

Populus trichocarpa cell wall chemistry and ultrastructure trait variation, genetic control and genetic correlations

Ilga Porth^{1*}, Jaroslav Klápště^{2*}, Oleksandr Skyba¹, Ben S. K. Lai¹, Armando Geraldés³, Wellington Muchero⁴, Gerald A. Tuskan⁴, Carl J. Douglas³, Yousry A. El-Kassaby² and Shawn D. Mansfield¹

¹Department of Wood Science, Faculty of Forestry, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada; ²Department of Forest Sciences, Faculty of Forestry, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada; ³Department of Botany, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada; ⁴BioSciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, 37831, USA

Summary

Authors for correspondence:
Shawn D. Mansfield
Tel: +1 604 822 0196
Email: shawn.mansfield@ubc.ca

Yousry A. El-Kassaby
Tel: +1 604 822 1821
Email: y.el-kassaby@ubc.ca

Received: 25 July 2012
Accepted: 18 September 2012

New Phytologist (2012)
doi: 10.1111/nph.12014

Key words: biofuel traits, genetic correlation, narrow-sense heritability, *Populus trichocarpa*, wood chemistry, wood composition, wood physical properties.

- The increasing ecological and economical importance of *Populus* species and hybrids has stimulated research into the investigation of the natural variation of the species and the estimation of the extent of genetic control over its wood quality traits for traditional forestry activities as well as the emerging bioenergy sector. A realized kinship matrix based on informative, high-density, biallelic single nucleotide polymorphism (SNP) genetic markers was constructed to estimate trait variance components, heritabilities, and genetic and phenotypic correlations.
- Seventeen traits related to wood chemistry and ultrastructure were examined in 334 9-yr-old *Populus trichocarpa* grown in a common-garden plot representing populations spanning the latitudinal range 44° to 58.6°. In these individuals, 9342 SNPs that conformed to Hardy–Weinberg expectations were employed to assess the genomic pair-wise kinship to estimate narrow-sense heritabilities and genetic correlations among traits.
- The range-wide phenotypic variation in all traits was substantial and several trait heritabilities were > 0.6. In total, 61 significant genetic and phenotypic correlations and a network of highly interrelated traits were identified.
- The high trait variation, the evidence for moderate to high heritabilities and the identification of advantageous trait combinations of industrially important characteristics should aid in providing the foundation for the enhancement of poplar tree breeding strategies for modern industrial use.

Introduction

Populus species (poplars, aspens and cottonwoods; referred to as poplar throughout) are ecologically and economically important tree species that span a substantial geographic area in temperate and boreal regions across the Northern Hemisphere (Burns & Honkala, 1990; Mansfield & Weineisen, 2007). Although gymnosperms are predominantly employed for conversion to commodity products, poplars are becoming increasingly important to the forest products industry and in the new green economy, particularly as a short-rotation species for the provision of fiber for pulp and paper production, oriented strand-board (OSB), moldings and trim, and furniture components (Semple *et al.*, 2007). More recently, these species have emerged as prospective bioenergy crops because of their inherent rapid growth rates and favorable cell wall chemistry (Dinus, 2000). Lignocellulosic feedstocks are considered to be more sustainable than starch-

and sugar-based feedstocks for bioenergy production in terms of ecological benefits, energy and carbon dioxide balance (Yuan *et al.*, 2008). Although poplar can generate relatively high yields and significant energy output (Yuan *et al.*, 2008; Carroll & Somerville, 2009), its innate cell wall composition requires costly pretreatment methods to access the complex cell wall carbohydrates (Hendriks & Zeeman, 2009; Mansfield, 2009). Many factors have been proposed to contribute to this natural recalcitrance of plant cell walls to deconstruction (Chang & Holtzapfel, 2000; Dinus, 2000; Mosier *et al.*, 2005; Studer *et al.*, 2011) and, at present, several research strategies are underway to attempt to optimize and enhance saccharification efficiency (Mansfield *et al.*, 1999, 2012; Yuan *et al.*, 2008; Carroll & Somerville, 2009; Shen *et al.*, 2011).

Biomass accumulation, wood composition and ultrastructural traits are known to be highly complex (Dinus, 2000; Novaes *et al.*, 2009). It is also known that poplars show substantial variation in many important wood properties, such as lignin content and syringyl to guaiacyl (S:G) lignin monomer ratio

*These authors contributed equally to this work.

(Wegrzyn *et al.*, 2010; Studer *et al.*, 2011), cellulose and hemicellulose composition (Sannigrahi *et al.*, 2010), cellulose crystallinity (Kumar *et al.*, 2009) and wood density (Zhang *et al.*, 2012). Poplar tree improvement programs can capitalize on both the existing natural variation in wood traits as well as the vast genomic resources that have become available for this species in recent years (Tuskan *et al.*, 2006; Jansson & Douglas, 2007; Ralph *et al.*, 2008; Geraldès *et al.*, 2011; Geraldès *et al.*, in press; Slavov *et al.*, 2012) to accelerate breeding strategies for improved cell wall and growth traits. Quantitative trait locus (QTL) identification and other traditional quantitative genetics approaches have been implemented to study the complex genetic architecture underlying wood formation (Pot *et al.*, 2002; Kirst *et al.*, 2004; Novaes *et al.*, 2009; Thumma *et al.*, 2010; Wegrzyn *et al.*, 2010; Beaulieu *et al.*, 2011). However, knowledge on the mode of genetic control (heritability) of the wood traits and their genetic relationships is essential for effective selective breeding programs. The aim of breeding programs is to maximize the genetic gain in industrially important wood traits and, concurrently, to effectively manage undesirable trait correlations whilst sustaining genetic variability for future selection and adaptation. For this, a better understanding of the inheritance and relationships among traits of interest is needed. With the exception of wood density, which is commonly studied and generally shows moderate to high heritability (Yanchuk *et al.*, 1984; Pliura *et al.*, 2007; El-Kassaby *et al.*, 2011), little is known about the genetic parameters for other wood-related traits in poplar. This article provides a comprehensive assessment of wood traits considered to be key to improve the utility of *Populus* spp. for industrial applications.

The availability of informative molecular markers provides the opportunity to estimate variance components and genetic parameters, such as heritability and genetic merit, in wild populations (Ritland, 1996; Frentiu *et al.*, 2008; Blonk *et al.*, 2010). In the past, simple sequence repeats (microsatellites) have been used to construct marker-based relationship matrices (i.e. co-ancestry) to infer effective population size in conservation programs (Oliehock *et al.*, 2006; Nomura, 2008), as well as to estimate heritability in wild populations (Van de Castele *et al.*, 2001; Frentiu *et al.*, 2008; Blonk *et al.*, 2010), with mixed results. The main drawbacks of employing microsatellite markers to estimate reliable genetic parameters is their limited number, and thus limited information, for the construction of reliable relationship matrices (Frentiu *et al.*, 2008). Generally, hundreds of polymorphic markers are needed to precisely estimate relatedness (Visscher *et al.*, 2008). In addition, factors such as sample size, complexity of the population structure, level of relatedness in the population, nonrandom mating and the implementation of phenotypic selection (i.e. truncation selection), individually or in concert, can affect the precision of the relatedness coefficient and, consequently, all ensuing genetic parameters, including heritability estimates (Van de Castele *et al.*, 2001). However, regardless of the accuracy of genetic parameters, breeding value estimates based on a molecular-based relationship matrix can still provide useful information for selection (Blonk *et al.*, 2010). The development of high-throughput

single nucleotide polymorphism (SNP) genotyping platforms promises drastic improvements in the estimates of genetic parameters generated from relationship-based methods. The constantly increasing number of available SNPs is expected to yield more accurate genetic parameters, thus increasing the expected selection efficiency of breeding programs in either wild or artificial populations with completely unknown or partially known pedigrees.

Furthermore, once a reasonable number of SNPs are available, the resulting, more accurate, genomic relationship matrices can, in turn, substitute the conventional pedigree matrices in best linear unbiased prediction (BLUP) analyses. Hence, the genetic parameters, such as additive genetic variance, heritability and genetic correlations, can be estimated with greater accuracy. This is known as the genomic best linear unbiased prediction (GBLUP) estimation method (VanRaden, 2008; Nadaf & Pong-Wong, 2011). However, it should be stated that the efficiency of selection programs that are focused on multiple traits is not only dependent on the extent of genetic control of these traits individually (i.e. heritability estimates), but also on the genetic correlations among them (Falconer & Mackay, 1996), which are driven by the relative fitness of the traits (Lande & Arnold, 1983). This study employs a dense SNP marker array to genotype 9342 SNPs to construct a genomic relationship matrix among 334 *P. trichocarpa* individuals with unknown genealogy, and to simultaneously estimate the narrow-sense heritabilities and genetic correlations among 17 cell wall chemistry and ultrastructural attributes that are important to the use of poplar as an industrial feedstock. Our results clearly demonstrate the substantial variation in these diverse traits in natural populations, and provide the underpinnings for breeding strategies to capitalize on genetic and phenotypic correlations among traits.

Materials and Methods

Black cottonwood population and wood core sampling

Complete details of the studied black cottonwood (*Populus trichocarpa* Torr. & Gray) population are described in Xie *et al.* (2009). In 2008, a 10-mm increment core was taken at breast height from each of 334 9-yr-old trees (originating from 105 subpopulations) in a north–south direction. The northern half of the core was used for ultrastructure determinations, as well as for wood chemistry, whereas the southern half of the core was used for fiber quality assessments. The geographical distribution of the subpopulations spanned 15° in latitude (range, 44.0–58.6°).

