Genes and gene clusters related to genotype and drought induced variation in saccharification potential, lignin content, and wood anatomical traits in Populus nigra

Wildhagen, Henning ; Paul, Shanty; Allwright, Mike; Smith, Hazel; Malinowska, Marta; Schnabel, Sabine; Paulo, Joao; Cattonaro, Federica; Vendramin, Vera ; Scalabrin, Simone ; Janz, Dennis; Douthe, Cyril; Brendel, Oliver; Buré, Cyril; Cohen, David; Hummel, Irène; Le Thiec, Didier; van Eeuwijk, Fred; Keurentjes, Joost; Flexas, Jaume

Published in:
Tree Physiology

DOI:
10.1093/treephys/tpx054

Publication date:
2018

Citation for published version (APA):

Document License
CC BY

General rights
Copyright and moral rights for the publications made accessible in the Aberystwyth Research Portal (the Institutional Repository) are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Aberystwyth Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Aberystwyth Research Portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

tel: +44 1970 62 2400
email: is@aber.ac.uk

Download date: 13. Aug. 2021
Genes and gene clusters related to genotype and drought-induced variation in saccharification potential, lignin content and wood anatomical traits in *Populus nigra*†

Henning Wildhagen1,10, Shanty Paul1, Mike Allwright2, Hazel K. Smith2, Marta Malinowska3, Sabine K. Schnabel4, M. João Paulo4, Federica Cattonaro5, Vera Vendramin5, Simone Scalabrin5, Dennis Janz1, Cyril Douthe6, Oliver Brede17, Cyril Buré7, David Cohen7, Irène Hummel7, Didier Le Thiec7, Fred van Eeuwijk4, Joost J.B. Keurentjes8, Jaume Flexas6, Michele Morgante9, Paul Robson3, Marie-Bénatrice Boget-Triboulot7, Gail Taylor2 and Andrea Polle1,11

1Forest Botany and Tree Physiology, Georg-August University of Goettingen, Büsgenweg 2, 37077 Göttingen, Germany; 2Center for Biological Sciences, University of Southampton, University Road, Southampton SO17 1BI, UK; 3Institute of Biological, Environmental & Rural Sciences (IBERS), Aberystwyth University, Gogerddan, Aberystwyth, SY23EE, UK; 4Biometrics, Wageningen University and Research, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands; 5IGA Technology Services, via Jacopo Linusso 51, 33100 Udine, Italy; 6Universidad de les Illes Balears, Carretera de Valldemossa Km 7.5, 07122 Palma de Mallorca, Illes Balears, Spain; 7EEF, INRA, Université de Lorraine, rue d’Amanse, 54280 Champenoux, France; 8Laboratory of Genetics, Wageningen University and Research, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands; 9Università Di Udine, Istituto di Genomica Applicata, via Jacopo Linusso 51, 33100 Udine, Italy; 10Present address: HAWK University of Applied Sciences and Arts, Faculty of Resource Management, Büsgenweg 1a, 37077 Göttingen, Germany; 11Corresponding author (apolle@gwdg.de)

Received January 17, 2017; accepted May 3, 2017; handling Editor Janice Cooke

Wood is a renewable resource that can be employed for the production of second generation biofuels by enzymatic saccharification and subsequent fermentation. Knowledge on how the saccharification potential is affected by genotype-related variation of wood traits and drought is scarce. Here, we used three *Populus nigra* L. genotypes from habitats differing in water availability to (i) investigate the relationships between wood anatomy, lignin content and saccharification and (ii) identify genes and co-expressed gene clusters related to genotype and drought-induced variation in wood traits and saccharification potential. The three poplar genotypes differed in wood anatomy, lignin content and saccharification potential. Drought resulted in reduced cambial activity, decreased vessel and fiber lumina, and increased the saccharification potential. The saccharification potential was unrelated to lignin content as well as to most wood anatomical traits. RNA sequencing of the developing xylem revealed that 1.5% of the analyzed genes were differentially expressed in response to drought, while 67% differed among the genotypes. Weighted gene correlation network analysis identified modules of co-expressed genes correlated with saccharification potential. These modules were enriched in gene ontology terms related to cell wall polysaccharide biosynthesis and modification and vesicle transport, but not to lignin biosynthesis. Among the most strongly saccharification-correlated genes, those with regulatory functions, especially kinases, were prominent. We further identified transcription factors whose transcript abundances differed among genotypes, and which were co-regulated with genes for biosynthesis and modifications of hemicelluloses and pectin. Overall, our study suggests that the regulation of pectin and hemicellulose metabolism is a promising target for improving wood quality of second generation bioenergy crops. The causal relationship of the identified genes and pathways with saccharification potential needs to be validated in further experiments.

**Keywords**: biofuels, cell wall, drought, genotypic variation, glucose release, lignin, *Populus*, saccharification, wood traits.

†Dedicated to the memory of Carl Douglas.

© The Author 2017. Published by Oxford University Press. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
Introduction

Wood is an attractive feedstock for biofuels because it is a renewable resource that can be produced in a sustainable manner (Polle et al. 2013, Wei and Polle 2016). Wood is composed mainly of cellulose (35–50%), hemicelluloses (15–35%), lignin (15–35%) and pectins (<10%) (Pomion et al. 2001). Glucose can be released from cellulolytic compounds in the cell wall by saccharification and subsequently used for conversion into ethanol (Galbe and Zacchi 2002). Lignin and hemicelluloses and their interactions with pectin or cellulose can negatively impact the efficiency of biofuel production (Mansfield et al. 1999, Nookaraju et al. 2013, Xiao and Anderson 2013). For example, transgenic lines of Populus trichocarpa with suppressed expression of 4-coumarate:coenzyme A ligase contained lower lignin contents and exhibited greater saccharification potential than the wild type (Min et al. 2012). In Populus × canescens, in which cinnamoyl-CoA reductase, a key enzyme for lignin biosynthesis, was down-regulated, the cell walls were more easily digestible, which resulted in higher saccharification and ethanol yield (Acker et al. 2014).

The saccharification potential of wood is also affected by environmental conditions such as drought (Fiasconaro et al. 2012, van der Weijde et al. 2016). Drought brings about changes in cell wall composition including an increased abundance of hemicellulose, pectins and lignin, which results in strengthening of the cell wall (Moore et al. 2008, Le Gall et al. 2015). Fiasconaro et al. (2012) reported that the potential for biofuel conversion in drought-treated alfalfa plants decreased due to a higher lignin content compared with the well-watered plants. On the other hand, in a study comprising 50 different Miscanthus genotypes, drought treatment caused a reduction of cell wall cellulose content and an increase in cell wall hemicellulose content, but nevertheless resulted in a higher efficiency of the conversion of cellulose to glucose during saccharification (van der Weijde et al. 2016). These contrasting findings suggest that the effect of drought on saccharification potential depends largely on the particular species studied, and on the relative changes of cell wall components under drought. In poplar, drought stress has massive consequences for the wood anatomy and cell wall metabolism (Harvey and Van Den Driessen 1997, Arend and Fromm 2007, Beniwal et al. 2010, Fichot et al. 2009, 2010, Schreiber et al. 2011, Cao et al. 2014, Guet et al. 2015, Le Gall et al. 2015), but studies on the effects of drought on the saccharification potential of Populus and the underlying anatomical and molecular responses have not yet been reported, but are greatly needed (Studer et al. 2011).

