value (Pearson correlation coefficient) of 0.73; bar, 50  $\mu\text{m}$ . **B**) Chemical determinations of hydroxyproline (hyp) level. Hyp level is expressed as percentage of Hyp per mg of DW xylems. Data are means  $\pm$  SD from three biological replicates. **C**) Mass-spectra of EXT O-glycans (Hyp-arabinosides, HA1-HA5) analyzed with MALDI-MS in transgenic *Populus*. Hyp-arabinosides are detected as follows: HA1, Hyp-(Ara)<sub>1</sub>; HA2, Hyp-(Ara)<sub>2</sub>; HA3, Hyp-(Ara)<sub>3</sub>; HA4, Hyp-(Ara)<sub>4</sub>; and HA5, Hyp-(Ara)<sub>5</sub>. Ions important for the interpretation of the results are highlighted. %int, relative intensity (%). **D**) Quantitative measurement of xylem cells wall thickness using TEM images. Data are means  $\pm$  SD from individual cells of three individual plants. **E**) AFM mapping of three-dimensional topography overlaid with elastic modulus of stem xylem cells. Colors indicate elasticity. **F**) Quantification of apparent Young's modulus using the Peak Force QNM mode. Each measurement represents the average of value of consecutively indented three times as the Young's modulus of each position. Data are means  $\pm$  SD of 50 positions from three individual plants. WT; *p4h9*-KD: *PtoP4H9* knocked down line; *P4H9*-OE: *PtoP4H9* overexpressed line. The letters a, b, and c indicate significant differences.  $P < 0.01$  (ANOVA LSD test,  $n = 3$  for **B**,  $n > 30$  for **D**,  $n = 50$  for **F**). The phenotype analysis of the representative line of *p4h9*-KD (#1) and *P4H9*-OE (#10) is from three independent transgenic lines respectively. Results from other independent lines are shown in Supplemental Fig. S111 to K.

equal across all growth stages, as gene expression data showed a number of distinct temporal-spatial expression patterns (Bac-Molenaar et al. 2015). In this study, we modeled the growth trajectories of each accession using a non-linear growth logistic model, which aimed at dissecting the genetic basis of such dynamics in stem perennial growth that account for year-to-year variation across annual growth. We identified 14 MTAs that were involved in shaping dynamic growth by genetic dissection of the stem growth rate (Supplemental Data Set 5).

Several studies have suggested that integrating the time dimension of stem growth should remarkably improve the

statistical power of dissecting the genetic architecture and facilitate biological interpretations in trees (Campbell et al. 2015). For instance, functional mapping has been implemented to identify candidate genes affecting the dynamic stem growth in a full-sib population of 64 progeny derived from two *Populus* species (Jiang et al. 2016; Xu et al. 2016). The method modeled growth trends have integrated information over multiple time points and detected candidate genes responsible for the whole developmental process, instead of being contributed to any single observation (Li and Sillanpää 2015). Moreover, we also showed that the major candidate gene responsible for DBH growth rate, *PtoP4H9*,



**Figure 5.** Distribution of natural variation in *PtoP4H9* during local adaptation of *P. tomentosa*. **A–D**) Putative sweep regions around *PtoP4H9* locus were evaluated by different summary statistics. The bar denotes the location of the *PtoP4H9* gene. The statistics were calculated separately for accessions from three geographic populations: southern, S; northeastern, NE; northwestern, NW. **A**) Genetic differentiation ( $F_{st}$ ), **B**) nucleotide diversity ( $\pi$ ), **C**) Tajima's D, and **D**) CLR test for the presence of a selective sweep. **E**) Frequency distribution of two haplotypes of *PtoP4H9* in three geographic populations. **F**) Relative expression level of *PtoP4H9* in the three geographic populations. In box plots, center line represents the median, box limits denote the upper and lower quartiles, whiskers indicate the 1.5 $\times$  interquartile range, and points are outliers. The letters a, b, and c indicate significant differences.  $P < 0.001$  (ANOVA LSD test).

played a pivotal role in regulation of single time-point traits and explained more than 5% of the phenotypic variation of DBH across seven consecutive years. These results highlighted that integrating time-dimensional models facilitated the genetic dissection of complex growth traits and provided an attempt to unveil the genetic mechanisms that controlled the trait formation.

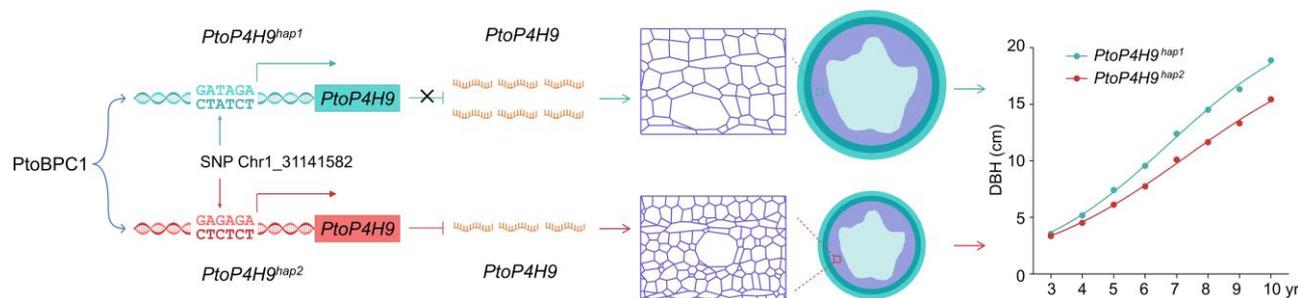
Collectively, our study provided a comprehensive understanding of the genetic architectures of stem perennial growth in a natural population and could enable detection of candidate genes responsible for whole growth trajectory and possible dynamic genetic contributions of candidate genes across growth stages. By adding a time dimension, dynamic data contain more information on the genetic

mechanisms controlling the growth process, which will help elucidate the genetic regulatory pattern of plant growth and alleviate many of the limitations seen in previous studies.

### *PtoBPC1-P4H9* module involved in the cell wall modification to regulate cell size and stem growth in *Populus*

Stem growth is one of the most important traits in perennial trees and molecular genetics studies have characterized a series of candidate genes that are involved in regulating stem growth. In this study, we focused on *PtoP4H9* which is responsible for the growth trajectories of DBH traits in *Populus*. *PtoP4H9* is a member of the prolyl 4-hydroxylases (P4Hs) that is localized to the Golgi (Fig. 4A). P4Hs catalyze proline hydroxylation that is an early post-translational modification of HRGPs, such as EXTENSINS (EXTs), PROLINE RICH PROTEINS, and ARABINO GALACTAN PROTEINS, in the cell wall (Kieliszewski and Lampert 1994; Xu et al. 2008; Showalter et al. 2010; Velasquez et al. 2011). Several members of the P4Hs family are capable of regulating cell expansion and plant growth (Fragkostefanakis et al. 2014; Velasquez et al. 2015). These studies provide clues to how P4Hs regulate cell expansion in roots and leaves but their roles in stem growth of *Populus* are still largely unknown. Here, we provide reliable evidence, including morphological analysis of natural accessions and transgenic poplars (Supplemental Fig. S10; Fig. 3E), supporting the crucial role of *PtoP4H9* in modulating xylem cell size and stem growth in *Populus*.

HRGPs are the major class of cell wall protein that occur in the extracellular space or that are linked to the outer membrane of plant cells (Marzol et al. 2018). Extensive studies have shown that correct O-glycosylation of HRGPs is essential for cell wall assembly. Suppression of *CrP4H1* in *Chlamydomonas reinhardtii* results in improper cell wall assembly and the inability of cells to regenerate their walls (Keskiaho et al. 2007). *AtP4H5* trigger angle differences of the oriented fibrils in the root cell wall, which links HRGPs to changes in the cell wall architecture at the root hair tip in *Arabidopsis* (Velasquez et al. 2015). A phylogenetic tree reveals that the P4Hs show strong syntenic relationships between *Arabidopsis* and *Populus*, suggesting conserved or similar functions in these two species (Supplemental Fig. S12). Tissue expression pattern and co-expression analysis were also performed to show that *PtoP4H9* is involved in cell wall assembly of the cambium in *Populus* (Supplemental Figs. S13 and S14). The modification of EXT glycopeptides and mechanical characteristics of the xylem cell wall were significantly altered in the *PtoP4H9* transgenic *Populus* (Fig. 4, C, E, and F). It is therefore plausible that *PtoP4H9* may improve stem growth by regulating cell wall assembly and cell expansion during perennial growth of *Populus*. Additionally, the thickness and chemical components content of the secondary cell wall was unchanged even when the expression of *PtoP4H9* was altered (Fig. 4D; Supplemental Fig. S15), indicating that *PtoP4H9* confers a



**Figure 6.** A proposed model of *PtoP4H9* regulates stem growth in *Populus*. The natural variation in *PtoP4H9* promoter leads to different binding affinities for PtoBPC1 transcription factor. The *PtoP4H9*<sup>hap2</sup> promoter containing the GAGA motif is bound by PtoBPC1, while *PtoP4H9*<sup>hap1</sup> haplotype lacked this sequence feature. This sequence variation cause PtoBPC1 binds to the *PtoP4H9* promoter with lower affinity in *PtoP4H9*<sup>hap1</sup> than in *PtoP4H9*<sup>hap2</sup> accessions, leading to differences in *PtoP4H9* expression. Thus, *PtoP4H9*<sup>hap1</sup> accessions show larger cell size and faster stem radial growth rate than *PtoP4H9*<sup>hap2</sup> accessions. The rectangle is the fiber cells in stem xylem from different allele of *PtoP4H9*. The scatter plot shows the average values of the eight time-point of DBH in accessions carried *PtoP4H9*<sup>hap1</sup> and *PtoP4H9*<sup>hap2</sup> respectively, and the curves represent the growth trajectories of DBH fitted by logistic model.

growth advantage under perennial growth without any obvious negative effects on wood quality.