Phenotypic trait data

Wood chemical composition Holocellulose, alpha cellulose and hemicellulose amounts were determined according to Browning (1967), with minor modifications. In brief, the woody material was ground using a Wiley mill to pass a 40-mesh screen, and then extracted overnight in hot acetone and allowed to dry overnight at 50°C. Aliquots of 3.5 ml of solution A (60 ml glacial acetic

acid + 1.3 g l⁻¹ NaOH) and 1.5 ml of 20% sodium chlorite solution (20 g NaClO₂ in 80 ml distilled water) were added to exact amounts (150 mg) of extract-free wood meal to initiate the chlorite delignification. The reaction tube was tightly sealed and then gently shaken at 50°C for 16 h. The reaction was quenched by placing the tubes in an ice bath, and the reaction solution was then thoroughly removed by pipetting whilst not disturbing the settled reacted wood meal. This procedure was repeated. Finally, the reacted wood meal was transferred to a preweighed coarse sintered crucible, and washed twice with 50 ml of 1% glacial acetic acid (under vacuum suction), followed by a wash with 10 ml of acetone under applied vacuum, and then dried at 50°C overnight to obtain the holocellulose yield. To obtain the alpha cellulose content of the woody material, alkaline extractions using two different sodium hydroxide extractions were performed sequentially to remove the hemicelluloses. Exact weights of 100 mg of holocellulose were transferred to a small beaker and left at room temperature for 30 min to allow moisture equilibration. To this, 8 ml of 17.5% NaOH (from sodium hydroxide 50% w/w) was added and the material was left to react for 30 min. Then, 8 ml of distilled water were added and the material was stirred for 1 min, and left to react for 29 min. The reaction solution was carefully removed, and the process was repeated with fresh reactants. After the second reaction, all retentate was filtered through a preweighed coarse sintered crucible by washing with distilled water (3 × 50 ml). Subsequently, the reaction was neutralized by soaking in 1.0 M acetic acid for 5 min. After washing with distilled water (3 × 50 ml), the material was dried at 50°C overnight to obtain the alpha cellulose content. Subtracting the alpha cellulose fraction recovered from the holocellulose provided the hemicellulose yield.

Cell wall carbohydrates and total lignin were determined according to Huntley *et al.* (2003), with the concentrations of neutral cell wall-associated carbohydrates (glucose, xylose, mannose, galactose, rhamnose and arabinose) being quantified by high-performance anion exchange liquid chromatography using a Dionex (DX-600; Sunnyvale, CA, USA) high-performance liquid chromatography system equipped with a PA1 (Dionex) column, a pulsed amperometric detector with a gold electrode and a SpectraAS3500 auto injector (Spectra-Physics, Santa Clara, CA, USA). The column was equilibrated with 250 mM NaOH and eluted with deionized water at a flow rate of 1 ml min⁻¹. The total lignin content was calculated as the sum of the Klason lignin and soluble lignin fractions. Lignin and cell wall sugar contents were determined as a percentage of dry wood, based on the initial weight of the wood sample analyzed.

The lignin monomer composition (S : G ratio) was determined according to Robinson & Mansfield (2009), and was analyzed by gas chromatography on a Hewlett Packard 5890 series II instrument (Agilent Technologies, Santa Clara, CA, USA), equipped with an autosampler, splitless injector, flame ionizing detector and a 30-m 5% diphenyl-95% dimethyl polysiloxane-coated RTX-5MS capillary column (inner diameter, 0.25 mm).

Wood physical properties Fiber length was analyzed using the woody material isolated from the southern half of the increment

core. In brief, each sample was macerated in a solution consisting of 30% hydrogen peroxide and glacial acetic acid in a 1 : 1 ratio (Franklin solution) at 70°C for 48 h. Following the reaction, the remaining fibrous material was washed with deionized water until the samples had been neutralized. Subsamples were resuspended in 10 ml of deionized, distilled water and analyzed on a fiber quality analyzer (FQA; Optest Equipment, Inc., Hawkesbury, Ontario, Canada). The fiber length was recorded in millimeters.

Wood density was analyzed using the northern half of the increment core. A 1.67-mm-thick section was precision cut from each increment core using a custom-built twin-blade pneumatic saw. After extraction with acetone, the wood sections were acclimated to 7% moisture content before density analysis. The samples were then scanned from pith to the bark by X-ray densitometry (QTRS-01X; Quintek Measurement Systems Inc., Knoxville, TN, USA). Wood density was obtained by averaging the density measures across all growth rings. Ten randomly selected *P. trichocarpa* samples from the same site, that were not part of the study, were sectioned as described above, their density determined directly by accurately recording the weight over volume and then scanned by X-ray densitometry to establish an extinction coefficient empirically, which was used to estimate the wood density for all samples from the population.

The microfibril angle (MFA) was measured on two growth rings, the first growth ring adjacent to the pith and the most recent growth ring immediately adjacent to the bark, on the precision-cut samples employed for density determination. MFA was determined using a Bruker D8 Discover (Bruker AXS Inc., Madison, WI, USA) wide-angle X-ray diffraction instrument equipped with an area detector (GADDS) to collect the diffraction patterns of the 002 crystal planes of cellulose. These strong 002 diffraction arcs of the cellulose I reflections contain information on all the orientations of microfibrils in the wood sample. The measurements were collected in transmission mode using a CuK α radiation source emitting at a wavelength of 0.154 nm, with X-ray optics $\theta_1 = \theta_2 = 0^\circ$ (θ_1 = source, θ_2 = detector). The diffraction intensities were collected at Bragg angles $2\theta = 20\text{--}24^\circ$. The X-ray diffraction profile was integrated at Chi between -180° and $+180^\circ$. We used the poplar-specific standard curve (MFA = 0.9583 T - 3.8364, $R^2 = 0.9686$) that related the empirically assessed MFA (as described in Ukrainetz *et al.*, 2008) to that quantified by the X-ray diffraction pattern. The mean T value of the two 002 diffraction peaks was used in the equation to determine the MFA.

The specifications of the crystallinity measurements using X-ray diffraction were the same as for MFA measurements, except that $\theta_1 = 17^\circ$. Diffraction intensities were collected at Bragg angles $2\theta = 4\text{--}40^\circ$. The X-ray diffraction profile was integrated at Chi between -180° and 0° using GADDS software (Bruker AXS Inc., Madison, WI, USA). The data were normalized and resolved using the crystallinity calculation method of Vonk (1973). First, the background diffraction signal was subtracted, and then the amorphous curve was fitted to the diffraction pattern; finally, linear regression analysis was conducted to obtain the cell wall crystallinity.

SNP genotyping

The SNP genotyping of 334 individuals from the *P. trichocarpa* population was carried out using an Illumina Infinium[®] genotyping array with a set of 34 131 SNPs in 3543 candidate genes, as described by Geraldes *et al.* (in press). We eliminated SNPs within candidate genes with a minor allele frequency below 0.05, > 10% missing values and an Illumina GeneTrain score below 0.5. These three selection criteria reduced the number of SNPs to 29 233 for the 334 trees genotyped, 9342 of which were used to estimate the genomic pair-wise kinship. These 9342 SNPs were also selected because of their conformity to Hardy–Weinberg expectations. The genotypes of the 334 individuals for 9342 SNPs on the genotyping array are given in Supporting Information Table S1. Details regarding each locus and the genotyping methods can be found in Geraldes *et al.* (in press).

Estimation of heritabilities and genetic correlations between wood traits

The analysis of variance and trait heritability estimation followed the animal model in Henderson (1984) following:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e} \quad \text{Eqn 1}$$

(\mathbf{y} , vector of observations; \mathbf{X} and \mathbf{Z} , incidence matrices relating fixed effects in vector $\boldsymbol{\beta}$ and random effects in vector \mathbf{u} (breeding values) assuming $U \sim N(0, \sigma_a^2)$ and $\text{Var}(\mathbf{u}) = 2\mathbf{K}\sigma_a^2$; \mathbf{e} , vector of residual effects assuming $E \sim N(0, \sigma_e^2)$). \mathbf{K} is the kinship matrix computed on the basis of marker information following Loiselle *et al.* (1995), which substituted the commonly used average numerator relationship matrix \mathbf{A} based on pedigree information (Henderson, 1976) after substituting the negative values with zeros (Yu *et al.*, 2006).

Genetic correlation between pairs of traits was estimated as:

$$r = \frac{\text{cov}_{AB}}{\sqrt{\sigma_{a1}^2 \sigma_{a2}^2}} \quad \text{Eqn 2}$$

where cov_{AB} , σ_{a1}^2 and σ_{a2}^2 are the genetic covariance between traits and the additive genetic variance of both traits estimated in the bivariate model (Falconer & Mackay, 1996). Assuming $\mathbf{Y}_i = \begin{bmatrix} \mathbf{Y}_i^1 \\ \mathbf{Y}_i^2 \end{bmatrix}$ as the response vector for the individual i and that the bivariate model is performed as follows:

$$\mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{u}_i + \mathbf{e}_i \quad \text{Eqn 3}$$

where $\begin{cases} \mathbf{u}_i \sim N(0, G) \\ \mathbf{e}_i \sim N(0, R) \end{cases}$, the covariate matrix of measurements errors is defined by $\mathbf{R} = \boldsymbol{\Sigma} \otimes \mathbf{I}$, where $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{e1}^2 & \sigma_{e1e2} \\ \sigma_{e2e1} & \sigma_{e2}^2 \end{bmatrix}$ and symbol \otimes represents the Kronecker product. The covariance matrix of random effects is $\mathbf{G} = \begin{bmatrix} \sigma_{a1}^2 & \sigma_{a1a2} \\ \sigma_{a2a1} & \sigma_{a2}^2 \end{bmatrix} \otimes 2\mathbf{K}$, where \mathbf{K} is the kinship matrix. With the assumption of mutual independence between \mathbf{u} and \mathbf{e} , $\text{Var}(\mathbf{Y}_i) = \mathbf{Z}_i\mathbf{G}_i\mathbf{Z}_i^T + \mathbf{R}$. Analyses were

performed in ASReml (Gilmour *et al.*, 2002). The test of relationship between matrices of genetic and phenotypic correlations was performed using the ‘mantel.test’ function in R package ‘ncf’ computing (Bjornstad, 2009).