The saccharification potential is also subject to genetic variation, as shown for Miscanthus (Souza et al. 2015) and poplar (Studer et al. 2011). Similarly, natural genetic variation in lignin content was reported for many taxa, including Arabidopsis thaliana (Capron et al. 2013), eucalyptus (Klash et al. 2010, Elissetche et al. 2011), conifers (Gonzalez-Martinez et al. 2007, Gaspar et al. 2011) and poplars (Wegrzyn et al. 2010, Zhou et al. 2011, Guerra et al. 2013, Porth et al. 2013, Muchero et al. 2015). Given the importance of the genus Populus as a second generation bioenergy crop (Allwright and Taylor 2016), it is of great interest to study whether these variations translate into variation in saccharification potential.

Attempts to improve the saccharification potential of bioenergy crops require insights into the molecular control of wood properties. Several studies using quantitative trait locus (QTL) mapping or association mapping approaches identified candidate genes related to cell wall composition, as well as lignin content and composition (Ranjan et al. 2009, Guerra et al. 2013, Porth et al. 2013, Fahrenkrog et al. 2016). However, such approaches have only rarely been applied to uncover candidate genes or QTLs related to sugar release (Berreton et al. 2010, Muchero et al. 2015). Knowledge on transcriptional networks controlling diverse aspects of wood formation including secondary cell wall formation and biosynthesis of its components has accumulated during recent years (for recent reviews: Zhong et al., 2010, Hussey et al., 2013, Nakano et al., 2015, Ye and Zhong 2015). However, a systematic transcriptome-wide investigation of genes and gene clusters underlying genetic variation in saccharification potential in economically important bioenergy crops such as Populus is missing.

The aim of this study was to identify clusters of co-expressed genes, as well as candidate genes and their putative transcriptional regulators, related to genotype- and drought-induced variation in wood anatomy, lignin content and saccharification potential. We hypothesized (i) that drought results in increased lignification and decreased saccharification, (ii) that genotypes originating from different environments differ in lignin content and saccharification potential and (iii) that these drought and genotype-effects on saccharification potential are underpinned by distinct differences in wood anatomical traits and transcript abundances of genes involved in wood formation and cell wall metabolism, especially lignin biosynthesis.

To this end, we used three genotypes that were selected from large-scale common garden experiments with up to 13 different Populus nigra L. populations (DeWoody et al. 2015, Viger et al. 2016). The P. nigra populations originated from habitats with different climatic conditions across Europe and showed significant genetic differentiation as well as phenotypic variation in growth rates, plant architecture and leaf size (DeWoody et al. 2015, Viger et al. 2016). The most pronounced differences among the phenotypic traits were observed between P. nigra genotypes originating from a dry habitat in Spain and those from a wet habitat in Italy, while a French genotype exhibited intermediate properties (Viger et al. 2016). Here, we exposed these contrasting genotypes to a highly controlled drought treatment, applied RNA sequencing and weighted gene correlation network analysis (WGCNA) to identify novel candidate genes related to wood properties and saccharification potential.
Materials and methods

Plant material

Three genotypes of *P. nigra* L., originating from individual trees of natural populations in France (Drôme 6; FR-6), Italy (La Zelata; IT1) and Spain (Ebro 2; SP-2; see DeWoody et al. 2015) were studied. Cuttings from these genotypes were purchased from INRA Orleans (Orléans, France), where the materials are available for research purposes under the accession numbers N-38 (‘Italy’), RIN2-new (‘Spain’) and 6-J29 (‘France’). The cuttings were planted in 10 l plastic pots filled with a 1:1 (v/v) mixture of peat and sand, amended with a slow-release fertilizer (4 g l⁻¹ of Nutricote T100, 13:13:13 NPK and micronutrients; FERTIL S.A.S., Boulogne Billancourt, France) and 1 g l⁻¹ CaMg(CO₃)₂. Plants were grown in two chambers of a fully automated glasshouse for phenotyping located at Champenoux, France (48°45’09.3”N, 6°20’27.6”E), under natural light conditions with daily maxima of irradiance ranging from 154 to 1011 μmol m⁻² s⁻¹ photosynthetically active radiation. Environmental conditions in the greenhouse were affected by weather conditions, but the temperature was adjusted to not exceed 28 °C. Plants were watered three times a day to 85% of field capacity with an automated watering system.

Drought experiment

After 6 weeks of growth, 32 plants of each genotype were randomly assigned to either a control or a drought treatment. The plants were randomized across the two greenhouse chambers, so that each chamber contained balanced proportions of each genotype and treatment. The position of plants was changed automatically.

Plants were exposed to drought by gradually decreasing the available soil water content (SWC). The control of SWC was based on pot weight and a calibration between volumetric SWC measured by Time Domain Reflectometry and pot weight. Available water was then expressed as relative extractable water content (REWsoil), which is defined as:

\[
\text{REW}_{\text{soil}} = \left( \frac{\text{SWC} - \text{SWC}_{\text{wilting point}}}{\text{SWC}_{\text{field capacity}} - \text{SWC}_{\text{wilting point}}} \right) \times 100\%.
\]

with SWC at wilting point = 3%; SWC at field capacity = 32%. Sixteen control plants per genotype were watered to 85% REWsoil three times a day by the automated system for the whole 5-week period of the experiment. For drought-treated plants, REWsoil was progressively decreased to reach a target level of 20% REWsoil in 2 weeks. The decrease in REWsoil was controlled and adjusted in each individual pot because bigger plants used more water than smaller plants. Thereby, all sixteen drought-treated plants per genotype experienced the same decline in REWsoil. Plants were then watered by the automated system to keep REWsoil at this target level for the following 3 weeks.

After 5 weeks of control or drought treatment, all plants were harvested. The developing secondary xylem was collected as described in Teichmann et al. (2008) from 10 plants per genotype and treatment, frozen immediately in liquid nitrogen and stored at −80 °C until further analysis. For wood anatomical analyses, the bottommost straight upright segment of 2–3 cm length of the stem was collected and immediately fixed in FAE (2% formaldehyde, 5% acetic acid and 63% ethanol). For the analysis of saccharification potential and lignin content, a stem segment of 3–4 cm length above the piece of wood collected for anatomical analysis was debarked and dried to constant weight at 50 °C.