Transcriptional regulation is an important mechanism for modulating plant growth. We show that the transcript level of *PtoP4H9* is substantially altered by the transcription factor PtoBPC1 which binds directly to the *PtoP4H9* promoter (Fig. 2, I to K). BPC proteins compose a plant-specific transcription factor family that contains a conserved DNA-binding domain that directly binds to GA-rich sequences in gene promoters (Kooiker et al. 2005). Previous study has revealed that BPC1 regulated the expression levels of several *HOMEODOMAIN* genes to control meristem activity (Simonini and Kater 2014). BPC1 also functions in cell wall biosynthesis by inhibiting the expression of *GALACTAN SYNTHASE1* (*GALS1*) in *Arabidopsis* (Yan et al. 2021). Here, PtoBPC1 repressed *PtoP4H9* expression and the expansion of cells in the xylem. Based on these results, we propose a regulatory mechanism that PtoBPC1-P4H9 module regulated cell expansion to improve stem growth in *Populus* (Fig. 6).

### Natural variation of *PtoP4H9* increased the efficiency of molecular breeding in perennial tree species

Improving stem growth has remained a major challenge in *Populus* breeding programs because of its complex genetic basis and the large genotype-by-environment effects that are commonly affecting this trait. Employing a GWAS strategy and taking a dynamic perspective of growth traits have allowed us to uncover the genetic architectures of growth in *Populus*. We have been able to quantify the genetic regulatory mechanisms of growth trajectories that determine pivotal candidate gene in regulation of stem perennial growth.

Our study has demonstrated that *PtoP4H9* is a major gene responsible for growth rate of whole growth trajectory of DBH (Fig. 2). In particular, we observe that a haplotype block located in the promoter of *PtoP4H9* is significantly associated with perennial growth of DBH (Fig. 2D). We also show that this haplotype as a core-promoter element corresponding

to the binding site of PtoBPC1 which directly alter *PtoP4H9* expression (Fig. 2H). *PtoP4H9* is continuously expressed at high level in the cambium and developing xylem in differently aged poplar (Fig. 2E). However, morphological comparisons reveal significant differences in stem xylem fiber cells size between the two haplotypes at *PtoP4H9*, *PtoP4H9*<sup>hap1</sup>, and *PtoP4H9*<sup>hap2</sup> (Supplemental Fig. S10). These conclusions are consistent with the phenotype observed in transgenic poplar lines, demonstrating that *PtoP4H9* is a major factor affecting stem growth in *Populus*.

Growth trajectories can differ during stem development and between populations, and such diversity of growth trajectories can reflect local adaptation of intraspecific genetic variation. In species of *Populus* that have large geographic ranges spanning wide environmental gradients, stem growth traits could conceivably be linked to local climate regimes through natural selection creating local adaptation. Indeed, stem growth shows a high degree of adaptive phenotypic variation across natural population of *P. tomentosa*. Previous study has demonstrated that local climate is often a fundamental driver of phenotypic variation of growth in forest trees (Evans et al. 2014). And we find that phenotypic differentiation among populations of *P. tomentosa* is consistent with a response to climate related variation (Fig. 1).

The geographic distribution of natural variants in *PtoP4H9* is also consistent with a strong influence of natural selection among populations (Fig. 5E). Low temperatures and limited annual rainfall in the northwest and northeast regions have favored smaller stature stems and there is a concomitant reduction in *PtoP4H9* expression in these populations, which is achieved by natural variation in the promoter of *PtoP4H9*. The importance of temperature and precipitation as critical environmental factors driving local adaptation in dynamic stem growth has also been established in *Pinus sylvestris*, which highlights the growth-limiting effects of drought and chilling (Peltola et al. 2002; McLane et al. 2011). Local adaptation has evolved as a response to climate by altering the frequency or by fixing of different alleles associated with lower

trunk biomass, resulting in adaptive population differentiation (O'Neill et al. 2008; Pautasso 2009). Collectively, GWAS and evolutionary effects of natural variation in *PtoP4H9* suggested that it is a potentially useful gene for marker-aided selection or transgenic modification, if beneficial results within breeding populations can be demonstrated in future studies.

Our study illustrates a blueprint for the analysis of the genetic control of the perennial growth process, providing valuable information on the molecular underpinnings of variation in dynamic genetic contributions and gene expression during stem growth in *Populus*. We have focused on the major factor *PtoP4H9* and demonstrated how it can be utilized for future molecular breeding aimed at improving stem growth. However, we emphasized that tree growth neither responded equally to inter-annual climate variation nor will predicted climate change affect all annual growth cycles in a similar manner. More attention should be given to the effect of “local” QTLs responsible for a particular growth stage in future studies and especially genes associated with the residual variances traits that represent the nonlinear growth of perennial stem. These are all important consequences for the understanding inter-annual variability of perennial stem growth which have substantial contributions for predicting future forest productivity.

## Materials and methods

### Germplasm population and growth conditions

The *P. tomentosa* germplasm population used in this study is composed of 303 accessions that were asexually propagated via root segments and planted in a randomized complete block design with three blocks in 2009 in Guan Xian County, Shandong Province, China (36°23'N, 115°47'E). The collection spans the majority of the natural distribution of *P. tomentosa* (30 to 40°N, 105 to 125°E) (Du et al. 2014).

The poplar “84K” (*Populus alba* × *P. glandulosa*) was used for genetic transformation. For clonal propagation, shoot segments of 3 to 4 cm with two or three young leaves were cut from sterilized plantlets and cultivated on solid woody plant medium at 25°C with 16 h light of 5,000 lux and 8 h dark. Then, 30-day-old clonally propagated plantlets were transplanted into soil (potting soil, turfy soil, and vermiculite 1:1:1) and then maintained in a greenhouse (light cycle: 16.0 h of light, 8.0 h of dark; temperature: 24/20°C; relative humidity: 40% to 45%). Light was provided by white fluorescent bulbs with an intensity of 120  $\mu\text{mol m}^{-2} \text{s}^{-1}$ .

### Dynamic growth traits measurement and growth trajectories analysis

We measured the HEI and DBH at eight time points with at least three replicates of 303 *P. tomentosa* accessions during December in 2011 to 2018, which represented a time span from 3 to 10 yr of tree growth (Supplemental Fig. S17). The growth trait data of the association panel are the mean of the three replicates.

The stem DBH and HEI were measured over time, which can be analyzed by fitting growth curves to the phenotypic values then analyzing the parameters describing growth's trajectory (Yang et al. 2006; Camargo et al. 2018). Herein, we plotted the phenotypes against year to produce time trajectories for DBH and HEI respectively (Fig. 1). A previous study demonstrated that the perennial growth of stem is broadly classified into three phases (juvenile, adult, and senescence) (Xu et al. 2016). Our time-series data belonged to the first phase, juvenile, which can actually be described as an “S”-shaped phase. The classical nonlinear growth expression, logistic, has been proven to be an accurate method for fitting growth-age phenotype in plants (Li and Sillanpää 2015; Camargo et al. 2018; Jiang et al. 2018). Thus, we applied this model to describe the growth trajectories of 303 *P. tomentosa* accessions using the 8 yr phenotype of stem DBH and HEI. The logistic equation is expressed as

$$f_i(t) = \frac{K_i}{(1 + a_i e^{-b_i \times t})}$$

where  $f_i(t)$  is the phenotypic value of accession  $i$  ( $i = 1, 2, 3, \dots, 303$ ) measured at age  $t$  ( $t = 3, 4, 5, \dots, 10$ ),  $K$  is the upper asymptote represented the maximal value reached,  $a$  is the value of  $t$  of the sigmoid curve's midpoint, and  $b$  is the tangent in the inflection point represented growth rate. Growth rates calculated this way can capture age-dependent growth of DBH and HEI. Thus, we used the growth curve parameter  $b$  of each accession to identify genetic determinants of stem growth by performing GWAS analysis subsequently.