Results

Variation in wood chemistry and ultrastructural traits

To measure trait variation in *P. trichocarpa*, we quantified wood cell wall traits in an unrelated population of 334 individuals (Fig. 1) from 105 subpopulations, grown in a common garden for 9 yr, with provenances spanning 15° latitude. The range of variation in 17 cell wall traits was considerable, and provides substantial opportunity for selection (Table 1). For example, the alpha cellulose and hemicellulose contents ranged between 35% and 49% and 22% and 34%, respectively, whereas total lignin ranged from 15% to 26%, and the molar composition of syringyl lignin monomers ranged from 63% to 79%. Similarly, the glucose content ranged from 41% to 62%, whereas the xylose content ranged from 14% to 24%. The mean average wood density in the population, an important measure for the expected yields, was 407 kg m⁻³. However, across the population, the density ranged between 320 and 528 kg m⁻³ (Table 1). Significant correlations were observed between subpopulation means for MFA, % galactose content in the cell wall and density with latitudinal gradients ($r=0.34$, $P=0.0004$; $r=-0.27$, $P=0.0063$; and $r=0.20$, $P=0.0401$, respectively). Wood samples with lower MFA and higher cell wall galactose content tended to originate from the northern region of the population range. More southern individuals, however, tended to display a higher wood density than individuals originating from the north. No other traits showed significant trends with latitude.

Heritability of wood traits

With the large number of SNPs (9342) used to generate the marker-based relationship matrix (Materials and Methods) and

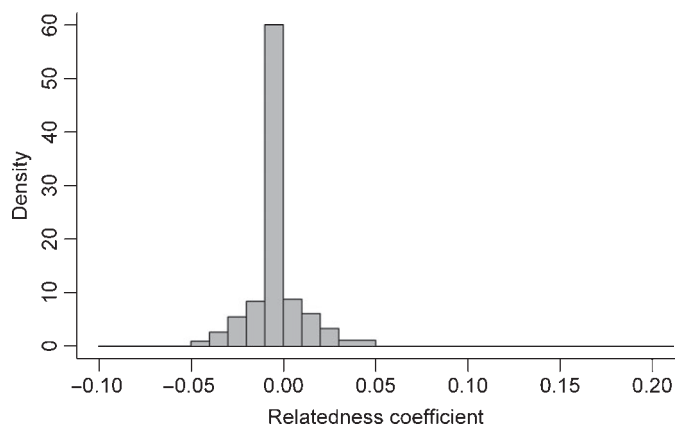


Fig. 1 Histogram of the distribution of relatedness coefficients among *P. trichocarpa* individuals showing pair-wise relationships clustering around zero, indicative of sparse natural population sampling.

Table 1 The mean, standard deviation (SD), minimum (MIN) and maximum (MAX) values for 12 xylem chemistry and five ultrastructural traits in 9-yr-old *Populus trichocarpa* grown in a common-garden plot

Traits	Mean	SD	MIN	MAX	<i>n</i>
Arabinose (%)	0.593	0.118	0.369	1.211	334
Galactose (%)	0.298	0.212	0.001	1.114	333
Glucose (%)	50.753	3.327	40.662	61.661	334
Xylose (%)	18.527	1.553	13.879	24.220	334
Mannose (%)	2.974	0.454	1.612	4.602	334
Insoluble lignin (%)	18.300	1.533	10.204	23.138	334
Soluble lignin (%)	2.528	0.407	1.532	3.524	333
Total lignin (%) ¹	20.852	1.489	14.659	25.724	333
Syringyl monomers (%)	71.495	2.349	62.784	79.231	332
Holocellulose (%) ¹	69.159	2.165	62.825	75.203	334
Alpha cellulose (%)	40.828	2.069	35.087	48.522	334
Hemicellulose (%)	28.331	1.777	22.008	33.652	334
Fiber length (mm)	0.895	0.114	0.443	1.168	310
MFA bark (°)	17.775	2.121	10.418	24.553	330
MFA pith (°)	19.618	1.918	13.653	25.512	330
Average wood density (kg m ⁻³)	406.805	33.569	320.420	528.230	330
Cellulose crystallinity (%)	49.357	3.363	42.017	61.728	319

¹Composite trait (total lignin = soluble + insoluble lignin; holocellulose = alpha cellulose + hemicelluloses). MFA bark, microfibril angle at most recent growth ring; MFA pith, microfibril angle at first growth ring; *n*, sample size; %, dry wood-based percentage of wood components.

the sample size ($n = 334$) studied, we were able to estimate significant narrow-sense heritability (h^2) for 16 of the 17 traits examined (Table 2). The significance of h^2 estimates was determined by the magnitude of their standard errors (SEs). In addition, no sex effect was considered, as it was assumed that this effect did not appreciably contribute to the variance in the traits under study (as the trees were only 9 yr of age, not all individuals were sexually reproductive and, hence, their genders were unknown; Q. Cronk, University of British Columbia, pers. comm.). With the exception of the holocellulose (combination of alpha cellulose and hemicellulose) content, which showed a low h^2 (0.11; SE = 0.135), the remaining h^2 estimates (Table 2) ranged from 0.24 (fiber length) to 0.97 (soluble lignin). High h^2 estimates (> 0.6) were detected for many of the cell wall chemistry traits, including arabinose, galactose, xylose, insoluble lignin and total lignin, whereas the wood average density showed the highest h^2 estimate among the ultrastructural traits (0.56).

Phenotypic and genetic correlations

We next determined the extent of trait–trait genetic (r_g) and phenotypic (r_p) correlations. These analyses showed that 78 (57%) and 61 (45%) of 136 possible phenotypic or genetic trait pairs, respectively, were significant (Tables 3–6). We identified 50 trait pairs for which both phenotypic and genetic correlations were significant; in addition, 11 trait pairs showed significant genetic correlations, but no significant phenotypic correlations, and 28 showed significant phenotypic correlations, but no significant genetic correlations (Tables 3–6). Matrix similarities between genetic and phenotypic correlations were determined using the

Table 2 Estimated narrow-sense heritability (h^2) with standard error (SE) for 12 xylem chemistry and five ultrastructural traits in 9-yr-old *Populus trichocarpa* grown in a common-garden plot

Traits	h^2	SE
Arabinose (%)	0.685	0.149
Galactose (%)	0.643	0.139
Glucose (%)	0.456	0.161
Xylose (%)	0.689	0.156
Mannose (%)	0.385	0.166
Insoluble lignin (%)	0.657	0.163
Soluble lignin (%)	0.965	0.142
Total lignin (%) ¹	0.652	0.163
Syringyl monomers (%)	0.388	0.153
Holocellulose (%) ¹	0.110	0.135
Alpha cellulose (%)	0.416	0.157
Hemicellulose (%)	0.347	0.149
Fiber length (mm)	0.240	0.177
MFA bark (°)	0.441	0.167
MFA pith (°)	0.407	0.164
Average wood density (kg m ⁻³)	0.558	0.166
Cellulose crystallinity (%)	0.399	0.174

¹Composite trait (total lignin = soluble + insoluble lignin; holocellulose = alpha cellulose + hemicelluloses). MFA bark, microfibril angle at recent growth ring; MFA pith, microfibril angle at first growth ring; %, dry wood-based percentage of wood components.

Mantel test with 1000 replications to uncover the patterns of correlation for 12 wood chemistry ($r = 0.82$) and five ultrastructural ($r = 0.80$) traits, and all tests were significant, indicating similar magnitude and trajectory between genetic and phenotypic correlations. To compare the correlation pattern between wood chemistry and ultrastructural traits, we conducted an overall Mantel test for the entire correlation matrix (i.e. 17 traits), which included both phenotypic and genetic correlations, and obtained a similar correlation pattern ($r = 0.76$). This was required as the wood chemistry and ultrastructural trait matrices are not symmetrical; hence, the overall matrix test was performed.

We then used network analysis (Waitt & Levin, 1998) to visually represent the structures of phenotypic and genetic correlations (Fig. 2a,b). Strong and positive phenotypic and genetic correlations existed between clusters of lignin chemistry traits (insoluble lignin, soluble lignin, total lignin) and hemicellulose traits (xylose, mannose, hemicellulose) (cluster I, Fig. 2). A cluster of positive correlations involving glucose, density, alpha cellulose and holocellulose contents was also observed (cluster II), and the genetic correlations for this cluster (Fig. 2b) were much stronger than their respective phenotypic correlations (Fig. 2a).

Although glucose and xylose traits were (weakly) positively correlated, both phenotypically and genetically ($r_p = 0.119$, $r_g = 0.295$), alpha cellulose and density were negatively correlated with xylose, and total and insoluble lignin traits (Fig. 2a,b). By contrast, strong positive genetic correlations were found between alpha cellulose and wood density. The strongest correlations observed among wood chemistry traits were the negative correlation between alpha cellulose and hemicellulose content ($r_g = -0.809$; $r_p = -0.374$) and the positive correlation between xylose and mannose content ($r_g = 0.699$; $r_p = 0.484$). Strong

Table 3 Phenotypic (below diagonal) and genetic (above diagonal) correlations (r_p and r_g , respectively) for xylem chemistry traits

Traits	Arabinose (%)	Galactose (%)	Glucose (%)	Xylose (%)	Mannose (%)	Insoluble lignin (%)	Soluble lignin (%)	Total lignin ¹ (%)	Syringyl (%)	Holocellulose ¹ (%)	Alpha cellulose (%)	Hemicellulose (%)
Arabinose (%)	1											
Galactose (%)	-0.475	1										
Glucose (%)	0.016ns	0.299	1									
Xylose (%)	-0.188	-0.250	0.119	1								
Mannose (%)	-0.038ns	-0.242	-0.025ns	0.484	1							
Insoluble lignin (%)	0.078ns	-0.210	-0.241	0.141	0.289	1						
Soluble lignin (%)	0.038ns	-0.220	-0.019ns	0.419	0.206	0.952	1					
Total lignin ¹ (%)	0.044ns	-0.276	-0.259	0.250	0.316	0.962	0.184	1				
Syringyl (%)	0.054ns	0.046ns	0.067ns	-0.010ns	-0.014ns	-0.254	0.273	-0.176	1			
Holocellulose ¹ (%)	0.160	0.072ns	0.114	-0.117	-0.067ns	-0.181	0.041ns	-0.212	0.219	1		
Alpha cellulose (%)	0.063ns	0.313	0.281	-0.370	-0.309	-0.309	-0.108ns	-0.398	0.167	0.648	1	
Hemicellulose (%)	0.122	-0.277	-0.189	0.288	0.235	0.140	0.174	0.202	0.074ns	0.463	-0.374	1

¹Composite trait (total lignin = soluble + insoluble lignin; holocellulose = alpha cellulose + hemicelluloses). Bold, significant; DNC, did not converge; ns, not significant at 5% level (for r_p) or not significant if $r_g < SE$.

genetic correlations were also found between MFA (most recent growth ring) and holocellulose (negative; $r_g = -0.532$), between MFA (both growth rings) and arabinose (positive; $r_g = 0.561$ and 0.671), and between cellulose crystallinity and fiber length (positive; $r_g = 0.588$). In addition, cellulose crystallinity showed strong positive correlations with insoluble and total lignin (both phenotypically and genetically; Fig. 2a,b). MFA (most recent growth ring) again showed a positive genetic correlation with density ($r_g = 0.260$), and a negative genetic correlation with glucose ($r_g = -0.479$), whereas density and glucose were positively correlated ($r_g = 0.527$; Fig. 2b), suggesting a trade-off between increased MFA and glucose content.