Plant radial growth

The diameter of the stem was measured twice a week on six biological replicates by taking photographs of the stem base with a scale bar attached to a fixed position. Stem diameter was then determined by image analysis using the software ImageJ (Schneider et al. 2012). Based on these measurements, the daily radial area increment \( A_{\text{inc}} \) for the period when target level of drought was reached (Day 13) until final harvest (Day 36) of the experiment was calculated as follows:

\[
A_{\text{inc}} = \frac{\pi r^2_{\text{Day 36}} - \pi r^2_{\text{Day 13}}}{\text{Day 36} - \text{Day 13}}.
\]

Wood anatomical analyses

FAE-fixed stem pieces were transferred into water to remove FAE. Stem cross sections of 10 μm thickness were cut with a freezing microtome (Jung AG, Heidelberg, Germany). The cross-sections were transferred into water and stained in 0.05% (w/v, pH = 7.0) toluidine blue O solution (O’Brien et al. 1964) for 3 min. The sections were then washed and mounted on glass slides using glycerol (28% v/v) as mounting medium. The prepared sections were viewed under a microscope (Axioptan Observer.Z1, Carl Zeiss GmbH, Oberkochen, Germany). Images were taken with a digital camera (Axio Cam MRC, Carl Zeiss Microimaging GmbH, Göttingen, Germany) attached to the microscope.

Images were analyzed with ImageJ (Schneider et al. 2012). Areas of 300 μm × 300 μm or of 100 μm × 200 μm in the mature secondary xylem adjacent to the developing xylem region were used for trait measurements at 100x (vessel frequency, vessel lumen area) or 400x (vessel cell wall thickness, fiber double wall thickness, fiber frequency, fiber lumen area, percentage of cell wall area of vessels and fibers, percentage of ray area, number of cambial cell layers) magnifications, respectively. These areas were selected because all cells here were newly formed during the treatment period. Five biological replicates per genotype and treatment were analyzed.

Five measurements of vessel wall thickness were taken for each sample. To determine fiber double wall thickness, 14 measurements were made for each sample. Vessel or fiber frequency was defined as number of vessels or fibers per unit area. The number...
of vessels or fibers was calculated per mm$^{-2}$. The lumen area of each vessel in the analyzed xylem area (300 μm × 300 μm) of a sample was measured using the ‘Analyze particles’ option of ImageJ. Fiber lumen area was measured manually with the help of ‘freehand selection tool’ in ImageJ. To obtain total fiber lumen area of a specified area, the average fiber lumen area obtained for a subsample of 14 fibers was multiplied with the total fiber number. The fraction of cell wall area (CWA) per total area (%) of analyzed xylem (TAX) of each sample was calculated as follows:

$$\text{CWA}(\%) = \frac{[\text{TAX} - (\text{vessel lumen area} + \text{fibre lumen area} + \text{ray area})]}{\text{TAX}} \times 100.$$

For calculating relative width of developing xylem, first the radius of the developing xylem ring was measured using images taken at 100X magnification while the radius of the whole xylem (developing and mature secondary xylem) was calculated from images taken at 25X magnification as the distance from the cambium to the pith. The relative width of the developing xylem was calculated as:

$$\text{Relative width of developing xylem (\%) = (radial width of developing xylem/ radial width of whole xylem)} \times 100.$$

**Analysis of lignin content**

Cell walls were isolated adapting the procedure by Foster et al. (2010) and Da Costa et al. (2014). For each 70 mg of plant material, cell walls were extracted following this sequence of steps: the material was incubated twice with 1.5 ml of aqueous ethanol for 20 min at 40 °C; once with 1.5 ml chloroform/methanol (1:1 v/v) for 10 min; and three times for 10 min with 1.5 ml acetone. During each incubation, samples were vortexed frequently. Between each step, the supernatant was collected by centrifugation at 15,000 r.p.m., and discarded. Following the last acetone wash, samples were left to dry at 35 °C overnight in a fume hood. Dried biomass was then re-suspended in 1.5 ml of 0.1 M sodium acetate buffer (pH 5.0) and heated to 80 °C for 20 min to induce starch gelatinization. Once cooled, 35 μl of 0.01% sodium azide was added to inhibit microbial growth, and starch was removed by incubation with 35 μl amylase (50 μg/ml in H$_2$O; from Bacillus sp., (Sigma, St Louis, MO, USA)) and 17 μl Pullulanase (18.7 units from Bacillus amylo-lyticus; Sigma). Tubes were incubated overnight at 37 °C and shaking at 400 r.p.m. The next day the digestion was terminated by heating the samples at 98 °C for 10 min. The cell wall material was collected by centrifugation. Purified cell walls were then washed three times with 1.5 ml deionized water and twice with 1.5 ml of acetone followed by drying in the fume hood.

For each sample, acetyl bromide soluble lignin was determined in triplicate following the procedures described in Da Costa et al. (2014) and Foster et al. (2010). Aliquots of 7.0 mg (±0.3) of the cell-wall samples were weighed into 10 ml Pyrex glass tubes. To solubilize lignin, 500 μl of the freshly prepared acetyl bromide solution (25% v/v acetyl bromide in glacial acetic acid) was added. Capped tubes were incubated for 3 h at 50 °C; during the last 30 min of incubation, samples were mixed three times using a vortex mixer. After the digestion tubes were cooled on ice and the samples were diluted by the addition of 2000 μl of 2 M NaOH and 350 μl 0.5 M freshly prepared hydroxyamine hydrochloride and gently mixed. The solution was then neutralized with 7150 μl glacial acetic acid. The tubes were recapped, and mixed several times by inversion and then centrifuged at 1000 r.p.m. for 5 min to pellet undissolved cell wall material. An aliquot of 200 μl of each sample in triplicate was transferred to UV-transparent 96-well plates (UV-Star; Greiner Bio-One, Gloucestershire, UK). The absorbance of each assay mixture was determined two to three times at 280 nm using a plate reader (mQuant; Bio-Tek Instruments, Winooski, VT, USA) and KC4 software (v. 3.3; Bio-Tek). Included in each plate was a standard cell wall preparation of poplar, which was a large mixed sample of cell wall extract that had been assayed 20 times to provide an accurate assay value to test for consistency in subsequent assays; also included was one negative control as the absorbance baseline, which contained all reactants but lacked cell wall material. Acetyl bromide-soluble lignin percentage content (ABSL%) was calculated using absorption reading at 280 nm (A280), path length determined for the 96-well microplates with a volume of 200 μl per well used during the analysis (0.556 cm) (PL), reaction volume in liters (VR) the sample weight in grams (WS), and the specific absorption coefficient (SAC, 17.78 g$^{-1}$ cm$^{-1}$) in the equation described by Da Costa et al. (2014):

$$\text{ABSL} = \left( \frac{A_{280}}{\text{SAC} \times \text{PL}} \right) \times \frac{\text{VR}}{\text{WS}} \times 100\%.$$