### Variant calling and SNP annotation

The 303 accessions of the association population were resequenced to a depth >30× (raw data) using the Illumina GA2 sequencing platform according to the manufacturer's instructions. Raw reads were trimmed through a series of quality control procedures using Trimmomatic software (Bolger et al. 2014). Reads with  $\geq 10\%$  unidentified nucleotides (N) or with >50% bases having phred quality <5 were removed. Reads with >10 nt aligned to the adapter, allowing for  $\leq 10\%$  mismatches were also removed. Finally, putative PCR duplicates generated by the PCR amplification in the library construction process (read 1 and read 2 of two paired-end reads that were completely identical) were also removed.

All clean reads from all samples were aligned to the *P. tomentosa* genome using the Burrows–Wheeler Aligner (v0.7.5a-r405) with default parameters (Li and Durbin 2010). Reads with low-quality mapping qualities (MQ < 20) were filtered out using SAMtools (v1.1). We called genome variants using Genome Analysis Toolkit 4.0 (GATK 4.0) with conservative parameters (SNP: QD < 5.0 || MQ < 40.0 || FS > 60.0 || SOR > 3.0 || MQRankSum < -12.5 || ReadPosRankSum < -8.0; Indel: QD < 5.0 || FS > 200.0 || SOR > 10.0 || MQRankSum < -12.5 || ReadPosRankSum < -8.0). The annotation of the identified SNPs was performed using SnpEff (Cingolani et al. 2012). Finally, we kept SNPs

with only bi-allelic positions and the missing rate  $<0.2$  for population genetic analysis.

### Population genomic analyses

To identify genomic selective sweep regions, we calculate genetic diversity  $\pi$  across three geographic populations (NW, NE, and S) using a 10-kb sliding-window with a step size of 5 kb according to the method described by Wang et al. (2018). If  $\pi$  for a window was  $<0.001$  in a population, the window was excluded. Then, we calculate the  $\pi$  ratios between pairs of geographic populations. Population differentiation was evaluated by the fixation index  $F_{ST}$ . Pairwise  $F_{ST}$  values were quantified among the three populations through the sliding-window approach (10-kb windows sliding in 5-kb steps) by VCFtools (Danecek et al. 2011). Chromosomal regions accounting for the top 5%  $\pi$  ratios and top 5%  $F_{ST}$  values were considered as sweep regions, and windows located 20 kb apart were merged into a single sweep region. In addition, the summary statistic, Tajima's D and CLR, were used to confirm the occurrence of selective sweeps in the identified regions. The Tajima's D values were determined in VCFtools (Danecek et al. 2011), and the CLR was calculated in SweeD software (Pavlidis et al. 2013) with 10-kb windows.

### Population structure and phylogenetic analysis

For the population genetics analysis, we obtained a filtered subset consisting of 208,997 SNPs at four-fold degenerate synonymous sites (Minor Allele Frequency (MAF)  $>0.05$  and missing rate  $<0.2$ ). These SNPs were used to perform population structure analysis using Admixture v1.3.0 (Alexander et al. 2009). We conducted a search for the most likely value of K by performing 10 replicate runs with random seed at each value of K ranging from 1 to 10 with 10-fold cross-validation, and  $K=3$  was chosen because it was the minimal value of the cross-validation error (Supplemental Fig. S18, A and B). The SNP set was also used to construct the neighbor joining phylogenetic tree by SNPhylo software (Lee et al. 2014) with a bootstrap value of 1,000 and visualized using FigTree v1.4.3 (<http://tree.bio.ed.ac.uk/software/figtree/>). The resulting phylogenetic tree clearly encompassed three clades (Supplemental Fig. S18C), which was consistent with the results from the Admixture analyses.

### GWAS strategy

We filtered sites using  $MAF > 0.05$  and missing rate  $<0.1$  and obtained 9,689,355 high confidence SNPs that were used for GWAS. We used EMMAX to generate the kinship matrix for all samples and to test for SNP associations with all traits (Kang et al. 2010). The kinship matrix was used as the variance-covariance matrix for the random effect and the population structure components from the Admixture analysis ( $K=3$ ) were included as fixed effects. The growth rate parameter  $b$  from logistic model of 303 accessions was mapped as quantitative traits. Because of the non-independence of SNPs caused by strong LD, it is usually

too strict for significant association detection when multiple-test correction is performed based on the total number of markers (Li et al. 2012; Wang et al. 2016). To alleviate some of these issues, the effective number of independent markers ( $n$ ) was calculated using Genetic type 1 Error Calculator (GEC) software (Li et al. 2012), resulting in a suggestive threshold for the control of the type 1 error rate of  $1.84 \times 10^{-8}$  ( $0.05/n$ ,  $n=2,721,994$ ). LD analysis using the R package LD heatmap was used to define LD blocks surrounding significant SNPs by intervals (Shin et al. 2006). To reduce the noise of candidate gene identification, we selected a single gene candidate for each LD block, as whichever gene contains the SNPs within the coding region or is closest to the intergenic SNPs.

### RNA-seq of different aged *Populus*

Eight tissues (leaves, root, bark, apex, cambium, phloem, mature xylem, and developing xylem) were collected from different aged (1, 2, 3, 4, 5, and 10 year-old) copies of the *P. tomentosa* clone "LM50" with three independent biological replicates grown in the field in Guan Xian County, China. Tissue samples were immediately frozen in liquid nitrogen. Only seven tissues were collected in 1-year-old *Populus* as root tissue was not sampled for that age group. Total RNA was extracted from all samples using a Qiagen RNeasy kit according to the manufacturer's instructions (Qiagen China, Shanghai, China). DNase digestion was performed at the time of RNA purification using an RNase Free DNase set (Qiagen). A total of 3  $\mu\text{g}$  RNA per sample was used to construct strand-specific RNA-seq libraries using the NEBNext UltraTM RNA Library Prep Kit for Illumina (NEB, USA) following manufacturer's recommendations and index codes. Clustering of the index-coded samples was performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumina) according to the manufacturer's instructions. After quantification on a Qubit 2.0 Fluorometer and Agilent 2100 Bioanalyzer, the strand-specific libraries were sequenced on an Illumina HiSeq 2500 instrument, which generated 100-nt paired-end reads. To obtain clean data, the reads containing adapter, reads containing ploy-N, and low-quality reads were removed from the raw data. Clean reads were uniquely mapped to the *P. tomentosa* genome using TopHat 2.1.1 with default options (Trapnell et al. 2009). The isoform levels and gene level counts of the assembled transcripts were computed and normalized based on fragments per kilobase of transcript per million fragments values using Cufflinks v2.1.1 with default options (Trapnell et al. 2012).

### Reverse transcription quantitative PCR (Rt-qPCR)

Total RNA was extracted from each sample using a Plant Total RNA Isolation Kit (Vazyme, Nanjing, China) and reverse transcribed into cDNA using the Reverse Transcription System (Promega Corporation, WI, USA) according to the manufacturer's protocol. RT-qPCR was performed on the 7500 Fast Real-Time PCR System using SYBR Premix Ex Taq

(TaKaRa, Dalian, China) according to the manufacturer's protocol with three technical replicates. The relative gene expression level was calculated using the  $2^{-\Delta\Delta C_t}$  method (Livak and Schmittgen 2021). The 18S gene was used as an internal control. Three biological replicates of RT-qPCR were performed.

### Overexpression and RNAi vector construction and transformation of poplar

To construct the overexpression vector of *PtoP4H9*, the coding sequence (CDS) of *PtoP4H9* was amplified from *P. tomentosa*, cloned into the *pCXS*N vector under the control of the 35S promoter (Chen et al. 2022), and introduced into *Agrobacterium* (*Agrobacterium tumefaciens*) GV3101. To construct the RNAi vector of *PtoP4H9*, we cloned a 226 bp fragment of *PtoP4H9* into the *pCXS*N vector and introduced it into *Agrobacterium* GV3101. The overexpression and RNAi vectors were transformed into poplar 84 K (*P. alba* × *P. glandulosa*) (designated as WT) using *Agrobacterium*-mediated infiltration of leaf disks as described previously (Song et al. 2021). PCR genotyping with the primers for hygromycin/kanamycin-resistance genes was performed to identify positive transgenic individuals. Six independent overexpressing transgenic lines and six independent knockdown transgenic lines were used for further functional analysis. Transgenic and non-transgenic 84 K poplars were propagated via *in vitro* micro-cutting.