Discussion

To our knowledge, this is the first forest tree study to comprehensively analyze hundreds of natural accessions of *Populus* spp. (poplar) for xylem chemistry and ultrastructural traits simultaneously. Previous studies evaluating wood trait variation in large poplar populations have employed other analytical techniques to assess some of the wood chemistry traits evaluated in this study. These include near-infrared reflectance (NIR) and pyrolysis molecular beam mass spectrometry (PyrMBMS) (Wegrzyn *et al.*, 2010; Studer *et al.*, 2011). Although such techniques have the advantage of being rapid and amenable to high-throughput analysis, a drawback of both is that they provide only a limited number of trait estimates, and are based on training sets derived from wet chemistry analyses similar to the techniques used in the current study. For example, PyrMBMS requires precalibrations using wet chemistry and quantifies wood chemistry traits indirectly, and, although rapid and accurate, does not disclose the make-up of five- and six-carbon sugars, and only offers total amounts of each (e.g. glucose derived from hemicellulose and alpha cellulose). Although commonly applied, neither offers comprehensive insights into the chemistry of cell wall traits (Novaes *et al.*, 2009). By contrast, we applied the direct measurements of wood chemistry traits using wet laboratory approaches and direct measures of ultrastructure. Although laborious, these approaches provide the most direct and accurate phenotypic estimates of quantitative wood chemistry and ultrastructural traits and trait variation, making them highly amenable to the study of wood phenotype–genotype correlations in the *P. trichocarpa* populations studied.

Wood chemistry and ultrastructural traits show considerable variability

Wood quality is defined by the intrinsic morphological and chemical components of wood and, as such, plays a critical role in determining the overall economic value of a woody stem. In addition to growth rate, properties such as wood density, fiber length, MFA and cell wall chemistry are all industrially important traits, and ultimately can be used to select superior clonal material. However, trait selection is ultimately dependent on the end use of woody biomass. In this study, we present a comprehensive assessment of 334 *P. trichocarpa* genotypes spanning much of the

Table 4 Phenotypic (below diagonal) and genetic (above diagonal) (r_p and r_g , respectively) correlations for ultrastructural traits

Traits	Fiber length (mm)	MFA bark (°)	MFA pith (°)	Average wood density (kg m ⁻³)	Cellulose crystallinity (%)
Fiber length (mm)	1	-0.261ns	-0.177ns	-0.385ns	0.588
MFA bark (°)	-0.138	1	0.999	0.260	-0.097ns
MFA pith (°)	-0.004ns	0.620	1	0.189ns	-0.347ns
Average wood density (kg m ⁻³)	-0.015ns	0.100ns	0.140	1	-0.472
Cellulose crystallinity (%)	-0.019ns	0.066ns	0.029ns	-0.272	1

Bold, significant; MFA bark, microfibril angle at recent growth ring; MFA pith, microfibril angle at first growth ring; ns, not significant at 5% level (for r_p) or not significant if $r_g < SE$.

Table 5 Phenotypic correlations (r_p) for xylem chemistry and ultrastructural traits

Traits	Fiber length (mm)	MFA bark (°)	MFA pith (°)	Average wood density (kg m ⁻³)	Cellulose crystallinity (%)
Arabinose (%)	0.087ns	0.137	0.120	-0.120	-0.066ns
Galactose (%)	-0.096ns	-0.194	-0.185	0.143	-0.088ns
Glucose (%)	0.161	-0.108	-0.014ns	0.122	-0.238
Xylose (%)	-0.052ns	0.178	0.149	-0.014ns	0.125
Mannose (%)	-0.001ns	-0.002ns	-0.041ns	-0.065ns	0.222
Insoluble lignin (%)	-0.094ns	-0.001ns	0.002ns	-0.215	0.187
Soluble lignin (%)	0.054ns	0.170	0.102ns	-0.083ns	0.076ns
Total lignin ¹ (%)	-0.077ns	0.053ns	0.029ns	-0.247	0.205
Syringyl (%)	-0.064ns	0.035ns	-0.051ns	0.087ns	-0.178
Holocellulose ¹ (%)	-0.034ns	-0.015ns	-0.049ns	0.088ns	-0.043ns
Alpha cellulose (%)	0.085ns	-0.122	-0.113	0.173	-0.149
Hemicellulose (%)	-0.139	0.124	0.072ns	-0.093ns	0.120

¹Composite trait (total lignin = soluble + insoluble lignin; holocellulose = alpha cellulose + hemicelluloses). Bold, significant; MFA bark, microfibril angle at recent growth ring; MFA pith, microfibril angle at first growth ring; ns, not significant at 5% level.

Table 6 Genetic correlations (r_g) for xylem chemistry and ultrastructural traits

Traits	Fiber length (mm)	MFA bark (°)	MFA pith (°)	Average wood density (kg m ⁻³)	Cellulose crystallinity (%)
Arabinose (%)	0.137ns	0.561	0.671	0.010ns	-0.279
Galactose (%)	-0.233ns	-0.788	-0.441	0.232	0.176ns
Glucose (%)	0.161ns	-0.479	-0.094ns	0.527	-0.146ns
Xylose (%)	0.335	0.016ns	0.092ns	-0.227	0.068ns
Mannose (%)	0.389ns	-0.070ns	-0.015ns	-0.121ns	-0.115ns
Insoluble lignin (%)	-0.279ns	-0.034ns	-0.144ns	-0.430	0.381
Soluble lignin (%)	DNC	0.368	0.134ns	-0.089ns	0.158ns
Total lignin ¹ (%)	-0.231ns	0.191ns	-0.012ns	-0.418	0.447
Syringyl (%)	-0.228ns	0.166ns	0.056ns	0.228ns	-0.655
Holocellulose ¹ (%)	0.450ns	-0.532	-0.149ns	0.668	-0.301ns
Alpha cellulose (%)	0.142ns	-0.449	-0.127ns	0.735	-0.180ns
Hemicellulose (%)	0.120ns	0.112ns	0.079ns	-0.337	0.082ns

¹Composite trait (total lignin = soluble + insoluble lignin; holocellulose = alpha cellulose + hemicelluloses). Bold, significant; DNC, did not converge; MFA bark, microfibril angle at recent growth ring; MFA pith, microfibril angle at first growth ring; ns, not significant if $r_g < SE$.

native range of the species and, based on extensive phenotypic and genotypic diversity, conclusively demonstrate the potential for the selection of superior genotypes for targeted breeding for both bioenergy applications and traditional industrial uses.

Traditionally important wood quality characteristics showed substantial variation, with wood density ranging from 320 to 528 kg m⁻³ and MFA from 10.4 to 25.5°. Wood density has a considerable influence on the strength, machinability and rates of

conversion of solid wood products, and affects both the yield and fiber properties of the pulp produced (Mansfield *et al.*, 2007). Moreover, wood density influences significantly the overall biomass yield of trees (when growth parameters, such as volume, remain constant), which can affect the bioenergy applications of the total harvested material from a given stand. The S₂ MFA has been shown to influence the longitudinal and transverse shrinkage in sawn lumber during drying operations, resulting in a

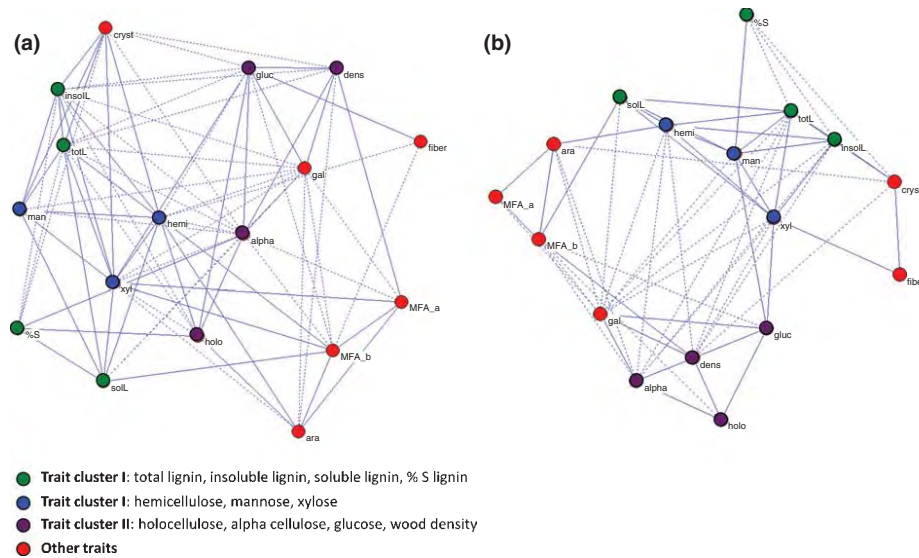


Fig. 2 Correlation structure for significant phenotypic and genetic correlations among 17 wood traits in *Populus trichocarpa*. (a) Phenotypic correlations; (b) genetic correlations. Correlations were computed as outlined in the Materials and Methods section; the social network was visualized using Pajek and the Fruchterman approach (<http://vlado.fmf.uni-lj.si/pub/networks/pajek/>); dashed lines represent negative correlations and solid lines represent positive correlations; the length of an edge between two vertices/traits indicates the relative strength of this pair-wise correlation within the network; arabinose (%), ara; galactose (%), gal; glucose (%), gluc; xylose (%), xyl; mannose (%), man; insoluble lignin (%), insolL; soluble lignin (%), solL; total lignin* (%), totL; syringyl monomers (%), %S; holocellulose* (%), holo; alpha cellulose (%), alpha; hemicellulose (%), hemi; fiber length (mm), fiber; microfibril angle (MFA) of most recent growth ring ($^{\circ}$), MFA_b; MFA of first growth ring ($^{\circ}$), MFA_a; average wood density (kg m^{-3}), dens; cellulose crystallinity (%), cryst. *Composite trait. Traits in cluster I and cluster II, as described in the text, are color coded as indicated.

substantially greater proportion of defects (when the angle is larger). In addition, MFA is also an important determinant of the overall wood strength and elasticity (Mansfield *et al.*, 2007, 2009). For example, in both *Pinus taeda* (Cramer *et al.*, 2005) and *Eucalyptus delegatensis* (Evans & Ilic, 2001), MFA has been shown to explain a large portion of the variation of the longitudinal modulus of elasticity.