**Analysis of saccharification potential**

Samples were assayed for saccharification potential according to the methodology described by Van Acker et al. (2013). Samples were dried overnight at 50 °C and milled for 1 min in a Retsch 300MM Mixer Miller (Retsch GmbH, Haan, Germany) at a frequency of 25 s$^{-1}$. Ground material was sieved and the fraction falling between 150 and 850 μm retained. An aliquot of 10–11 mg of ground material was used for each assay. Samples received acid pretreatment for 2 h at 80 °C in 1 M HCl and washing in 70% ethanol for 20 h at 55 °C. Samples were rinsed with water and acetone prior to re-drying overnight at 50 °C. Samples were reweighed before undergoing 48 h saccharification at 55 °C in a rotating thermomixer in acetic acid buffer with fungal cellulase (Trichoderma reesei) and cellobiase (Aspergillus niger) enzymes (Sigma) with 0.06 filter paper units activity (Xiao et al. 2004). Supernatant post-saccharification was assayed with GOD-POD (glucose oxidase, horseradish peroxidase and ABTS dye) solution permitting spectrophotometric (ELx800 Absorbance Reader, BioTek) glucose quantification from sample absorbance at 405 nm. Saccharification potential was calculated as sample glucose yield as a percentage of post-pretreatment, dry cell wall residue (PPT CWR).
**Statistical analysis of growth and wood anatomical data**

Statistical analyses were performed using R v3.1.1 (R Development Core Team 2015). Two-factorial mixed linear models with ‘genotype’ and ‘treatment’ included as main and interaction effect as well as a random effect for greenhouse chamber were fitted to the data using the function ‘lm’, package ‘nlme’ (Pinheiro et al. 2015). Normality and homogeneity of variance were tested visually by plotting the residuals. The data were transformed logarithmically (log2) if needed to achieve normality and homogeneity of variance of the residuals. In case of a significant ($P < 0.05$) genotype main effect, post-hoc tests were computed using the function ‘lsmeans’ (package ‘lsmeans’; Lenth 2015) called within the function ‘glht’ (package ‘multcomp’; Hothorn et al. 2008).

**RNA extraction, library preparation and sequencing**

Total RNA was extracted from homogenized samples of developing xylem of four biological replicates per genotype and treatment using the CTAB protocol (Chang et al. 1993) with minor modifications described in Janz et al. (2012). Isolated RNA was quality checked using Agilent 2100 Bioanalyzer RNA Nano assay (Agilent Technologies, Santa Clara, CA, USA). Two micrograms of total RNA (RNA Integrity Number RIN > 7.0) was used for library preparation using the ‘TruSeq mRNA Sample Prep kit v2’ (Illumina, San Diego, CA, USA), following the manufacturers’ instructions. Final libraries were quantified using the Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA, USA) and quality tested by Agilent 2100 Bioanalyzer High Sensitivity or the Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA, USA) and manufacturers

**Bioinformatic analyses**

Raw sequence reads were processed with the Python package Cutadapt v1.4.2 (Martin 2011) to remove any residual adapter contamination. Reads were subsequently trimmed to remove low-quality reads (option -trim_qual_left/right = 25) and reads shorter than 40 nucleotides, using the PRINSEQ software v.1.08-0.20.4 (Schmieder and Edwards 2011). Trimmed reads were aligned to the P. trichocarpa v3.0 transcriptome database (Tuskan et al. 2006) using TopHat2 v2.0.10 (Kim et al. 2013). A count table was generated using the Python package HTSeq v0.6.1 (Anders et al. 2015). Further inferential analyses of the count data were done with R package DESeq2, version 1.6.3 (Love et al. 2014) unless otherwise stated. Normalization of count tables was done based on the ‘median ratio method’ (Anders and Huber 2010) implemented in the function ‘estimateSizeFactors’. The analyzed libraries were part of a larger set of libraries generated from various plant tissues. Prior to any down-stream analyses, we identified genes significantly affected by a greenhouse effect across all tissues. By this procedure, three genes were identified and removed from the initial count table. Prior to the analysis of differential expression, we applied an unspecific filtering (Bourgon et al. 2010) to keep only those genes to which at least one read per million reads of library size aligned in at least four samples. The analysis of differential expression was carried out by fitting two-factorial generalized linear models of the negative binomial family (NB) (function ‘DESeq’) to the read counts $K_{ij}$ for gene $i$ in sample $j$ with a logarithmic link (Love et al. 2014):

$$K_{ij} \sim \text{NB}(\text{mean} = \mu_{ij}, \text{dispersion} = \alpha_i)$$

$$\text{log} q_{ij} = \sum x_{ij} \beta_r$$

The computations of the normalization constant $s_j$ and the dispersion parameter $\alpha_i$ are detailed in Love et al. (2014): $q_{ij}$ is proportional to the expectation value of the true concentration of fragments from gene $i$ in sample $j$, $x_{ij}$, denotes the elements of the design matrix $X$, and $\beta_r$ denotes the coefficients for gene $i$ and parameters (corresponding to columns of the design matrix) $r$.

To test for the significance of the ‘genotype-by-treatment’ interaction effect, the full model including ‘genotype’ and ‘treatment’ main and interaction effects (design matrix $X_1$ in Table S1 available as Supplementary Data at *Tree Physiology* Online) was compared against a reduced model without interaction effect (design matrix $X_2$ in Table S1 available as Supplementary Data at *Tree Physiology* Online) by applying a likelihood ratio test implemented in the function ‘nbinomLRT’. To assess the significance of the ‘treatment’ and ‘genotype’ main effect, a full model without interaction effect was set up (design matrix $X_3$ in Table S1 available as Supplementary Data at *Tree Physiology* Online), against which the respective reduced models without ‘treatment’ (design matrix $X_4$ in Table S1 available as Supplementary Data at *Tree Physiology* Online) or ‘genotype’ main effect (design matrix $X_5$ in Table S1 available as Supplementary Data at *Tree Physiology* Online) were tested with the function ‘nbinomLRT’. ‘Treatment’ referred to drought treatment as main effect; genes with significant differences (false discovery rate [FDR] = 0.05) in transcript abundances between drought-treated and control tissue were denominated as drought-related differentially expressed genes (DDEGs). Genes with a significant genotype main effect (FDR = 0.05) in transcript abundance were denominated as genotype-related differentially expressed genes (GDEGs). Pairwise genotype contrasts were computed for the set of GDEGs by specifying the contrasts in the function ‘results’.
Gene co-expression analysis and correlation to wood traits

The GDEGs and DDEGs, which had been obtained by the bioinformatics analyses described above, were used to re-construct transcriptional networks using the R package WGCNA version 1.47 (Langfelder and Horvath 2008, 2012). Gene co-expressions were summarized as adjacency matrices using soft-thresholding powers of 20. The soft-thresholding powers were selected based on low mean connectivity and a >90% model fit to scale-free topology. Adjacency matrices were transformed into topological overlap matrices, which were used to determine modules of co-expressed genes, with a cut height of 0.99 and the minimal module size set to 20 and 30 for the network of DDEGs and GDEGs, respectively. Modules whose correlation of eigengene values was >0.8 were merged. Eigengene values (corresponding to the first principal component of the modules’ gene expression data) of modules of DDEGs or GDEGs were then correlated with wood traits.