### Subcellular location of PtoP4H9

The *PtoP4H9* and phosphate transporter (*PHT4; 6*) coding sequences were respectively fused into GFP/mRFP vectors (Hassler et al. 2012). *Agrobacterium* GV3101 cells harboring the *Pro35S:PtoP4H9-GFP* and *Pro35S:PHT4; 6-mRFP* constructs were mixed and co-infiltrated into *Nicotiana benthamiana* leaf epidermal cells with a syringe. The cells were observed 3 d later under an LSM 710 laser scanning confocal microscope (Zeiss). GFP was excited with an Argon laser at 488 nm and the signal was detected at 505 to 540 nm. The mRFP was excited at 543 nm laser and its emission was collected at 560 to 615 nm. Co-localization between *PtoP4H9-GFP* and *PHT4; 6-mRFP* was analyzed using the ImageJ plugin Coloc2 (Analyze/Co-localization/Coloc2).

### EMSA

The EMSA was performed as previously described by Chen et al. (2022). The full-length CDS of *PtoBPC1* was cloned into the *pET-32a* expression vector (Liu et al. 2022b). Bacterial growth and protein induction were performed as described by the manufacturer (Novagen). Following the induction of protein production in cells with 100 mL of 0.3 mM isopropyl- $\beta$ -D-thiogalactoside and after washing and suspending in phosphate-buffered saline (PBS) (135 mM NaCl, 4.7 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, and 2 mM Na<sub>2</sub>HPO<sub>4</sub>, pH 7.4), bacteria were stored at  $-80^{\circ}\text{C}$ . The purity of the protein was verified by SDS-PAGE analysis. Protein

concentration was quantified and adjusted by dilution in annealing buffer (100 mM Tris-HCl [pH 7.5], 10 mM EDTA, and 1 M NaCl), so that the probe was used at a concentration of 10 mM. The annealing conditions were  $75^{\circ}\text{C}$  for 30 min followed by slow annealing at room temperature for 2 h. Subsequently, 10 mL was diluted to a final concentration of 1 mM (in 100 mL of water). EMSA detection was performed as described by the manufacturer of Chemiluminescent Nucleic Acid Detection Module Kit (ThermoFisher 89880). The probe sequences are listed in Supplemental Table S1.

### SPR

SPR measurements were performed using six dilutions of the oligonucleotides of *PtoP4H9*<sup>hap1</sup> and *PtoP4H9*<sup>hap2</sup> on a Biacore 2000 platform using purified fused PtoBPC1-His proteins immobilized on CM5 chips (GE Healthcare). The dilutions of the oligonucleotides of *PtoP4H9*<sup>hap1</sup> were 0.042, 0.084, 0.167, 0.335, 0.67, and 1.34  $\mu\text{M}$ . The dilutions of the oligonucleotides of *PtoP4H9*<sup>hap2</sup> were 0.209, 0.419, 0.838, 1.675, 3.35, 6.7  $\mu\text{M}$ . Data were analyzed with the Scrubber2-T200 software (BioLogic Software).

### Microscopic observations of poplar stems

Stem segments were sampled from the 8th internode and fixed in a 70% (v/v) formaldehyde-acetic acid (FAA) fixative. Samples were dehydrated in a graded ethanol series (60, 70, 80, 90, 95, and 100%, v/v), and then dehydrated in a 100% dimethylbenzene. Samples were embedded in methyl methacrylate and cross sections of 10  $\mu\text{m}$  thickness were cut using a rotary microtome (HistoCore AUTOCUT; Leica Microsystems) and stained with Safranin O-Fast Green. Sections were visualized using a Leica DMR microscope with a NIKON Eclipse CI (NIKON Biological Microsystems, Japan). The layers and sizes of xylem cells were counted from three biological replicates using ImageJ. At least 100 cell sizes and 10 cell layers were measured.

### TEM of poplar stem cell walls

Stem segments were sampled from the 8th internode and prefixed in 3% glutaraldehyde in 0.2 N sodium phosphate buffer, pH 7.0, and were post-fixed in 2% OsO<sub>4</sub> in PBS, pH 7.2. Following ethanol dehydration, samples were embedded in acrylic resin (London Resin Company). Ultra-thin sections (50 to 70 nm) were double stained with 2% (w/v) uranyl acetate and 2.6% (w/v) lead citrate aqueous solution and examined with a JEM-1230 transmission electron microscope (JEOL) at 80 kV. Quantitative analysis of cell wall thickness was performed using Fiji (Schindelin et al. 2012). Three biological replicates were performed and at least 30 cells were collected.

### Chemical determinations of hydroxyproline (hyp) content

Hydroxyproline (hyp) levels in stem xylem were determined colorimetrically using L-Hyp as a standard according to the

method described by Velasquez et al. (2011). Briefly, 1 mg of sample was hydrolyzed with 6 M HCl-0.1% phenol for 24 h at 105°C. After evaporating the HCl, the sample was ion exchanged on an Amberlite IR-120 (H+) column and free amino acids were eluted with 4 M NH<sub>4</sub>OH. The solution was evaporated and washed with methanol and dried again. The sample was resuspended in 1 mL of distilled water plus 1 mL 0.05 M CuSO<sub>4</sub> and 1 mL 2.5 M NaOH (40°C, 3') followed by the addition of 1 mL 6% H<sub>2</sub>O<sub>2</sub> (40°C, 10'), 4 mL 1.5 M H<sub>2</sub>SO<sub>4</sub>, and 2 mL 5% *p*-dimethylaminobenzaldehyde in *n*-propanol (70°C, 16'). Absorbance at 555 nm and L-Hyp (5 to 20 µg/mL) was used as a reference. Hyp was expressed as percentage of Hyp per mg of dry weight (DW) xylems. Three biological replicates were performed.

### Analysis of cell wall components

Stems from at least three independent 8-week-old WT and transgenic poplar individuals were collected and ground in liquid nitrogen to a fine powder to prepare the alcohol-insoluble residue (AIR) as previously described (Yamaguchi et al. 2010). After a de-starching procedure (Yamaguchi et al. 2010), Klason lignin and carbohydrate contents analyses were carried out based on the National Renewable Energy Laboratory (NREL) protocol (Sluiter et al. 2008). Monosaccharide concentrations (glucose and xylose) were analyzed using HPLC (Agilent 1260, USA), coupled with an evaporative light-scattering detector and an Aminex HPX-87P column (300 × 7.8 mm; Bio-Rad Laboratories, USA) at 85°C. The mobile phase was ultra-pure water at a flow rate of 0.6 mL/min. Three biological replicates were performed.

### MALDI-MS analysis for hydroxyproline arabinosides linkages in HRGPs

The samples were prepared for MALDI-MS analysis as previously described with minor modifications (Velasquez et al. 2011). Briefly, the AIR of stem cells was treated to remove the matrix wall polysaccharides using pectin methylesterase and endo-polygalacturonase followed by endo-cellulase. The resulting residue was hydrolyzed to solubilize the extensins by refluxing for 6 hr in excess 0.22 M Ba(OH)<sub>2</sub>. The hydrolysate was neutralized with H<sub>2</sub>SO<sub>4</sub> and the precipitates were collected and evaporated. MALDI-MS detection was performed on an AXIMA Performance MALDI-MS (Shimadzu, Kyoto, Japan) with 337 nm laser. 20 mg 2,5-Dihydroxybenzoic acid standard was dissolved in 1 mL acetonitrile/H<sub>2</sub>O (1:1, v/v) as matrix solution. 5 µL of matrix solution and 5 µL of the sample were mixed and then spotted 1.0 µL on the stainless plate. The detection was operated in positive reflection mode and the *m/z* range was set from 200 to 1,000 Da. Three biological replicates were performed.

### AFM of poplar stem cell walls

AFM analysis was performed as previously described although some modifications were made to match the protocol suitable for poplar stem cell analysis (Zhang et al. 2021).

Hand-cut cross sections from freshly stem segments of the 8th internode were fixed to glass slide and submerged under 10% mannitol. A BioScope Resolve atomic force microscope equipped with a ScanAsyst-Fluid cantilever (Bruker) with a 20 nm tip radius and 0.7 N m<sup>-1</sup> spring constant was used to scan the topographical structure images of stem fiber cells. The Peak Force QNM mode of the acquisition software was used to detect the topography, peak force error, and Derjaguin–Muller–Toporov modulus image with peak force frequency at 2 kHz and peak force set-point at 3 nN. The topology image size was 10 × 10 µm<sup>2</sup> with a resolution of 128 × 128 pixels recorded at a scan rate of 0.2 Hz. To map apparent Young's modulus, 1 to 2-mm-deep indentations were created along the topological skeletons of fiber cells to ensure relatively normal contact between the probe and sample surface. At least ten indentation positions were chosen for each sample, with each position consecutively indented three times (the average value of three times as the Young's modulus of each position). Three biological replicates were performed. Data were analyzed with Nanoscope Analysis version 1.8.