Fiber length also varied dramatically, from 0.44 to 1.17 mm. It is well recognized that paper properties are highly correlated with inherent fiber morphology, such as fiber length, fiber diameter and wall thickness (coarseness), with fiber length contributing the greatest influence (Mansfield & Weineisen, 2007). Hardwood fibers generally impart good optical properties and act as a filler at a small percentage to softwood-derived carrier, or reinforcing fibers, in the manufacture of fine paper. It has been shown that paper strength is generally enhanced by longer fibers, and the importance of long fibers is predominant in sheets with low bonding strength, such as those made of only slightly beaten pulp or wet sheets (Mansfield *et al.*, 2004). Variations in wood chemistry also impart significant effects on overall pulp yield and quality, as well as the requisite energy and chemicals required to attain target residual lignin (H-factor) content before bleaching during chemical pulping (Stewart *et al.*, 2006).

We also found substantial variation in total lignin content (14.7–25.7%) and mol.% syringyl monomers (62.8–79.2), both of which have been shown to influence the pulping efficiency of poplar (Huntley *et al.*, 2003; Stewart *et al.*, 2006). As shown previously (Stewart *et al.*, 2009), poplar lignin with elevated syringyl content is less condensed, has more β -O-4 bonds per aromatic subunit and has a lignin fraction with a lower average molecular

mass. Wood with such lignin generates pulp with higher yield and lower kappa values (residual lignin). Similarly, wood traits that are of particular interest for chemical modification during pulping operations are also of interest in the production of biofuels from lignocellulosic feedstocks, particularly those that aim to use biochemical unit operations (i.e. enzymatic hydrolysis following a pretreatment). For example, biomass pretreatment efficiency is influenced by total lignin content, S:G ratio and hemicellulose content and distribution, whereas carbohydrate saccharification efficacy and yield are affected by wood density, % alpha cellulose, % glucose and cell wall crystallinity (Wegrzyn *et al.*, 2010; Studer *et al.*, 2011; Mansfield *et al.*, 2012). Thus, the variation in these lignin traits provides a potential opportunity for genetic improvement of poplar for both traditional and bio-energy uses.

The observed level of variability in wood traits in the poplar populations studied (Table 1) mirrors that reported in previous studies of poplar (Kumar *et al.*, 2009; Sannigrahi *et al.*, 2010; Wegrzyn *et al.*, 2010; Studer *et al.*, 2011). The 334 individuals in the *P. trichocarpa* collection have provenances spanning almost 15° latitude (Xie *et al.*, 2009). Unlike one study of *Eucalyptus globulus* accessions collected over a geographical range of 5° latitude (Stackpole *et al.*, 2011), we observed no significant correlation between the molar composition of syringyl lignin monomers, and thus the S:G monolignol ratio, and the geographical distribution of the subpopulation provenances from which the wood was sampled, with respect to latitude. However, both lignin content and syringyl subunit composition are important for the adaptation of trees to the environment, because lignin performs important roles in stem integrity (strength), hydraulic

conductivity and defenses against biotic challenges (Stackpole *et al.*, 2011). Thus, these properties may represent important fitness traits. We identified a significant trend for average wood density with latitude, as reported in the same *Eucalyptus* study (Stackpole *et al.*, 2011). Wood density, however, was shown to be of less importance for a tree's fitness (Pliura *et al.*, 2007). Density also has the advantage of high repeatability in clonal trials; hence, high genetic gains can be relatively easily achieved in breeding and selection for this trait. Interestingly, both MFA and % cell wall galactose showed a clear geographical cline. Based on the high correlations on an individual tree basis ($r=0.5$) of the monomeric sugars galactose, arabinose and rhamnose, we assumed that the determined galactose fraction in the cell wall originated from arabinogalactan (AG), arabinogalactan proteins (AGPs) and rhamnagalacturonan I (RGI, pectin), but not from galactomannan (hemicellulose). Recently, it has been shown that a noncellulosic polysaccharide matrix, consisting of hydrophilic AG, AGP and RGI, is present within the G-layer of tension wood (reviewed in Mellerowicz & Gorshkova, 2012). It is likely that these polymers contribute to the higher moisture retention associated with these cells. However, when the tension wood is dried and the water is removed rapidly, the resultant wood displays a greater level of shrinkage and causes substantial deformation of the cellulosic microfibrils, imparting negative effects on wood quality (Bowling & Vaughn, 2008). In addition, the MFA of the G-layer of tension wood is characteristically low, often approaching 0° in extreme cases (Mellerowicz & Gorshkova, 2012). Consistent with these features, the current findings clearly demonstrate a high negative genetic correlation between MFA and galactose content (Table 6). Furthermore, there is a strong negative correlation between MFA and alpha cellulose (Table 6), and a strong positive correlation between alpha cellulose and galactose (Table 3). Thus, our results suggest that individuals from the north, which have lower MFAs and higher cell wall galactose content, may contain a higher proportion of tension wood, which is also characterized by higher cellulose crystallinity (Mellerowicz & Gorshkova, 2012). In general, northern trees display faster growth rates, but have a shorter growing season (Soolanayakanahally *et al.*, 2009), and deposit tension-like wood with more AGP, more alpha cellulose and lower MFA. By contrast, southern trees show slower growth rates, but have a longer growing season, resulting in the deposition of a greater proportion of latewood-type fibers known to have higher density, which is a function of the smaller lumen diameters, as growth slows and the demand for water dissipates.

Wood chemistry and ultrastructural traits show moderate to high heritability

The use of the 'animal model' has been widely applied in evolutionary genetics studies to obtain reliable estimates for the additive genetic effects in natural populations with a mixture of relationships between individuals (Kruuk, 2004). Estimates of the derived strict-sense heritabilities provide accurate assessments of the proportion of the total phenotypic variance explained by the underlying genetic factors, as this genetic variance is mainly

additive for a quantitative trait (Hill *et al.*, 2008). Trait heritability with unknown pedigrees can be inferred by inferring relatedness between pairs of individuals and correlating this genetic information with information on their phenotypic similarities. The (assumed) relationships between individuals dictate the amount of genetic markers needed to reconstruct the relatedness matrix (Visscher *et al.*, 2008). In our case, individuals were largely unrelated (Fig. 1); thus, a large number of polymorphic molecular markers had to be employed to infer the relationships with higher precision. This is feasible with high-throughput genotyping techniques, which were readily available for our study. This approach has been successfully applied in other evolutionary studies, and also in plants (Ritland & Ritland, 1996), to estimate heritability 'in the wild'.

A knowledge of heritability is important and central to the direct prediction of the response of a quantitative trait to selection, its 'evolvability' (Houle, 1992). Changes in allelic frequencies and environmental conditions affect the heritable variation (Hoffmann & Merilä, 1999), and thus heritabilities are not necessarily constant estimates, but are environment specific (i.e. dependent on population sampling) and can even vary over time (Visscher *et al.*, 2008). If selection occurs on multiple traits, the genetic correlations between these traits also need to be known and taken into account. An understanding of the extent of genetic control over a particular trait (i.e. its heritable component) is also invaluable for studies that depend on strong phenotype-genotype correlations to identify individual loci with a significant effect on the trait or aim at the elucidation of the genetic architecture underlying the trait, as in association genetic studies; for example, when a larger percentage of a trait's variance is under genetic control (i.e. high heritability), an SNP marker-associated polymorphism is expected to be more meaningful than an equivalent association with a trait of lower heritability, even if the observed association expresses the same percentage of the trait's variance. Thus, our study represents an excellent resource for *P. trichocarpa* tree improvement strategies and facilitates the prediction of an advantageous phenotype in individuals through SNP markers.

With the exception of holocellulose, for which narrow-sense heritability was low ($h^2 < SE$), all other wood traits investigated were under moderate or high genetic control (Table 2). Moderate to high heritabilities in traits suggest that satisfactory genetic gains could be obtained in these traits and that there exists a prospective potential for genetic tree improvement. The reasons for low h^2 could be higher contributions of nonadditive genetic variance and substantial genome by environment (G × E) interactions. Traits that are highly responsive to environmental variation, such as life history traits, are well known to have low h^2 , because the genetic variance in total fitness is low (Price & Schluter, 1991). Specifically, fitness-related traits have greater contributions of dominance and epistatic variance (Hoffmann & Merilä, 1999), and fitness traits have confirmed low heritability with high residual variability (Houle, 1992). By contrast, the amount of genetic variation for many phenotypes related to fitness can be substantial (Visscher *et al.*, 2008). This suggests higher adaptability of the population for increasing population

fitness, as genetic variation is related to environmental variation (Dobzhansky, 1951). However, the response of every trait to directional selection depends on both estimates: the heritability and the genetic variation. We found that c. 60% of the xylem chemistry traits showed high heritability ($h^2 > 0.5$). By contrast, only one of five measured ultrastructural traits had $h^2 > 0.5$ (density), but moderate levels of narrow-sense heritability were observed for all other ultrastructural traits. Previous studies on *E. globulus* found heritability estimates for density, cellulose and lignin that were comparable with our *P. trichocarpa* results (Stackpole *et al.*, 2011). As with our results, very high heritability for acid-soluble lignin was reported previously in *E. globulus* (Poke *et al.*, 2006). These findings suggest that heritabilities, at least for certain wood traits, allow meaningful comparisons across hardwood species.