Enrichment of gene ontology (GO) categories (Ashburner et al. 2000) within modules was tested using ‘The Ontologist’ (Bauer et al. 2008). Genes assigned to individual modules were used as study sets, and the sets of all genes showing a drought- or genotype-main effect, respectively, as the reference sets (‘gene population’). The transcript abundances of the individual genes in modules related to the saccharification potential were tested for significant correlation with the saccharification potential. The resulting set was mapped onto MapMan categories (MapMan v3.5.1R2; Thimm et al. 2004) based on their Arabidopsis best hits. MapMan’s built-in Wilcoxon rank sum test with Benjamini-Yekutieli correction was used to identify bins enriched in genes with high correlations to saccharification potential. Arabidopsis best hits assigned to the MapMan bin ‘cell wall’ (bin 10) were used as the co-expression database CressExpress (Srinivasasaianganendra et al. 2008; http://cressexpress.org/) to identify genes commonly co-regulated with these bait genes across multiple experiments. CressExpress v3.0 was used with default quality-control statistic D to search for co-regulated genes in the microarray data from experiments NASC_137 (‘AtGenExpress Stress Treatments, (control plants); GSE5620; Kilian et al. 2007), NASC_141 (‘AtGenExpress Stress Treatments’ (drought stress); GSE5624; Kilian et al. 2007), NASC_153 (‘AtGenExpress Developmental Series shoots and stems’; GSE5633), NASC_14 (‘control of lignification’; Rogers et al. 2005) and NASC_27 (‘Assembly of the cell wall pectic matrix’; GSE6181). Genes were considered co-expressed with the saccharification-related bait genes when the gene-expression data across the selected microarray experiments were correlated with \( r^2 > 0.8 \) and \( P < 0.05 \). Co-expressed genes annotated as transcription factors were identified according to the list of A. thaliana transcription factors available at the Plant Transcription Factor Database v4.0 (Jin et al. 2016; http://planttfdb.cbi.pku.edu.cn/). The identified transcription factors were compared with the transcription factors in the list of DEGs and genes present in both lists were extracted as candidate genes.

Quantitative real time polymerase chain reaction (qRT-PCR) of selected genes

RNAseq data were validated by quantitative real time polymerase chain reaction (qRT-PCR) on a Light Cycler 480 system (Roche Diagnostics, Mannheim, Germany). Primers were designed with Oligo Explorer (Gene Link, Hawthorne, NY, USA, http://www.genelink.com/). We used three genes without significant changes across our conditions as the references (Potri.012G141400, Potri.001G006400, Potri.005G074900; primer sequences (5’−3’, forward and reverse): AAATTGAGTTGGGAAAAGGC, ACAACATCTGGCATGCAATCC for Potri.012G141400; GATCCCTATGATGCTGCTGA, GCAACTTGACGGTCCTTGTG for Potri.001G006400; ACCA CGAGACAAAGTGTCTT, CTTTTGGGCTTCTTGCAAAC for Potri.005G074900), and four genes as targets (Potri.014G035500, Potri.012G097900, Potri.010G083600, Potri.004G038700; primer sequences (5’−3’, forward and reverse): TTCGAGCCCCAACATTACAAAG, GATAGTAGTTGGGATGGCGAT for Potri.014G035500; AATCTCCCTATGGCATCAG, ATGATGATCAGACGCCGCTGG for Potri.012G097900; GGTGTGAGGAAATTAGGACGC, AGCTTGGCCAGTTCAATGTGGAATC for Potri.010G083600; CCGTTC ACTCTCCTCCATATGG for Potri.004G038700). Primer efficiencies and Cq-values were obtained using LinRegPCR v2016.2.0.0 (Ruijter et al. 2009). The expression of target genes was related to the expression of the reference genes as the inverse of the normalized relative quantity defined in Hellemans et al. (2007). The correlation of the relative expression of target genes determined by qRT-PCR with normalized transcript abundance obtained by RNAseq was analyzed. The following correlations were obtained: Potri.014G035500: \( r = 0.63, P = 0.001 \), Potri.012G097900: \( r = 0.73, P < 0.001 \), Potri.010G083600: \( r = 0.79, P < 0.001 \), Potri.004G038700: \( r = 0.38, P = 0.071 \) (Figure S1 available as Supplementary Data at Tree Physiology Online).

Results

Genotypic and drought-induced variation in wood anatomical traits

The three P. nigra genotypes originating from habitats with increasing precipitation (Spain < France < Italy; DeWoody et al. 2015), differed significantly in radial area increment \(< P < 0.0001 \). Genotype France showed a significantly higher radial area increment as compared with the genotypes Italy and Spain (Table 1). Drought resulted in a significant decline in radial growth \(< P < 0.0001 \), while the genotype–drought interaction effect was not significant \( P = 0.662 \).

In order to elucidate whether the variation in radial area production was related to changes in cambial activity or the
differentiation of newly formed xylem cells, we counted the cambial cell layers and determined the radial width of the developing xylem. The number of cambial cell layers did not vary among genotypes (Table 1). Significant genotypic variation was observed in the relative radial width of the developing xylem ($P = 0.014$), with genotype Italy having a smaller relative width of the developing xylem compared with genotype France (Table 1). The genotypes differed significantly in vessel ($P = 0.007$) and fiber frequencies ($P = 0.02$) as well as in the vessel ($P = 0.007$) and fiber lumen areas ($P < 0.0001$, Table 1). The genotype France showed lower vessel and fiber frequencies and significantly higher vessel and fiber lumen areas compared with the other two genotypes (Table 1). Neither vessel nor fiber cell wall thickness or cell wall area of vessels and fibers differed among the *P. nigra* genotypes studied here (Table 1). Genotype Spain showed a significantly higher percentage of ray area as compared with genotype France ($P = 0.034$). Under drought, both the number of cambial cell layers and the relative radial width of the developing xylem decreased (Table 1), reflecting the significant reduction in radial area growth under drought ($P < 0.0001$). The vessel frequency increased, while the decrease in vessel lumen area under drought was marginally significant ($P = 0.062$, Table 1). Among the fiber traits, only the fiber lumen area was reduced under drought ($P = 0.023$), whereas the fiber frequency was unaffected in the three genotypes (Table 1). The fraction of cell wall area of vessels, fibers and rays

---

**Table 1.** Wood anatomical traits and lignin of three genotypes of *Populus nigra* exposed to a control or drought treatment for 5 weeks.