### Co-expressed genes with *PtoP4H9*

To explore the molecular function of *PtoP4H9*, we carried out co-expression analysis in cambium tissues from differently aged (1, 2, 3, 4, 5, and 10 year-old) copies of the *P. tomentosa* accession "LM50." The co-expressed genes were identified based on pairwise gene expression correlations determined using Pearson's correlation coefficients (*r*) calculated using R (<https://www.r-project.org/>). This identified 530 co-expressed genes using  $|r| \geq 0.85$  and  $P < 0.05$ . Enrichment analysis of GO was performed using the TBtools software (Chen et al. 2020).

### Phylogenetic analysis of P4Hs

Amino acid sequences of genes from the P4Hs family were obtained from *Amborella trichopoda*, *A. thaliana*, *C. reinhardtii*, *Physcomitrium patens*, *P. tomentosa*, and *Solanum lycopersicum* and then aligned using Multiple Alignment using Fast Fourier Transform (Katoh et al. 2019). An unrooted neighbor joining tree was constructed using MEGA7 with the JTT model (Kumar et al. 2016). Bootstrap percentages at the branch points were estimated from 1,000 bootstrap replications. The evolutionary distances (expressed as number of amino acid substitutions per site) were computed using the JTT matrix-based method. An alignment for generating the tree and the tree is provided in Supplemental File 1 and File 2.

### Quantification and statistical analyses

For the annual growth traits, the broad-sense heritability ( $h^2$ ) was estimated by treating accessions as a random effect and the biological replication as a replication effect using the following formula:  $h^2 = Vg/(Vg + Ve)$ , where *Vg* and *Ve* are the variances derived from genetic and environmental effects, respectively. For the growth increment traits, slope and

residual variance, we estimated the narrow-sense heritability using the restricted maximum likelihood method as implemented in GCTA 1.94 (Yang et al. 2011). This analysis estimates the proportion of phenotypic variance explained by all SNPs used in a GWAS in conventionally unrelated accessions, that is, the SNP-based heritability ( $h_{\text{SNP}}^2$ ) (Yang et al. 2017).

All experiments were repeated at least three times. Two-tailed *t* tests and ANOVA least significant difference (LSD) test were used to analyze the significance of differences between the control and treatment groups. A *P*-value <0.05 was considered to be significant. Test statistics are shown in Supplemental Data Set 9.

### Accession numbers

The raw data from the genome resequencing of the 303 *P. tomentosa* accessions and all RNA-seq data for the differently aged *P. tomentosa* clones have been submitted to the Genome Sequence Archive in the BIG Data Center, Beijing Institute of Genomics (BIG), Chinese Academy of Sciences (CAS) under accession numbers CRA000903 and CRA004084. Both data sets are publicly accessible at <http://bigd.big.ac.cn/gsa/>. The reference genome of *P. tomentosa* is downloaded from CNGB Sequence Archive (CNSA) of China National GeneBank DataBase (CNGBdb) with CNSA project ID CNP0004290, and is publicly accessible at <https://db.cngb.org/>.

### Acknowledgments

We thank Dr. Siying Qin at the National Center for Protein Sciences at Peking University for technical help with AFM, Dr. Xiaoyu Zhang and Mr. Ziyang Zhou at the College of Science at Beijing Forestry University for assistance with modeling statistics.

### Author contributions

D.Z. designed the research. L.X., Y.F., and H.Z. performed the research. L.X., Y.F., H.Z., M.Q., J.Z., P.L., D.W., and L.J. collected, analyzed or interpreted the data. L.X. wrote the manuscript. P.K.I., H.X.W., Y.A.E., and Q.D. revised the manuscript. D.Z. obtained funding and he is responsible for this article. All authors read and approved the manuscript.

### Supplemental data

The following materials are available in the online version of this article.

**Supplemental Figure S1.** Phenotypic correlation matrix of dynamic growth traits in *P. tomentosa* natural population.

**Supplemental Figure S2.** Phenotypic variation of dynamic growth traits in *P. tomentosa* natural population.

**Supplemental Figure S3.** The fitting curves of growth trajectories of 303 accessions.

**Supplemental Figure S4.** Data distribution of observed data and predicted values of logistic model.

**Supplemental Figure S5.** *PtoBIGE1B* is a candidate gene associated with growth rate of HEI.

**Supplemental Figure S6.** PVE by significant SNPs and expression pattern of candidate genes.

**Supplemental Figure S7.** GWAS results of *PtoP4H9* associated with eight measured traits of DBH separately.

**Supplemental Figure S8.** Temporal expression pattern of *PtoP4H9* between different haplotype accessions of *PtoP4H9*.

**Supplemental Figure S9.** Negative expression pattern of *PtoP4H9* and *PtoBPC1*.

**Supplemental Figure S10.** Morphological variation of stem xylem fiber cells between different haplotype accessions of *PtoP4H9*.

**Supplemental Figure S11.** Phenotypic comparison of *PtoP4H9*-OE and *ptop4h9*-KD lines.

**Supplemental Figure S12.** Phylogenetic tree of P4H proteins.

**Supplemental Figure S13.** Expression patterns of P4Hs in *P. tremula*.

**Supplemental Figure S14.** Time-course expression of co-expressed genes with *PtoP4H9* in cambium tissues of different aged *P. tomentosa*.

**Supplemental Figure S15.** Quantitative measurement of xylem cell wall components.

**Supplemental Figure S16.** Environment factors in the natural distribution region of *P. tomentosa*.

**Supplemental Figure S17.** The images of the field site and measured trees.

**Supplemental Figure S18.** Population structure of the natural population of *P. tomentosa*.

**Supplemental Table S1.** Oligonucleotide sequences of the primers used in this study.

**Supplemental Data Set S1.** Stem DBH and HEI growth trajectories of 303 *P. tomentosa* accessions in 3 to 10 yr referred to Fig. 1.

**Supplemental Data Set S2.** Phenotypic correlation of time-point growth traits referred to Supplemental Fig. S1.

**Supplemental Data Set S3.** Phenotypic variations and heritabilities of time-point growth traits.

**Supplemental Data Set S4.** Growth trajectories analysis by logistic model of 303 *P. tomentosa* accessions.

**Supplemental Data Set S5.** Genome-wide association signals of growth rate traits.

**Supplemental Data Set S6.** Gene annotation of the candidate genes underlying growth rate traits.

**Supplemental Data Set S7.** Expression pattern of candidate genes underlying growth trajectory traits in eight tissues from different aged *P. tomentosa* referred to Supplemental Fig. S6B.

**Supplemental Data Set S8.** Co-expressed genes with *PtoP4H9* in cambium tissues referred to Supplemental Fig. S14.

**Supplemental Data Set S9.** Statistical analysis results for each figure.

**Supplemental File S1.** Sequence alignment of P4H proteins from *A. trichopoda*, *A. thaliana*, *C. reinhardtii*, *P. patens*, *P. tomentosa*, and *S. lycopersicum*.

**Supplemental File S2.** Tree file for P4H proteins from *A. trichopoda*, *A. thaliana*, *C. reinhardtii*, *P. patens*, *P. tomentosa*, and *S. lycopersicum*.

## Funding

This work was supported by the State “14.5” National Key Research and Development Program of China (No. 2021YFD2200101), the Project of the National Natural Science Foundation of China (Nos. 32170370, 31872707, 32201584, and 31901336), Key Research and Development Project of Zhejiang Province (No. 2021C02054), and the Program of Introducing Talents of Discipline to Universities (111 Project, No. B20050).

*Conflict of interest statement.* None declared.

## Data availability

The raw data from the genome resequencing of the 303 *P. tomentosa* accessions and all RNA-seq data for the differently aged *P. tomentosa* clones have been submitted to the Genome Sequence Archive in the BIG Data Center, Beijing Institute of Genomics (BIG), Chinese Academy of Sciences (CAS) under accession numbers CRA000903 and CRA004084. Both data sets are publicly accessible at <http://bigd.big.ac.cn/gsa/>. The reference genome of *Populus tomentosa* is download from CNGB Sequence Archive (CNSA) of China National GeneBank DataBase (CNGBdb) with CNSA project ID CNP0004290, and is publicly accessible at <https://db.cnbg.org/>.