Different patterns of correlations among wood traits in *P. trichocarpa*

Genetic correlations between traits can arise from pleiotropic effects or from strong linkage disequilibrium of causative genes. Generally, genetic correlations are expected to be higher than phenotypic correlations because $G \times E$ interactions can be prevalent in the latter. However, low trait heritabilities can manifest extreme genetic correlations relative to their phenotypic counterparts (Cheverud, 1988). The average matrix correlations between genetic and phenotypic correlations observed here were similar to those reported previously for plants (Waitt & Levin, 1998). Comparisons between chemical and ultrastructural traits, as opposed to comparisons within chemical and ultrastructural traits, indicated the presence of stronger correlations within each trait set, as expected, as they tend to be more biologically related. Similarly, a lack of strong correlations between certain traits could be an indication of traits that are functionally or developmentally less related, and are therefore genetically and phenotypically less integrated, but functionally or developmentally less related traits can yield higher degrees of phenotypic correlation (Waitt & Levin, 1998). In any case, information on genetic correlations is invaluable for understanding the correlated response of traits to selection, that is trait co-evolution (Cheverud, 1988). From a practical perspective, this knowledge can be applied to indirectly select for and predict genetic gain for an advantageous phenotype that is difficult or costly to assess otherwise. It can also be applied to select for multiple traits simultaneously. However, the influence of environment on the expression of genetic covariation has also been highlighted (Sgro & Hoffmann, 2004), and needs to be considered when breeding strategies are developed.

Advantageous trait combinations

Genetic correlations provide insights into the evolutionary constraints of a species, and are best studied in the natural environment of the species (Lynch, 1999). However, tree improvement trials are widely established as common gardens to facilitate the testing of genotypes under comparable environmental conditions. Thus, it is possible to model the environmental effects and

distinguish these from the true genetic differences between the individuals. Here, we studied phenotypic traits important for cell wall characteristics in a *P. trichocarpa* common garden. To identify potential advantageous traits that may be combined in single individuals, we focus the following discussion on correlations that involve trait pairs for which both genetic and phenotypic correlations were significant.

We identified two important trait clusters, each involving multiple traits: cluster I, a set of hemicellulose- and lignin-related traits; and cluster II, consisting of cellulose and density traits (Fig. 2, Tables 3, 6). Genetic correlations within each of these clusters were strong and positive (Fig. 3a), but many genetic correlations between traits in the two clusters were negative, particularly between alpha cellulose and wood density (cluster I) and lignin and hemicellulose traits (cluster II; Fig. 3b).

These correlations are important considerations when selecting for desired phenotypic combinations in breeding (e.g. for improved biofuel traits). In this case, an important target for genetic improvement is the recalcitrance of plant cell walls to deconstruction. The natural resistance of wood cell walls to biological conversion (enzymatic hydrolysis) is attributed to the heterogeneity of the biomass components and the occlusion of cellulose microfibrils by lignin and hemicelluloses, and the resulting inaccessibility of hydrolytic enzymes to the cellulose surface area (Mansfield *et al.*, 1999; Mosier *et al.*, 2005). Although xylan and lignin can alternatively serve as resources for biomaterials, increased bioethanol yield generally requires feedstocks with lowered lignin and xylan content, fewer or modified hemicellulose–lignin interactions and above average syringyl lignin fraction ($S : G \geq 2.0$) (Studer *et al.*, 2011; Mansfield *et al.*, 2012). Thus, the identification of a concerted negative genetic correlation of lignin and hemicellulose content with density and alpha cellulose content (Fig. 3, Tables 3, 6) appears to be highly advantageous for breeding strategies directed at optimizing poplar as a lignocellulosic feedstock for bioethanol production. The data suggest that breeding for lower lignin or hemicellulose content could simultaneously select for higher density and alpha cellulose content, and vice versa.

We also found that lignin with a higher mol.% syringyl subunits is negatively correlated with the amount of insoluble and total lignin in the cell wall. These findings are consistent with those observed in *Eucalyptus* (Stackpole *et al.*, 2011), and suggest that simultaneous breeding for high mol.% syringyl lignin and low total lignin is feasible.

The properties of cellulose are also important factors to be considered in feedstock development (Mosier *et al.*, 2005). Extensive intra- and intermolecular hydrogen bonds promote the formation of crystalline structures. The intrinsic cellulosic ultrastructure has long been recognized as an important parameter in the efficiency of enzymatic hydrolysis (Mansfield *et al.*, 1999; Chang & Holtzapfle, 2000). We found that low cell wall crystallinity, which is advantageous for biofuel, chemical cellulose and chemical pulping applications, is correlated with lower lignin, higher mol.% syringyl lignin and higher density, suggesting that it should be possible to simultaneously select for these traits in breeding for improved feedstocks. Alternatively, MFA, which is

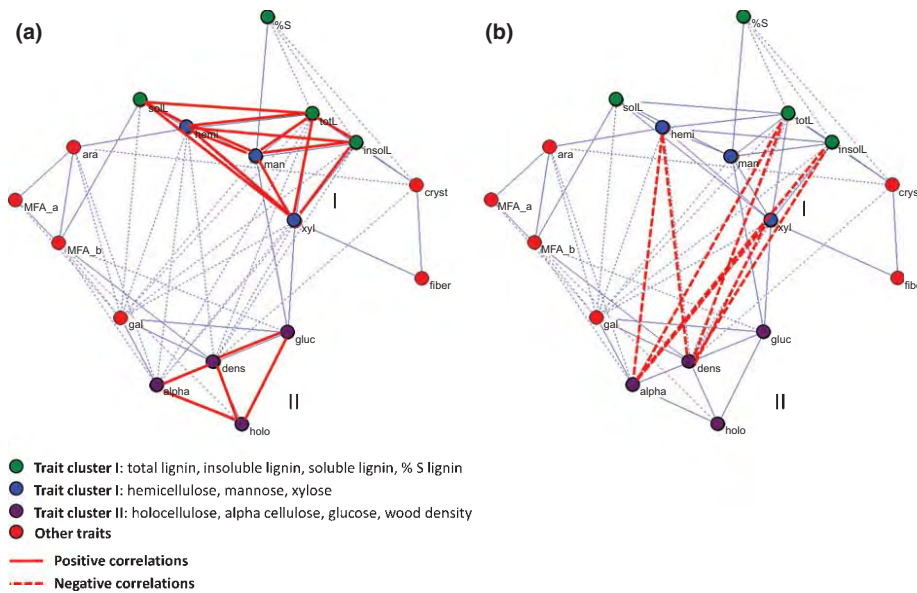


Fig. 3 Summary of genetic correlations between cluster I and cluster II traits relevant to biofuels applications. (a) Positive genetic correlations within cluster I and cluster II traits are highlighted (solid red lines). (b) Negative genetic correlations between cluster I and cluster II traits are highlighted (dashed red lines). Correlations were computed as outlined in the Materials and Methods section; the social network was visualized using Pajek and the Fruchterman approach (<http://vlado.fmf.uni-lj.si/pub/networks/pajek/>); dashed lines represent negative correlations and solid lines represent positive correlations; arabinose (%), ara; galactose (%), gal; glucose (%), gluc; xylose (%), xyl; mannose (%), man; insoluble lignin (%), insOLL; soluble lignin (%), solL; total lignin* (%), totL; syringyl monomers (%), %S; holocellulose* (%), holo; alpha cellulose (%), alpha; hemicellulose (%), hemi; fiber length (mm), fiber; microfibril angle (MFA) of most recent growth ring ($^{\circ}$), MFA_b; MFA of first growth ring ($^{\circ}$), MFA_a; average wood density (kg m^{-3}), dens; cellulose crystallinity (%), cryst. *Composite trait. Traits in cluster I and cluster II, as described in the text, are color coded as indicated.

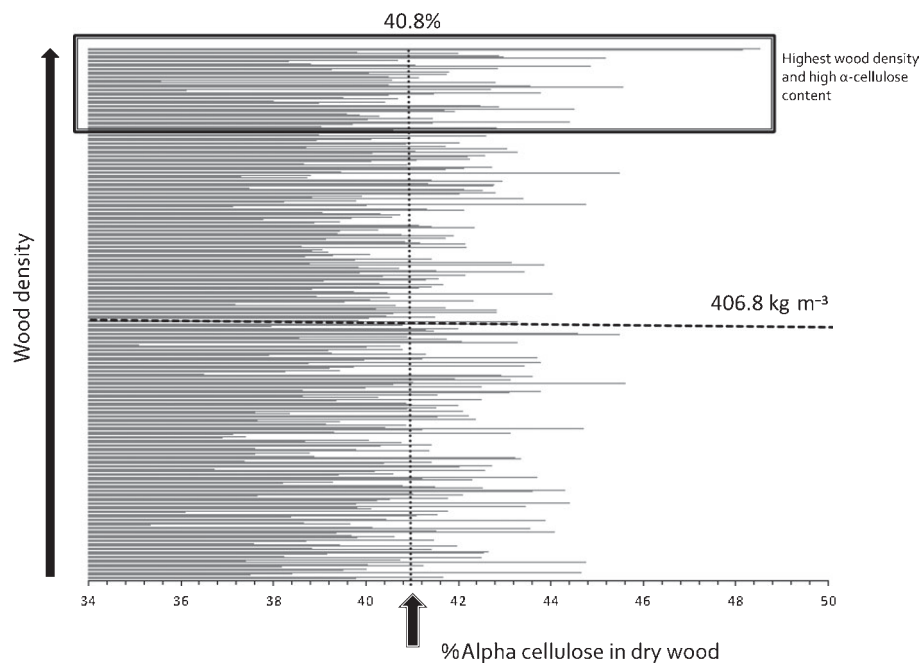


Fig. 4 Representation of the natural variation in wood density and alpha cellulose content in 9-yr-old *Populus trichocarpa* ($n = 330$) grown in a common-garden plot. Individuals were ranked by density and their alpha cellulose contents (% alpha cellulose in dry wood) were plotted. Population means for density (kg m^{-3}) and alpha cellulose (%) are depicted by dashed lines.

the primary angle of orientation of cellulose in the secondary wall relative to the vertical axis of elongating cells, is an important solid wood property. We found that lower MFA, a desirable trait for solid wood applications given its association with wood

stiffness and shrinkage, is correlated with a lower xylose content (Fig. 2a). Therefore, individuals selected on the basis of lower MFA would be predicted to produce offspring with higher glucose/alpha cellulose content and lower xylose content (Table 6). Two

advantageous solid wood and biochemical processing traits for both pulping and bioenergy applications are wood density and alpha cellulose content. Given the strong positive genetic correlation between these two traits (Table 6), and the existence of accessions within the population sampled with high values for both traits (Fig. 4), it appears to be possible to simultaneously select for both traits. Our results provide evidence that natural selection in *P. trichocarpa* already acts in favor of certain desirable industrial trait combinations, and would facilitate the optimization of the feedstock in this species through selective breeding.