<table>
<thead>
<tr>
<th>Anatomical trait</th>
<th>Genotype</th>
<th>Mean ± SE</th>
<th>Control</th>
<th>Drought</th>
<th>$P$-value</th>
<th>Post-hoc</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radial area growth (mm$^2$ day$^{-1}$)</td>
<td>Spain</td>
<td>1.9 ± 0.12</td>
<td>0.9 ± 0.12</td>
<td>$&lt;0.0001$</td>
<td>$&lt;0.0001$</td>
<td>0.663</td>
</tr>
<tr>
<td></td>
<td>France</td>
<td>2.9 ± 0.04</td>
<td>1.9 ± 0.13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Italy</td>
<td>2.1 ± 0.12</td>
<td>1.3 ± 0.09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative radial width of developing xylem (%)</td>
<td>Spain</td>
<td>5.3 ± 0.63</td>
<td>4.0 ± 0.40</td>
<td>0.005</td>
<td>0.014</td>
<td>0.294</td>
</tr>
<tr>
<td></td>
<td>France</td>
<td>5.3 ± 0.54</td>
<td>4.8 ± 0.57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Italy</td>
<td>4.5 ± 0.27</td>
<td>2.9 ± 0.24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of cambial cell layers</td>
<td>Spain</td>
<td>6.4 ± 0.2</td>
<td>4.8 ± 0.4</td>
<td>$&lt;0.0001$</td>
<td>0.659</td>
<td>0.241</td>
</tr>
<tr>
<td></td>
<td>France</td>
<td>7.0 ± 0.3</td>
<td>4.4 ± 0.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Italy</td>
<td>7.2 ± 0.4</td>
<td>4.6 ± 0.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vessel frequency (vessel number mm$^{-2}$)</td>
<td>Spain</td>
<td>170 ± 12</td>
<td>199 ± 13</td>
<td>0.007</td>
<td>0.007</td>
<td>0.704</td>
</tr>
<tr>
<td></td>
<td>France</td>
<td>124 ± 12</td>
<td>163 ± 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Italy</td>
<td>176 ± 16</td>
<td>264 ± 59</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vessel lumen area per vessel (μm$^2$)</td>
<td>Spain</td>
<td>1285 ± 61</td>
<td>1075 ± 68</td>
<td>0.062</td>
<td>0.007</td>
<td>0.720</td>
</tr>
<tr>
<td></td>
<td>France</td>
<td>1643 ± 137</td>
<td>1371 ± 125</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Italy</td>
<td>1173 ± 42</td>
<td>1048 ± 213</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vessel wall thickness (μm)</td>
<td>Spain</td>
<td>1.10 ± 0.06</td>
<td>1.15 ± 0.08</td>
<td>0.222</td>
<td>0.139</td>
<td>0.841</td>
</tr>
<tr>
<td></td>
<td>France</td>
<td>1.26 ± 0.09</td>
<td>1.30 ± 0.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Italy</td>
<td>1.16 ± 0.07</td>
<td>1.28 ± 0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fiber frequency (fiber number mm$^{-2}$)</td>
<td>Spain</td>
<td>3894 ± 210</td>
<td>4383 ± 290</td>
<td>0.595</td>
<td>0.02</td>
<td>0.230</td>
</tr>
<tr>
<td></td>
<td>France</td>
<td>3925 ± 248</td>
<td>3607 ± 314</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Italy</td>
<td>5003 ± 431</td>
<td>4411 ± 293</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fiber lumen area per fiber (μm$^2$)</td>
<td>Spain</td>
<td>98 ± 11</td>
<td>82 ± 3.7</td>
<td>$0.023$</td>
<td>$&lt;0.0001$</td>
<td>0.235</td>
</tr>
<tr>
<td></td>
<td>France</td>
<td>125 ± 9</td>
<td>124 ± 5.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Italy</td>
<td>97 ± 7</td>
<td>71 ± 2.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fiber double wall thickness (μm)</td>
<td>Spain</td>
<td>3.8 ± 0.16</td>
<td>3.9 ± 0.19</td>
<td>0.495</td>
<td>0.464</td>
<td>0.998</td>
</tr>
<tr>
<td></td>
<td>France</td>
<td>3.6 ± 0.23</td>
<td>3.7 ± 0.27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Italy</td>
<td>3.5 ± 0.19</td>
<td>3.7 ± 0.24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell wall area of vessels and fibers (%)</td>
<td>Spain</td>
<td>33.8 ± 2.8</td>
<td>36.8 ± 2.2</td>
<td>$0.023$</td>
<td>0.13</td>
<td>0.828</td>
</tr>
<tr>
<td></td>
<td>France</td>
<td>27.8 ± 0.1</td>
<td>33.1 ± 3.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Italy</td>
<td>31.0 ± 2.6</td>
<td>36.7 ± 1.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage of ray area (%)</td>
<td>Spain</td>
<td>9.9 ± 1.7</td>
<td>9.5 ± 1.0</td>
<td>0.803</td>
<td>0.025</td>
<td>0.988</td>
</tr>
<tr>
<td></td>
<td>France</td>
<td>6.8 ± 1.1</td>
<td>6.5 ± 1.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Italy</td>
<td>7.0 ± 1.2</td>
<td>6.9 ± 0.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lignin content per dry cell wall (%)</td>
<td>Spain</td>
<td>17.2 ± 0.56</td>
<td>16.9 ± 0.40</td>
<td>0.269</td>
<td>0.003</td>
<td>0.348</td>
</tr>
<tr>
<td></td>
<td>France</td>
<td>15.6 ± 0.49</td>
<td>16.1 ± 0.51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Italy</td>
<td>17.1 ± 0.57</td>
<td>18.2 ± 0.40</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Mean $\pm$ SE are given for $n = 5–6$ biological replicates per genotype and treatment. $P$-values as computed for 2-factorial linear models including genotype and drought main effects and the genotype–drought interaction effect (G × D). Significant effects with $P < 0.05$ are indicated in bold letters. Post hoc: letters indicate homogenous subgroups of genotypes identified by post-hoc tests for traits showing a significant genotype effect.
was significantly larger in drought-treated than in control trees \((P = 0.023)\). There was no significant genotype–drought interaction effect in any of the analyzed wood anatomical traits.

**Saccharification potential and lignin in *P. nigra* genotypes in relation to wood anatomical traits**

The lignin content (per cell wall mass) showed significant genotypic variation, with lower values in genotype France than in genotype Italy \((P = 0.003)\), but was unaffected by drought (Table 1). Saccharification potential was significantly higher in genotype Spain compared with genotypes France and Italy (Figure 1), and increased in response to drought \((P = 0.021, \text{Figure 1})\).

We tested whether relationships existed among the chemical and wood anatomical traits (Figure 2). The anticipated negative relationship between lignin and saccharification potential was not observed. The saccharification potential showed only a weak relationship with fiber lumen \((P = 0.047, \text{Figure 2})\). In contrast to the saccharification potential, wood traits and lignin were connected by various correlations. For example, the lignin content was negatively correlated with vessel \((P = 0.026)\) and fiber lumina \((P = 0.005)\), suggesting links between cell expansion and lignification. The strongest positive correlations were found between double wall thickness of fibers and the fraction of cell wall area \((P < 0.001)\) and between fiber lumina area and relative width of developing xylem \((P < 0.001, \text{Figure 2})\) indicating that fiber expansion may be a stronger driver of growth than vessel expansion \((r = 0.63\) for fiber lumina area compared with \(r = 0.49\) for vessel lumen area) and that thick fiber cell walls lead to a high fraction of cell wall area, i.e., dense wood (Figure 2). The strongest negative relationships were found between vessel frequency and vessel lumen \((P < 0.001)\) and between fiber lumen and fraction of cell wall area \((P < 0.001, \text{Figure 2})\), indicating that decreases in vessel lumen are strongly connected with increases in vessel frequencies, whereas decreases in fiber lumen apparently increased the fraction of cell wall area.