## References

- Alexander DH, Novembre J, Lange K.** Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 2009;**19**(9):1655–1664. <https://doi.org/10.1101/gr.094052.109>
- Bac-Molenaar JA, Vreugdenhil D, Granier C, Keurentjes JJ.** Genome-wide association mapping of growth dynamics detects time-specific and general quantitative trait loci. *J Exp Bot.* 2015;**66**(18):5567–5580. <https://doi.org/10.1093/jxb/erv176>
- Baison J, Vidalis A, Zhou L, Chen Z, Li Z, Sillanpää MJ, Bernhardsson C, Scofield D, Forsberg N, Grahn T, et al.** Genome-wide association study identified novel candidate loci affecting wood formation in Norway spruce. *Plant J.* 2019;**100**(1):83–100. <https://doi.org/10.1111/tbj.14429>
- Bhalerao RP, Fischer U.** Environmental and hormonal control of cambial stem cell dynamics. *J Exp Bot.* 2017;**68**(1):79–87. <https://doi.org/10.1093/jxb/erw466>
- Bolger AM, Lohse M, Usadel B.** Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;**30**(15):2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Bradshaw HD, Ceulemans R, Davis J, Stettler R.** Emerging model systems in plant biology: poplar (*Populus*) as a model forest tree. *J Plant Growth Regul.* 2000;**19**(3):306–313. <https://doi.org/10.1007/s003440000030>
- Bresadola L, Caseys C, Castiglione S, Buerkle CA, Wegmann D, Lexer C.** Admixture mapping in interspecific *Populus* hybrids identifies classes of genomic architectures for phytochemical, morphological and growth traits. *New Phytol.* 2019;**223**(4):2076–2089. <https://doi.org/10.1111/nph.15930>
- Carmago AV, Mackay I, Mott R, Han J, Doonan JH, Askew K, Corke F, Williams K, Bentley AR.** Functional mapping of quantitative trait loci (QTLs) associated with plant performance in a wheat MAGIC mapping population. *Front Plant Sci.* 2018;**9**:887. <https://doi.org/10.3389/fpls.2018.00887>
- Campbell MT, Knecht AC, Berger B, Brien CJ, Wang D, Walia H.** Integrating image-based phenomics and association analysis to dissect the genetic architecture of temporal salinity responses in rice. *Plant Physiol.* 2015;**168**(4):1476–1489. <https://doi.org/10.1104/pp.15.00450>
- Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R.** TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol Plant.* 2020;**13**(8):1194–1202. <https://doi.org/10.1016/j.molp.2020.06.009>
- Chen P, Song Y, Liu X, Xiao L, Bu C, Liu P, Zhao L, Ingvarsson PK, Wu HX, El-Kassaby YA, et al.** LncRNA *PMAT-PtoMYB46* module represses *PtoMATE* and *PtoARF2* promoting Pb(2+) uptake and plant growth in poplar. *J Hazard Mater.* 2022;**433**:128769. <https://doi.org/10.1016/j.jhazmat.2022.128769>
- Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM.** A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin).* 2012;**6**(2):80–92. <https://doi.org/10.4161/fly.19695>
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al.** The variant call format and VCFtools. *Bioinformatics.* 2011;**27**(15):2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>
- De Rybel B, Mähönen AP, Helariutta Y, Weijers D.** Plant vascular development: from early specification to differentiation. *Nat Rev Mol Cell Biol.* 2016;**17**(1):30–40. <https://doi.org/10.1038/nrm.2015.6>
- Ding J, Nilsson O.** Molecular regulation of phenology in trees—because the seasons they are a-changing. *Curr Opin Plant Biol.* 2016;**29**:73–79. <https://doi.org/10.1016/j.pbi.2015.11.007>
- Ding J, Zhang B, Li Y, André D, Nilsson O.** Phytochrome B and PHYTOCHROME INTERACTING FACTOR8 modulate seasonal growth in trees. *New Phytol.* 2021;**232**(6):2339–2352. <https://doi.org/10.1111/nph.17350>
- Du Q, Xu B, Gong C, Yang X, Pan W, Tian J, Li B, Zhang D.** Variation in growth, leaf, and wood property traits of Chinese white poplar (*Populus tomentosa*), a major industrial tree species in Northern China. *Can J Forest Res.* 2014;**44**(4):326–339. <https://doi.org/10.1139/cjfr-2013-0416>
- El-Lithy ME, Clercx EJ, Ruys GJ, Koornneef M, Vreugdenhil D.** Quantitative trait locus analysis of growth-related traits in a new *Arabidopsis* recombinant inbred population. *Plant Physiol.* 2004;**135**(1):444–458. <https://doi.org/10.1104/pp.103.036822>
- Evans LM, Slavov GT, Rodgers-Melnick E, Martin J, Ranjan P, Muchero W, Brunner AM, Schackwitz W, Gunter L, Chen JG, et al.** Population genomics of *Populus trichocarpa* identifies signatures of selection and adaptive trait associations. *Nat Genet.* 2014;**46**(10):1089–1096. <https://doi.org/10.1038/ng.3075>
- Fahrenkrog AM, Neves LG, Resende MJ, Vazquez AI, de Los CG, Dervinis C, Sykes R, Davis M, Davenport R, Barbazuk WB, et al.** Genome-wide association study reveals putative regulators of bioenergy traits in *Populus deltoides*. *New Phytol.* 2017;**213**(2):799–811. <https://doi.org/10.1111/nph.14154>
- Feng L, Jiang P, Li C, Zhao J, Dong A, Yang D, Wu R.** Genetic dissection of growth trajectories in forest trees: from FunMap to FunGraph. *Forest Res.* 2021;**1**(1):1–10. <https://doi.org/10.48130/FR-2021-0019>
- Fischer U, Kucukoglu M, Helariutta Y, Bhalerao RP.** The dynamics of cambial stem cell activity. *Annu Rev Plant Biol.* 2019;**70**(1):293–319. <https://doi.org/10.1146/annurev-arplant-050718-100402>