Our results also encourage further in-depth studies of trait co-evolution in long-lived perennials. For example, the presence of strong genetic correlations between certain wood traits emphasizes the presence of a common set of genes at the center of these correlations. The identified directionality of these trait correlations would suggest either synergistically or antagonistically pleiotropic gene effects on these traits, dependent on whether the two traits are positively or negatively correlated. Thus, it would be worthwhile to uncover such genetic pleiotropy and identify genes that are significantly correlated with both traits. Such an approach might involve the study of the genetics of genome-wide gene expression to collocate QTLs for gene expression and the quantitative traits under study (Porth *et al.*, 2012). Key to such an undertaking is a knowledge of the extent of interrelation between/among traits of interest, and our network analysis of genetic correlations among traits provides important insights into potential genetic pleiotropy for a comprehensive set of wood traits.

Conclusions

Here, we demonstrate the feasibility of using informative, high-density, biallelic SNP genetic markers to assess the extent of genetic variance and covariance among important wood traits in a natural population of *P. trichocarpa*. We found that the use of a realized kinship matrix in the absence of a structured pedigree was advantageous, an approach that presents a new avenue for the estimation of trait variance components and genetic correlations in natural populations. Our approach to obtain a reliable insight into the network of correlated quantitative attributes in a wild population on the basis of genomic relatedness using high-density marker information represents a novel methodology to the study of trait correlations in plants. Our results provide information on the extensive variation in a comprehensive set of wood cell wall chemistry and ultrastructural characteristics, which are key traits for biofuel applications as well as traditional forest product activities. Our study showed that these industrially important wood traits exhibit significant range-wide phenotypic variation in the species. Furthermore, these findings indicate that there exists substantial opportunity for the use of breeding and selection to capture multiple advantageous trait combinations of interest from natural populations into improved progeny. The respective heritability values for the traits assessed in this study, and data on phenotypic and genetic trait correlations, should help to inform poplar breeding programs that aim to optimize *Populus* lignocellulosic feedstock for solid wood and fiber-based

products, and for chemical processing applications, such as bio-energy.

Acknowledgements

This work was supported by the Genome British Columbia Applied Genomics Innovation Program (Project 103BIO). The authors acknowledge Shofiu Azam and Limin Lao for sample collection, and Eryang Li, Miki Fujita and Teaghan Mayers for technical support. We also acknowledge Lee Gunter and Ann Wymore, Oak Ridge National Laboratory (Oak Ridge TN, USA), for collaboration on the development of the SNP genotyping array and for performing hybridizations to the SNP genotyping array. The SNP array data were generated with support from the BioEnergy Science Center, a US Department of Energy (DOE) Bioenergy Research Center (Office of Biological and Environmental Research in the DOE Office of Science) under Contract No. DE-AC02-05CH11231.

References

- Beaulieu J, Doerksen T, Boyle B, Clement S, Deslauriers M, Beauseigle S, Blais S, Poulin PL, Lenz P, Caron S *et al.* 2011. Association genetics of wood physical traits in the conifer white spruce and relationships with gene expression. *Genetics* 188: 197–214.
- Bjornstad ON. 2009. ncf: spatial nonparametric covariance functions. R package version 1.1-3. <http://CRAN.R-project.org/package=ncf> [accessed on 20 July 2012].
- Blonk RJW, Komen H, Kamstra A, van Arendonk JAM. 2010. Estimating breeding values with molecular relatedness and reconstructed pedigrees in natural mating populations of Common Sole, *Solea solea*. *Genetics* 184: 213–219.
- Bowling AJ, Vaughn KC. 2008. Immunocytochemical characterization of tension wood: gelatinous fibers contain more than just cellulose. *American Journal of Botany* 95: 655–663.
- Browning BL. 1967. *Methods of wood chemistry*. New York, NY, USA: Wiley Interscience Publishers.
- Burns RM, Honkala BH. 1990. *Silvics of North America, Vol. 2, Hardwoods*. Washington, DC, USA: USDA Forest Service Agriculture Handbook.
- Carroll A, Somerville C. 2009. Cellulosic biofuels. *Annual Review of Plant Biology* 60: 165–182.
- Chang VS, Holtzapfel MT. 2000. Fundamental factors affecting biomass enzymatic reactivity. *Applied Biochemistry and Biotechnology* 84–86: 5–37.
- Cheverud JM. 1988. A comparison of genetic and phenotypic correlations. *Evolution* 42: 958–968.
- Cramer S, Kretschmann D, Lakes R, Schmidt T. 2005. Earlywood and latewood elastic properties in loblolly pine. *Holzforschung* 59: 531–538.
- Dinus RJ. 2000. *Genetic modification of short rotation Poplar biomass feedstock for efficient conversion to ethanol*. Bioenergy Feedstock Development Program, Environmental Sciences Division, Oak Ridge National Laboratory, TN, USA: 67.
- Dobzhansky T. 1951. *Genetics and the origin of species*. New York, NY, USA: Columbia University Press.
- El-Kassaby YA, Mansfield SD, Isik F, Stoehr M. 2011. *In situ* wood quality assessment in Douglas-fir. *Tree Genetics & Genomes* 7: 553–561.
- Evans R, Ilic J. 2001. Rapid prediction of wood stiffness from microfibril angle and density. *Forest Products Journal* 51: 53–57.
- Falconer DS, Mackay TFC. 1996. *Introduction to quantitative genetics, 4th edn*. Harlow, UK: Longmans Green.
- Frentiu FD, Clegg SM, Chittock J, Burke T, Blows MW, Owens IPF. 2008. Pedigree-free animal models: the relatedness matrix reloaded. *Proceedings of the Royal Society B: Biological Sciences* 275: 639–647.
- Geraldes A, DiFazio SP, Slavov GT, Ranjan P, Muchero W, Hannemann J, Gunter LE, Wymore A, Grassa C, Farzaneh N *et al.* In press. A 34k SNP