**The transcriptome of the developing xylem varied substantially among genotypes and was moderately affected by drought**

In order to identify genes and clusters of co-expressed genes underlying genotype or drought-induced variation in saccharification potential and wood traits, whole-transcriptome gene expression profiling of the developing xylem was done by RNA-sequencing. After trimming to remove adaptor contamination and low-quality reads we obtained 23.1 ± 0.95 M reads (mean ± SE) per sample. After alignment to the *P. trichocarpa* reference and removal of genes with low expression levels, a count table was obtained for 23,393 unique gene models.

Principal component analysis of the count data revealed a prominent separation of genotypes along both of the first two principal components, which together explained 60% of the variance, while clustering according to treatment was less apparent (Figure S2 available as Supplementary Data at Tree Physiology Online). Testing for differential expression confirmed strong genotype-related effects because a total of 15,657 GDEGs (67% of the unique gene models tested) were found. Pairwise genotype contrasts between the genotypes showed that ~20% of the GDEGs differed among all genotypes (Figure S3 available as Supplementary Data at Tree Physiology Online); the remaining GDEGs showed similar distributions between pairwise comparisons, suggesting similar divergence of the molecular regulation among the three genotypes (Figure S3 available as Supplementary Data at Tree Physiology Online).

A drought main effect was detected for only 347 DDEGs, of which 219 were down- and 128 were up-regulated in response to low water availability. A significant genotype–drought interaction effect was not evident for any of the tested genes.

**Gene co-expression modules showed distinct correlations with wood anatomical and chemical traits**

In order to relate the sets of GDEGs and DDEGs to genotype-related or drought-induced variation in wood traits, we first reconstructed transcriptional networks of GDEGs and DDEGs by WGCNA and then correlated modules of co-expressed genes with wood traits (Figure 3).

The GDEGs clustered in 18 co-expression modules (DX-G1 to DX-G18; Figure 3A). Wood traits that showed no genotype-related variation (number of cambial cell layers, thickness of fiber and vessel walls, and the fraction of cell wall area; Table 1) showed no or weak correlations with the co-expression modules of GDEGs (Figure 3A). Furthermore, we identified three modules DX-G1, DX-G12 and DX-G7 without any significant correlation with wood traits (Figure 3A). The wood anatomical traits

---

**Figure 1.** Saccharification potential of well-watered and drought-treated *Populus nigra* genotypes originating from contrasting habitats in Spain, Italy and France. Saccharification potential was expressed as glucose yield per dry cell wall residue \((n = 5\) biological replicates). In the boxes, circles denote the mean and horizontal lines the median. \(P\) values for the drought and genotype main effect, and for the genotype–drought interaction effect \((P_{G-D})\) are given.
showed significant positive or negative correlations with DX-G2, DX-G3, DX-G6, DX-G8, DX-G9, DX-G11 and DX-G18 with frequent overlaps (Figure 3A). For example, DX-G3 showed positive correlations with fiber and vessel frequencies and negative correlations with fiber and vessel lumina and the relative width of the developing xylem.

The modules of the GDEGs that were correlated with the saccharification potential (positive: DX-G13, DX-G14, DX-G15, DX-G16, DX-G17; negative: DX-G4, DX-G5, DX-G10) showed in most cases no overlap with wood traits (Figure 3A). Exceptions were DX-G13, which was inversely correlated with vessel lumen and DX-G10, which was inversely correlated with fiber lumen.

To obtain information on the functions that were correlated with the chemical or anatomical traits, the modules were subjected to GO term analyses. DX-G13 and DX-G10, the modules inversely correlated with vessel and fiber lumina, showed significant enrichment of GO terms related to ‘catechol-containing compound biosynthesis’ and ‘cinnamic acid metabolism’ (DX-G13), and in ‘small molecule catabolic process’, ‘coumarin biosynthetic process’, ‘flavonoid biosynthetic process’ and ‘cell wall biogenesis’ (DX-G10) (Table S2 available as Supplementary Data at Tree Physiology Online).

The modules positively correlated with saccharification potential showed significant enrichments in GO terms for ‘flavonoid biosynthesis’ and ‘anthocyanin-containing compound biosynthesis’ (DX-G14), ‘cellular protein modification processes’, ‘ethylene-related signaling’ (DX-G16) and ‘regulation of cell morphogenesis and growth’ (DX-G17; Table S2 available as Supplementary Data at Tree Physiology Online). Modules DX-G10, DX-G4 and DX-G5 were significantly negatively correlated with saccharification potential (see Table S3 available as Supplementary Data at Tree Physiology Online for GDEGs assigned to these modules) and were significantly enriched in GO terms for ‘carbohydrate metabolic process’ (DX-G4, DX-G5), ‘cellular metabolic compound salvage’ and ‘thioester metabolic process’ (DX-G5).

The drought-induced DEGs clustered in three co-expressions modules (DX-D1, DX-D2, DX-D3; Figure 3B). Eigengenes of these modules were correlated with those wood anatomical traits that showed significant differences between control and drought treated plants (Figure 3B, Table 1), and with saccharification potential (Figure 3B, Table 1). Module DX-D1 was significantly positively correlated with the number of cambial cell layers \((P = 0.017)\) and significantly enriched in GO terms related to ‘translation’ and ‘cellular macromolecule biosynthetic process’ (Table S2 available as Supplementary Data at Tree Physiology Online). Module DX-D2 was also positively correlated with the number of cambial cell layers \((P = 0.019)\) but in addition
significantly negatively correlated with saccharification potential (\( P = 0.017 \), Figure 3B). Module DX-D3 was significantly negatively correlated with the number of cambial cell layers (\( P = 0.002 \)), fiber lumen area (\( P = 0.032 \)) and vessel lumen area (\( P = 0.03 \)), but strongly positively correlated with saccharification potential (\( P = 0.007 \), Figure 3B). The GO terms associated with this module pointed to membrane-related processes ('membrane-bounded organelle' and 'intracellular membrane-bounded organelle').

To identify putative key genes involved in the control of saccharification potential, we ranked all GDEGs and DDEGs assigned to modules significantly correlated with saccharification potential (Figure 3) by the significance of correlation of transcript abundance with this wood property. The most strongly correlated GDEGs were protein kinase proteins, MAPKKK proteins, cytochrome P450 proteins, a heteroglucan glucosidase and others (Table S3 available as Supplementary Data at Tree Physiology Online, Figure 4A). The DDEGs most strongly correlated with saccharification potential included nucleotide-sugar transporter family proteins, a phosphomanomutase, a methyltransferase and others (Table S4 available as Supplementary Data at Tree Physiology Online, Figure 4B).

**Metabolic processes in the developing xylem underlying variation in saccharification potential**

Since variation in saccharification potential was largely independent from wood anatomical traits (Figure 2), we sought to get additional insights into the metabolic control of saccharification potential by mapping all GDEGs that were assigned to any of the saccharification-correlated modules and whose transcript abundance was significantly correlated with the saccharification potential onto MapMan functional categories ('bins').