- Fragkostefanakis S, Sedeek KE, Raad M, Zaki MS, Kalaitzis P.** Virus-induced gene silencing of three putative prolyl 4-hydroxylases enhances plant growth in tomato (*Solanum lycopersicum*). *Plant Mol Biol*. 2014;**85**(4–5):459–471. <https://doi.org/10.1007/s11103-014-0197-6>
- Gong H, Zhang XY, Zhu S, Jiang L, Zhu X, Fang Q, Wu R.** Genetic architecture of multiphasic growth covariation as revealed by a non-linear mixed mapping framework. *Front Plant Sci*. 2021;**12**:711219. <https://doi.org/10.3389/fpls.2021.711219>
- Han X, An Y, Zhou Y, Liu C, Yin W, Xia X.** Comparative transcriptome analyses define genes and gene modules differing between two *Populus* genotypes with contrasting stem growth rates. *Biotechnol Biofuels*. 2020;**13**(1):139. <https://doi.org/10.1186/s13068-020-01758-0>
- Hassler S, Lemke L, Jung B, Mohlmann T, Kruger F, Schumacher K, Espen L, Martinoia E, Neuhaus HE.** Lack of the Golgi phosphate transporter PHT4; 6 causes strong developmental defects, constitutively activated disease resistance mechanisms and altered intracellular phosphate compartmentation in *Arabidopsis*. *Plant J*. 2012;**72**(5):732–744. <https://doi.org/10.1111/j.1365-313X.2012.05106.x>
- Hill JJ, Hollender CA.** Branching out: new insights into the genetic regulation of shoot architecture in trees. *Curr Opin Plant Biol*. 2019;**47**:73–80. <https://doi.org/10.1016/j.pbi.2018.09.010>
- Jiang L, He X, Jin Y, Sang M, Chen N, Zhu J, Zhang Z, Li J, Wu R.** A mapping framework of competition–cooperation QTLs that drive community dynamics. *Nat Commun*. 2018;**9**(1):3010. <https://doi.org/10.1038/s41467-018-05416-w>
- Jiang L, Ye M, Zhu S, Zhai Y, Xu M, Huang M, Wu R.** Computational identification of genes modulating stem height-diameter allometry. *Plant Biotechnol J*. 2016;**14**(12):2254–2264. <https://doi.org/10.1111/pbi.12579>
- Johnsson C, Fischer U.** Cambial stem cells and their niche. *Plant Sci*. 2016;**252**:239–245. <https://doi.org/10.1016/j.plantsci.2016.08.002>
- Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E.** Variance component model to account for sample structure in genome-wide association studies. *Nat Genet*. 2010;**42**(4):348–354. <https://doi.org/10.1038/ng.548>
- Katoh K, Rozewicki J, Yamada KD.** MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform*. 2019;**20**(4):1160–1166. <https://doi.org/10.1093/bib/bbx108>
- Kauter D, Lewandowski I, Claupein W.** Quantity and quality of harvestable biomass from *Populus* short rotation coppice for solid fuel use—a review of the physiological basis and management influences. *Biomass Bioenergy*. 2003;**24**(6):411–427. [https://doi.org/10.1016/S0961-9534\(02\)00177-0](https://doi.org/10.1016/S0961-9534(02)00177-0)
- Keskiaho K, Hieta R, Sormunen R, Myllyharju J.** *Chlamydomonas reinhardtii* has multiple prolyl 4-hydroxylases, one of which is essential for proper cell wall assembly. *Plant Cell*. 2007;**19**(1):256–269. <https://doi.org/10.1105/tpc.106.042739>
- Kieliszewski MJ, Lampert DT.** Extensin: repetitive motifs, functional sites, post-translational codes, and phylogeny. *Plant J*. 1994;**5**(2):157–172. <https://doi.org/10.1046/j.1365-313X.1994.05020157.x>
- Knoch D, Abbadì A, Grandke F, Meyer RC, Samans B, Werner CR, Snowdon RJ, Altmann T.** Strong temporal dynamics of QTL action on plant growth progression revealed through high-throughput phenotyping in canola. *Plant Biotechnol J*. 2020;**18**(1):68–82. <https://doi.org/10.1111/pbi.13171>
- Kooiker M, Airoidi CA, Losa A, Manzotti PS, Finzi L, Kater MM, Colombo L.** BASIC PENTACYSSTEINE1, a GA binding protein that induces conformational changes in the regulatory region of the homeotic *Arabidopsis* gene *SEEDSTICK*. *Plant Cell*. 2005;**17**(3):722–729. <https://doi.org/10.1105/tpc.104.030130>
- Kumar S, Stecher G, Tamura K.** MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol*. 2016;**33**(7):1870–1874. <https://doi.org/10.1093/molbev/msw054>
- Lee TH, Guo H, Wang X, Kim C, Paterson AH.** SNPPhylo: a pipeline to construct a phylogenetic tree from huge SNP data. *BMC Genomics*. 2014;**15**(1):162. <https://doi.org/10.1186/1471-2164-15-162>
- Lehmann F, Hardtke CS.** Secondary growth of the *Arabidopsis* hypocotyl-vascular development in dimensions. *Curr Opin Plant Biol*. 2016;**29**:9–15. <https://doi.org/10.1016/j.pbi.2015.10.011>
- Li H, Durbin R.** Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010;**26**(5):589–595. <https://doi.org/10.1093/bioinformatics/btp698>
- Li Z, Lathe RS, Li J, He H, Bhalerao RP.** Towards understanding the biological foundations of perenniality. *Trends Plant Sci*. 2022;**27**(1):56–68. <https://doi.org/10.1016/j.tplants.2021.08.007>
- Li Z, Sillanpää MJ.** Dynamic quantitative trait locus analysis of plant phenomic data. *Trends Plant Sci*. 2015;**20**(12):822–833. <https://doi.org/10.1016/j.tplants.2015.08.012>
- Li MX, Yeung JM, Cherny SS, Sham PC.** Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Hum Genet*. 2012;**131**(5):747–756. <https://doi.org/10.1007/s00439-011-1118-2>
- Liu X, Liu H, Li H, An X, Song L, You C, Zhao L, Tian Y, Wang X.** MdMYB10 affects nitrogen uptake and reallocation by regulating the nitrate transporter MdNRT2.4–1 in red-fleshed apple. *Hortic Res*. 2022a;**9**:uhac016. <https://doi.org/10.1093/hr/uhac016>
- Liu R, Wen S, Sun T, Wang R, Zuo W, Yang T, Wang C, Hu L, Lu M, Wang L.** PagWOX11/12a positively regulates the PagSAUR36 gene that enhances adventitious root development in poplar. *J Exp Bot*. 2022b;**73**(22):7298–7311. <https://doi.org/10.1093/jxb/erac345>
- Livak KJ, Schmittgen TD.** Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  method. *Methods*. 2001;**25**(4):402–408. <https://doi.org/10.1006/meth.2001.1262>
- Marzol E, Borassi C, Bringas M, Sede A, Rodríguez GD, Capece L, Estevez JM.** Filling the gaps to solve the extensin puzzle. *Mol Plant*. 2018;**11**(5):645–658. <https://doi.org/10.1016/j.molp.2018.03.003>
- McLane SC, LeMay VM, Aitken SN.** Modeling lodgepole pine radial growth relative to climate and genetics using universal growth-trend response functions. *Ecol Appl*. 2011;**21**(3):776–788. <https://doi.org/10.1890/10-0131.1>
- Muraya MM, Chu J, Zhao Y, Junker A, Klukas C, Reif JC, Altmann T.** Genetic variation of growth dynamics in maize (*Zea mays* L.) revealed through automated non-invasive phenotyping. *Plant J*. 2017;**89**(2):366–380. <https://doi.org/10.1111/tpj.13390>
- O'Neill GA, Hamann A, Wang T.** Accounting for population variation improves estimates of the impact of climate change on species' growth and distribution. *J Appl Ecol*. 2008;**45**(4):1040–1049. <https://doi.org/10.1111/j.1365-2664.2008.01472.x>
- Paul L, Rinne P, van der Schoot C.** Shoot meristems of deciduous woody perennials: self-organization and morphogenetic transitions. *Curr Opin Plant Biol*. 2014;**17**:86–95. <https://doi.org/10.1016/j.pbi.2013.11.009>
- Pautasso M.** Geographical genetics and the conservation of forest trees. *Perspect Plant Ecol*. 2009;**11**(3):157–189. <https://doi.org/10.1016/j.ppees.2009.01.003>
- Pavlidis P, Živković D, Stamatakis A, Alachiotis N.** Sweed: likelihood-based detection of selective sweeps in thousands of genomes. *Mol Biol Evol*. 2013;**30**(9):2224–2234. <https://doi.org/10.1093/molbev/mst112>
- Peltola H, Kilpeläinen A, Kellomaki S.** Diameter growth of Scots pine (*Pinus sylvestris*) trees grown at elevated temperature and carbon dioxide concentration under boreal conditions. *Tree Physiol*. 2002;**22**(14):963–972. <https://doi.org/10.1093/treephys/22.14.963>
- Rathgeber CB, Cuny HE, Fonti P.** Biological basis of tree-ring formation: a crash course. *Front Plant Sci*. 2016;**7**:734. <https://doi.org/10.3389/fpls.2016.00734>
- Reich PB, Oleksyn J.** Climate warming will reduce growth and survival of Scots pine except in the far north. *Ecol Lett*. 2008;**11**(6):588–597. <https://doi.org/10.1111/j.1461-0248.2008.01172.x>
- Resende RT, Resende MD, Silva FF, Azevedo CF, Takahashi EK, Silva-Junior OB, Grattapaglia D.** Regional heritability mapping and genome-wide association identify loci for complex growth, wood and disease resistance traits in Eucalyptus. *New Phytol*. 2017;**213**(3):1287–1300. <https://doi.org/10.1111/nph.14266>