- genotyping array for *Populus trichocarpa*: design, application to the study of natural populations and transferability to other *Populus* species. *Molecular Ecology Resources*, in press.
- Geraldes A, Pang J, Thiessen N, Cezard T, Moore R, Zhao Y, Tam A, Wang S, Friedmann M, Birol I *et al.* 2011. SNP discovery in black cottonwood (*Populus trichocarpa*) by population transcriptome resequencing. *Molecular Ecology Resources* 11: 81–92.
- Gilmour AR, Gogel BJ, Cullis BR, Welham SJ, Thompson R. 2002. *ASReml user guide release 1.0*. Hemel Hempstead, UK: VSN International Ltd.
- Henderson CR. 1976. Simple method for computing inverse of a numerator relationship matrix used in prediction of breeding values. *Biometrics* 32: 69–83.
- Henderson CR. 1984. *Applications of linear models in animal breeding*. Guelph, ON, Canada: University of Guelph.
- Hendriks ATWM, Zeeman G. 2009. Pretreatments to enhance the digestibility of lignocellulosic biomass. *Bioresource Technology* 100: 10–18.
- Hill WG, Goddard ME, Visscher PM. 2008. Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genetics* 4: e1000008.
- Hoffmann AA, Merilä J. 1999. Heritable variation and evolution under favourable and unfavourable conditions. *Trends in Ecology and Evolution* 14: 96–101.
- Houle D. 1992. Comparing evolvability and variability of quantitative traits. *Genetics* 130: 195–204.
- Huntley SK, Ellis D, Gilbert M, Chapple C, Mansfield SD. 2003. Improved pulping efficiency in C4H–F5H transformed poplar. *Journal of Agricultural and Food Chemistry* 51: 6178–6183.
- Jansson S, Douglas CJ. 2007. *Populus*: a model system for plant biology. *Annual Review of Plant Biology* 58: 435–458.
- Kirst M, Myburg AA, De Leon JPG, Kirst ME, Scott J, Sederoff R. 2004. Coordinated genetic regulation of growth and lignin revealed by quantitative trait locus analysis of cDNA microarray data in an interspecific backcross of eucalyptus. *Plant Physiology* 135: 2368–2378.
- Kruuk LE. 2004. Estimating genetic parameters in natural populations using the 'animal model'. *Philosophical Transactions of the Royal Society London, B: Biological Sciences* 359: 873–890.
- Kumar R, Mago G, Balan V, Wyman CE. 2009. Physical and chemical characterizations of corn stover and poplar solids resulting from leading pretreatment technologies. *Bioresource Technology* 100: 3948–3962.
- Lande R, Arnold SJ. 1983. The measurement of selection on correlated characters. *Evolution* 37: 1210–1226.
- Loiselle BA, Sork VL, Nason J, Graham C. 1995. Spatial genetic-structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany* 82: 1420–1425.
- Lynch M. 1999. Estimating genetic correlations in natural populations. *Genetical Research* 74: 255–264.
- Mansfield SD. 2009. Solutions for dissolution – engineering cell walls for deconstruction. *Current Opinion in Biotechnology* 20: 286–294.
- Mansfield SD, Kang KY, Chapple C. 2012. Designed for deconstruction – Poplar trees altered in cell wall lignification improve the efficacy of bioethanol production. *New Phytologist* 194: 91–101.
- Mansfield SD, Kibblewhite RP, Riddell MJC. 2004. Characterisation of the reinforcement potential of different softwood Kraft fibres in softwood/hardwood pulp mixtures. *Wood and Fiber Science* 36: 344–358.
- Mansfield SD, Mooney C, Saddler JN. 1999. Substrate and enzyme characteristics that limit cellulose hydrolysis. *Biotechnology Progress* 15: 804–816.
- Mansfield SD, Parish R, Di Lucca CM, Goudie J, Kang KY, Ott P. 2009. Revisiting the transition between juvenile and mature wood: a comparison of fibre length, microfibril angle and relative wood density in lodgepole pine. *Holzforchung* 63: 449–456.
- Mansfield SD, Parish R, Goudie JW, Kang KY, Ott P. 2007. The effects of crown ratio on the transition from juvenile to mature wood production in lodgepole pine in western Canada. *Canadian Journal of Forest Research* 37: 1450–1459.
- Mansfield SD, Weineisen H. 2007. Wood fibre quality and kraft pulping efficiencies of trembling aspen (*Populus tremuloides* Michx) clones. *Journal of Wood Chemistry and Technology* 27: 135–151.
- Mellerowicz EJ, Gorshkova TA. 2012. Tensional stress generation in gelatinous fibres: a review and possible mechanism based on cell-wall structure and composition. *Journal of Experimental Botany* 63: 551–565.
- Mosier N, Wyman C, Dale B, Elander R, Lee YY, Holtzapple M, Ladisch M. 2005. Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresource Technology* 96: 673–686.
- Nadaf J, Pong-Wong R. 2011. Applying different genomic evaluation approaches on QTLMAS2010 dataset. *BMC Proceedings* 5(Suppl 3): S9.
- Nomura T. 2008. Estimation of effective number of breeders from molecular coancestry of single cohort sample. *Evolutionary Applications* 1: 462–474.
- Novaes E, Osorio L, Drost DR, Miles BL, Boaventura-Novaes CRD, Benedict C, Dervinis C, Yu Q, Sykes R, Davis M *et al.* 2009. Quantitative genetic analysis of biomass and wood chemistry of *Populus* under different nitrogen levels. *New Phytologist* 182: 878–890.
- Oliehock PA, Windig JJ, van Arendonk JAM, Bijma P. 2006. Estimating relatedness between individuals in general populations with a focus on their use in conservation programs. *Genetics* 173: 483–496.
- Pliura A, Zhang SY, MacKay J, Bousquet J. 2007. Genotypic variation in wood density and growth traits of poplar hybrids at four clonal trials. *Forest Ecology and Management* 238: 92–106.
- Poke FS, Potts BM, Vaillancourt RE, Raymond CA. 2006. Genetic parameters for lignin, extractives and decay in *Eucalyptus globulus*. *Annals of Forest Science* 63: 813–821.
- Porth I, White R, Jaquish B, Alfaro R, Ritland C, Ritland K. 2012. Genetical genomics identifies the genetic architecture for growth and weevil resistance in spruce. *PLoS ONE* 7: e44397.
- Pot D, Chantre G, Rozenberg P, Rodrigues JC, Jones GL, Pereira H, Hannrup B, Cahalan C, Plomion C. 2002. Genetic control of pulp and timber properties in maritime pine (*Pinus pinaster* Ait.). *Annals of Forest Science* 59: 563–575.
- Price T, Schluter D. 1991. On the low heritability of life-history traits. *Evolution* 45: 853–861.
- Ralph SG, Chun HJE, Cooper D, Kirkpatrick R, Kolosova N, Gunter L, Tuskan GA, Douglas CJ, Holt RA, Jones SJM *et al.* 2008. Analysis of 4,664 high-quality sequence-finished poplar full-length cDNA clones and their utility for the discovery of genes responding to insect feeding. *BMC Genomics* 9: 57.
- Ritland K. 1996. Estimators for pairwise relatedness and individual inbreeding coefficients. *Genetical Research* 67: 175–185.
- Ritland K, Ritland C. 1996. Inferences about contrast quantitative inheritance based on natural population structure in the yellow monkeyflower, *Mimulus guttatus*. *Evolution* 50: 1074–1082.
- Robinson AR, Mansfield SD. 2009. Rapid analysis of poplar lignin monomer composition by a streamlined thioacidolysis procedure and near-infrared reflectance-based prediction modeling. *Plant Journal* 58: 706–714.
- Sannigrahi P, Ragauskas AJ, Tuskan GA. 2010. Poplar as a feedstock for biofuels: a review of compositional characteristics. *Biofuels, Bioproducts & Biorefining-Biofr* 4: 209–226.
- Simple DE, Vaillant MH, Kang KY, Oh SW, Smith GD, Mansfield SD. 2007. Evaluating the suitability of hybrid poplar clones for the manufacture of oriented strand boards. *Holzforchung* 61: 430–438.
- Sgro CM, Hoffmann AA. 2004. Genetic correlations, tradeoffs and environmental variation. *Heredity* 93: 241–248.
- Shen H, He X, Poovaiah CR, Wuddineh WA, Ma J, Mann DGJ, Wang H, Jackson L, Tang Y, Stewart CN *et al.* 2011. Functional characterization of the switchgrass (*Panicum virgatum*) R2R3-MYB transcription factor PvMYB4 for improvement of lignocellulosic feedstocks. *New Phytologist* 193: 121–136.
- Slavov GT, DiFazio SP, Martin J, Schackwitz W, Muchero W, Rodgers-Melnick E, Lipphardt MF, Pennacchio CP, Hellsten U, Pennacchio L *et al.* 2012. Genome resequencing reveals multiscale geographic structure and extensive linkage disequilibrium in the forest tree *Populus trichocarpa*. *New Phytologist* 196: 713–725.
- Soolanayakanahally RY, Guy RD, Silim SN, Drewes EC, Schroeder WR. 2009. Enhanced assimilation rate and water use efficiency with latitude through increased photosynthetic capacity and internal conductance in balsam poplar (*Populus balsamifera* L.). *Plant, Cell & Environment* 32: 1821–1832.
- Stackpole DJ, Vaillancourt RE, Alves A, Rodrigues J, Potts BM. 2011. Genetic variation in the chemical components of *Eucalyptus globulus* wood. *G3: Genes, Genomes, Genetics* 1: 151–159.

- Stewart JJ, Akiyama T, Chapple CC, Ralph J, Mansfield SD. 2009. The effects on lignin structure of over-expression of ferulate 5-hydroxylase in hybrid poplar. *Plant Physiology* 150: 621–635.
- Stewart JJ, Kadla JF, Mansfield SD. 2006. The influence of lignin chemistry and ultrastructure on the pulping efficiency of clonal aspen (*Populus tremuloides* Michx.). *Holzforschung* 60: 111–122.
- Studer MH, DeMartini JD, Davis MF, Sykes RW, Davison B, Keller M, Tuskan GA, Wyman CE. 2011. Lignin content in natural *Populus* variants affects sugar release. *Proceedings of the National Academy of Sciences, USA* 108: 6300–6305.
- Thumma BR, Southerton SG, Bell JC, Owen JV, Henery ML, Moran GF. 2010. Quantitative trait locus (QTL) analysis of wood quality traits in *Eucalyptus nitens*. *Tree Genetics & Genomes* 6: 305–317.
- Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A *et al.* 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313: 1596–1604.
- Ukrainetz NK, Kang KY, Aitken SN, Stoehr M, Mansfield SD. 2008. Heritability, phenotypic and genetic correlations of coastal Douglas-fir (*Pseudotsuga menziesii*) wood quality traits. *Canadian Journal of Forest Research* 38: 1536–1546.
- Van de Castele T, Galbusera P, Matthysen E. 2001. A comparison of microsatellite-based pairwise relatedness estimators. *Molecular Ecology* 10: 1539–1549.
- VanRaden PM. 2008. Efficient methods to compute genomic predictions. *Journal of Dairy Science* 91: 4414–4423.
- Visscher PM, Hill WG, Wray NR. 2008. Heritability in the genomics era – concepts and misconceptions. *Nature Reviews Genetics* 9: 255–266.
- Vonk CG. 1973. Computerization of Rulands X-ray method for determination of crystallinity in polymers. *Journal of Applied Crystallography* 6: 148–152.
- Waitt DE, Levin DA. 1998. Genetic and phenotypic correlations in plants: a botanical test of Cheverud's conjecture. *Heredity* 80: 310–319.
- Wegrzyn JL, Eckert AJ, Choi M, Lee JM, Stanton BJ, Sykes R, Davis MF, Tsai CJ, Neale DB. 2010. Association genetics of traits controlling lignin and cellulose biosynthesis in black cottonwood (*Populus trichocarpa*, Salicaceae) secondary xylem. *New Phytologist* 188: 515–532.
- Xie CY, Ying CC, Yanchuk AD, Holowachuk DL. 2009. Ecotypic mode of regional differentiation caused by restricted gene migration: a case in black cottonwood (*Populus trichocarpa*) along the Pacific Northwest coast. *Canadian Journal of Forest Research* 39: 519–526.
- Yanchuk AD, Dancik BP, Micko MM. 1984. Variation and heritability of wood density and fiber length of trembling aspen in Alberta, Canada. *Silvae Genetica* 33: 11–16.
- Yu JM, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB *et al.* 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics* 38: 203–208.
- Yuan JS, Tiller KH, Al-Ahmad H, Stewart NR, Stewart CN Jr. 2008. Plants to power: bioenergy to fuel the future. *Trends in Plant Science* 13: 421–429.
- Zhang P, Wu F, Kan X. 2012. Genotypic variation in wood properties and growth traits of triploid hybrid clones of *Populus tomentosa* at three clonal trials. *Tree Genetics & Genomes* 8: 1041–1050.

Supporting Information

Additional supporting information may be found in the online version of this article.

Table S1. Genotypes for 9342 single nucleotide polymorphism (SNP) loci in 334 *Populus trichocarpa* accessions used to construct a kinship matrix

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.