- Rohde A, Bhalerao RP.** Plant dormancy in the perennial context. *Trends Plant Sci.* 2007;**12**(5):217–223. <https://doi.org/10.1016/j.tplants.2007.03.012>
- Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld B, Tinevez JY, et al.** Fiji: an open-source platform for biological-image analysis. *Nat Meth.* 2012;**9**(7):676–682. <https://doi.org/10.1038/nmeth.2019>
- Shim D, Ko JH, Kim WC, Wang Q, Keathley DE, Han KH.** A molecular framework for seasonal growth-dormancy regulation in perennial plants. *Hortic Res.* 2014;**1**(1):14059. <https://doi.org/10.1038/hortres.2014.59>
- Shin JH, Blay S, McNeney B, Graham J.** LDheatmap: an R function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms. *J Stat Softw.* 2006;**16**(Code Snippet 3):1–9. <https://doi.org/10.18637/jss.v016.c03>
- Showalter AM, Keppler B, Lichtenberg J, Gu D, Welch LR.** A bioinformatics approach to the identification, classification, and analysis of hydroxyproline-rich glycoproteins. *Plant Physiol.* 2010;**153**(2):485–513. <https://doi.org/10.1104/pp.110.156554>
- Simonini S, Kater MM.** Class I BASIC PENTACYSTEINE factors regulate *HOMEODOMAIN* genes involved in meristem size maintenance. *J Exp Bot.* 2014;**65**(6):1455–1465. <https://doi.org/10.1093/jxb/eru003>
- Singh RK, Maurya JP, Azeez A, Miskolczi P, Tylewicz S, Stojković K, Delhomme N, Busov V, Bhalerao RP.** A genetic network mediating the control of bud break in hybrid aspen. *Nat Commun.* 2018;**9**(1):4173. <https://doi.org/10.1038/s41467-018-06696-y>
- Singh RK, Miskolczi P, Maurya JP, Bhalerao RP.** A tree ortholog of *SHORT VEGETATIVE PHASE* floral repressor mediates photoperiodic control of bud dormancy. *Curr Biol.* 2019;**29**(1):128–133.e2. <https://doi.org/10.1016/j.cub.2018.11.006>
- Singh RK, Svystun T, Aldahmash B, Jönsson AM, Bhalerao RP.** Photoperiod- and temperature-mediated control of phenology in trees—a molecular perspective. *New Phytol.* 2017;**213**(2):511–524. <https://doi.org/10.1111/nph.14346>
- Sluiter A, Hames B, Ruiz R, Scarlata C, Sluiter J, Templeton D, Crocker DLAP.** Determination of structural carbohydrates and lignin in biomass. *Lab Anal Proced.* 2008;**1617**(1):1–16.
- Smet W, De Rybel B.** Genetic and hormonal control of vascular tissue proliferation. *Curr Opin Plant Biol.* 2016;**29**:50–56. <https://doi.org/10.1016/j.pbi.2015.11.004>
- Song Y, Chen P, Xuan A, Bu C, Liu P, Ingvarsson PK, El-Kassaby YA, Zhang D.** Integration of genome wide association studies and co-expression networks reveal roles of *PtoWRKY42-PtoUGT76C1-1* in *trans*-zeatin metabolism and cytokinin sensitivity in poplar. *New Phytol.* 2021;**231**(4):1462–1477. <https://doi.org/10.1111/nph.17469>
- Sun S, Li C, Paterson AH, Jiang Y, Xu R, Robertson JS, Snider JL, Chee PW.** In-field high throughput phenotyping and cotton plant growth analysis using LiDAR. *Front Plant Sci.* 2018;**9**:16. <https://doi.org/10.3389/fpls.2018.00016>
- Sundell D, Street NR, Kumar M, Mellerowicz EJ, Kucukoglu M, Johnsson C, Kumar V, Mannapperuma C, Delhomme N, Nilsson O, et al.** AspWood: high-spatial-resolution transcriptome profiles reveal uncharacterized modularity of wood formation in *Populus tremula*. *Plant Cell.* 2017;**29**(7):1585–1604. <http://dx.doi.org/10.1105/tpc.17.00153>
- Suzuki M, Sato Y, Wu S, Kang BH, McCarty DR.** Conserved functions of the MATE transporter *BIG EMBRYO1* in regulation of lateral organ size and initiation rate. *Plant Cell.* 2015;**27**(8):2288–2300. <https://doi.org/10.1105/tpc.15.00290>
- Trapnell C, Pachter L, Salzberg SL.** TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics.* 2009;**25**(9):1105–1111. <https://doi.org/10.1093/bioinformatics/btp120>
- Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L.** Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc.* 2012;**7**(3):562–578. <https://doi.org/10.1038/nprot.2012.016>
- Vaganov EA, Hughes MK, Shashkin AV.** Growth dynamics of conifer tree rings: images of past and future environments. Vol. 183. Berlin, Heidelberg: Springer; 2006.
- Velasquez SM, Ricardi MM, Dorosz JG, Fernandez PV, Nadra AD, Pol-Fachin L, Egelund J, Gille S, Harholt J, Ciancia M, et al.** O-glycosylated cell wall proteins are essential in root hair growth. *Science.* 2011;**332**(6036):1401–1403. <https://doi.org/10.1126/science.1206657>
- Velasquez SM, Ricardi MM, Poulsen CP, Oikawa A, Dilokpimol A, Halim A, Mangano S, Denita JS, Marzol E, Salgado SJ, et al.** Complex regulation of prolyl-4-hydroxylases impacts root hair expansion. *Mol Plant.* 2015;**8**(5):734–746. <https://doi.org/10.1016/j.molp.2014.11.017>
- Wang J, Ding J, Tan B, Robinson KM, Michelson IH, Johansson A, Nystedt B, Scofield DG, Nilsson O, Jansson S, et al.** A major locus controls local adaptation and adaptive life history variation in a perennial plant. *Genome Biol.* 2018;**19**(1):1–17. <https://doi.org/10.1186/s13059-018-1444-y>
- Wang S, Feng J, Ren W, Huang B, Zhou L, Wen Y, Zhang J, Dunwell JM, Xu S, Zhang Y.** Improving power and accuracy of genome-wide association studies via a multi-locus mixed linear model methodology. *Sci Rep.* 2016;**6**(1):19444. <https://doi.org/10.1038/srep19444>
- Wang J, Park MY, Wang L, Koo Y, Chen X, Weigel D, Poethig RS.** miRNA control of vegetative phase change in trees. *PLoS Genet.* 2011;**7**(2):e1002012. <https://doi.org/10.1371/journal.pgen.1002012>
- Wei S, Yang G, Yang Y, Yin T.** Time-sequential detection of quantitative trait loci and candidate genes underlying the dynamic growth of *Salix suchowensis*. *Tree Physiol.* 2022;**42**(4):877–890. <https://doi.org/10.1093/treephys/tpab138>
- Wu R, Jiang L.** Recovering dynamic networks in big static datasets. *Phys Rep.* 2021;**912**:1–57. <https://doi.org/10.1016/j.physrep.2021.01.003>
- Xiao L, Du Q, Fang Y, Quan M, Lu W, Wang D, Si J, El-Kassaby YA, Zhang D.** Genetic architecture of the metabolic pathway of salicylic acid biosynthesis in *Populus*. *Tree Physiol.* 2021;**41**(11):2198–2215. <https://doi.org/10.1093/treephys/tpab068>
- Xu M, Jiang L, Zhu S, Zhou C, Ye M, Mao K, Sun L, Su X, Pan H, Zhang S, et al.** A computational framework for mapping the timing of vegetative phase change. *New Phytol.* 2016;**211**(2):750–760. <https://doi.org/10.1111/nph.13907>
- Xu J, Tan L, Lampion DTA, Showalter AM, Kieliszewski MJ.** The O-Hyp glycosylation code in tobacco and *Arabidopsis* and a proposed role of Hyp-glycans in secretion. *Phytochemistry.* 2008;**69**(8):1631–1640. <https://doi.org/10.1016/j.phytochem.2008.02.006>
- Yamaguchi M, Goue N, Igarashi H, Ohtani M, Nakano Y, Mortimer JC, Nishikubo N, Kubo M, Katayama Y, Kakegawa K, et al.** VASCULAR-RELATED NAC-DOMAIN6 and VASCULAR-RELATED NAC-DOMAIN7 effectively induce transdifferentiation into xylem vessel elements under control of an induction system. *Plant Physiol.* 2010;**153**(3):906–914. <https://doi.org/10.1104/pp.110.154013>
- Yan J, Liu Y, Yang L, He H, Huang Y, Fang L, Scheller HV, Jiang M, Zhang A.** Cell wall  $\beta$ -1, 4-galactan regulated by the BPC1/BPC2-GALS1 module aggravates salt sensitivity in *Arabidopsis thaliana*. *Mol Plant.* 2021;**14**(3):411–425. <https://doi.org/10.1016/j.molp.2020.11.023>
- Yang J, Lee SH, Goddard ME, Visscher PM.** GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet.* 2011;**88**(1):76–82. <https://doi.org/10.1016/j.ajhg.2010.11.011>
- Yang R, Tian Q, Xu S.** Mapping quantitative trait loci for longitudinal traits in line crosses. *Genetics.* 2006;**173**(4):2339–2356. <https://doi.org/10.1534/genetics.105.054775>
- Yang J, Zeng J, Goddard ME, Wray NR, Visscher PM.** Concepts, estimation and interpretation of SNP-based heritability. *Nat Genet.* 2017;**49**(9):1304–1310. <https://doi.org/10.1038/ng.3941>
- Ye M, Jiang L, Mao K, Wang Y, Wang Z, Wu R.** Functional mapping of seasonal transition in perennial plants. *Brief Bioinform.* 2015;**16**(3):526–535. <https://doi.org/10.1093/bib/bbu025>

**Zhang H, Guo Z, Zhuang Y, Suo Y, Du J, Gao Z, Pan J, Li L, Wang T, Xiao L, et al.** MicroRNA775 regulates intrinsic leaf size and reduces cell wall pectin levels by targeting a galactosyltransferase gene in *Arabidopsis*. *Plant Cell*. 2021;**33**(3):581–602. <https://doi.org/10.1093/plcell/koaa049>

**Zhang X, Huang C, Wu D, Qiao F, Li W, Duan L, Wang K, Xiao Y, Chen G, Liu Q, et al.** High-throughput phenotyping and QTL mapping reveals the genetic architecture of maize plant growth. *Plant Physiol*. 2017;**173**(3):1554–1564. <https://doi.org/10.1104/pp.16.01516>