An over-representation of GDEGs with strong positive or negative correlations to saccharification potential was detected in five bins, namely 'cell/cell vesicle transport', 'cell wall', 'misc/cytochrome P450' and 'protein targeting secretory pathway' (Table 2). A closer inspection of the bin 'cell wall' revealed that transcript abundances of GDEGs involved in 'cell wall modification' (bin 10.7), 'hemicellulose synthesis' (bin 10.3), 'pectin synthesis' (bin 10.4), 'cellulose synthesis' (bin 10.2) and 'cell wall degradation' (bins 10.6.2 and 10.6.3) were mostly negatively correlated with saccharification potential (Figure 5A, Table S5 available as Supplementary Data at Tree Physiology Online). In contrast, bins 'pectin esterases' (bin 10.8) and 'cell wall/precursor synthesis' (bin 10.1) included specific processes that were positively correlated with saccharification

![Figure 3. Heatmap representing the correlation of eigengenes of co-expression modules with wood traits for genes showing significant genotype-related variation in transcript abundance (DX-G1 to DX-G18) (A) or drought-responsive differentially expressed genes (DX-D1 to DX-D3) (B). Eigengenes of the modules correspond to the first principal component. Numbers represent Pearson’s correlation coefficient \( r \) and the \( P \) value for the correlation (in brackets). For abbreviations of traits see Figure 2.](image-url)

<table>
<thead>
<tr>
<th>Module</th>
<th>r</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>DX-G1</td>
<td>-0.07</td>
<td>0.20</td>
</tr>
<tr>
<td>DX-G2</td>
<td>0.20</td>
<td>0.18</td>
</tr>
<tr>
<td>DX-G3</td>
<td>-0.60</td>
<td>0.05</td>
</tr>
<tr>
<td>DX-G4</td>
<td>-0.11</td>
<td>0.05</td>
</tr>
<tr>
<td>DX-G5</td>
<td>-0.35</td>
<td>0.01</td>
</tr>
<tr>
<td>DX-G6</td>
<td>-0.02</td>
<td>0.18</td>
</tr>
<tr>
<td>DX-G7</td>
<td>-0.05</td>
<td>0.18</td>
</tr>
<tr>
<td>DX-G8</td>
<td>-0.25</td>
<td>0.01</td>
</tr>
<tr>
<td>DX-G9</td>
<td>-0.24</td>
<td>0.01</td>
</tr>
<tr>
<td>DX-G10</td>
<td>-0.03</td>
<td>0.12</td>
</tr>
<tr>
<td>DX-G11</td>
<td>-0.01</td>
<td>0.12</td>
</tr>
<tr>
<td>DX-G12</td>
<td>-0.04</td>
<td>0.12</td>
</tr>
<tr>
<td>DX-G13</td>
<td>-0.05</td>
<td>0.12</td>
</tr>
<tr>
<td>DX-G14</td>
<td>-0.06</td>
<td>0.12</td>
</tr>
<tr>
<td>DX-G15</td>
<td>-0.07</td>
<td>0.12</td>
</tr>
<tr>
<td>DX-G16</td>
<td>-0.08</td>
<td>0.12</td>
</tr>
<tr>
<td>DX-G17</td>
<td>-0.09</td>
<td>0.12</td>
</tr>
<tr>
<td>DX-G18</td>
<td>-0.10</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Metabolic processes in the developing xylem underlying variation in saccharification potential

Since variation in saccharification potential was largely independent from wood anatomical traits (Figure 2), we sought to get additional insights into the metabolic control of saccharification potential by mapping all GDEGs that were assigned to any of the saccharification-correlated modules and whose transcript abundance was significantly correlated with the saccharification potential onto MapMan functional categories ('bins').

An over-representation of GDEGs with strong positive or negative correlations to saccharification potential was detected in five bins, namely 'cell/cell vesicle transport', 'cell wall', 'misc/cytochrome P450' and 'protein targeting secretory pathway' (Table 2). A closer inspection of the bin 'cell wall' revealed that transcript abundances of GDEGs involved in 'cell wall modification' (bin 10.7), 'hemicellulose synthesis' (bin 10.3), 'pectin synthesis' (bin 10.4), 'cellulose synthesis' (bin 10.2) and 'cell wall degradation' (bins 10.6.2 and 10.6.3) were mostly negatively correlated with saccharification potential (Figure 5A, Table S5 available as Supplementary Data at Tree Physiology Online). In contrast, bins 'pectin esterases' (bin 10.8) and 'cell wall/precursor synthesis' (bin 10.1) included specific processes that were positively correlated with saccharification
potential. The bin 'cell wall/precursor synthesis' (Figure 5B) contained transcripts encoding UDP-\(\beta\)-glucose/UDP-\(\beta\)-galactose epimerases (bin 10.1.2), UDP-glucuronic acid decarboxylases (bin 10.1.5), and enzymes involved in GDP-\(\alpha\)-mannose (bins 10.1.20, 10.1.21, 10.1.1) and UDP-rhamnose biosynthesis (bin 10.1.11) that were negatively correlated with saccharification potential. In contrast, transcripts coding for enzymes involved in the biosynthesis of UDP-\(\alpha\)-arabinose (bin 10.1.9, UDP-arabinose 4-epimerase), UDP-\(\beta\)-galacturonic acid (bin 10.1.6, UDP-\(\beta\)-glucuronate 4-epimerase 2) and UDP-\(\alpha\)-rhamnose (bin 10.1.10,
In order to identify regulatory genes commonly co-expressed with saccharification potential (Figure 5B; Table S5 available as Supplementary Data at Tree Physiology Online).

In contrast to our initial hypothesis that drought will cause an increase in lignification (Moore et al. 2008, Le Gall et al. 2015), the lignin content was not affected under drought. We had further hypothesized that along with an increased lignification, drought will cause a reduction in saccharification potential. Our data do not support these hypotheses, since we observed a significant increase in saccharification potential under drought across all genotypes (Figure 1) and no correlation with lignin or with wood anatomy. We are not aware of any previous studies on drought effects on saccharification potential in woody dicots, but van der Weijde et al. (2016) reported an increase in saccharification potential of Miscanthus under drought. These observations suggest that drought may cause alterations in cell wall chemistry apart from lignin that facilitate the enzymatic release of glucose.

Important bioenergy crops such as Miscanthus (Souza et al. 2015), Eucalyptus (Klash et al. 2010, Elissetche et al. 2011), and Populus (Wegrzyn et al. 2010, Studer et al. 2011, Zhou et al. 2011, Guerra et al. 2013, Porth et al. 2013, Muchero et al. 2015, Fahrenkrog et al. 2016) exhibit substantial natural genetic variation in lignin content. Likewise, saccharification potential of bioenergy crops showed genetic variation in earlier studies (Studer et al. 2011, Souza et al. 2015, van der Weijde et al. 2016). In agreement with these findings, P. nigra also showed genotypic variations in lignin content and in saccharification potential (Table 1, Figure 1). Our hypothesis that adaptation of poplars to low water availability in their native habitat correlates with high lignin content and low saccharification potential has to be refuted: saccharification potential was unexpectedly highest in genotype Spain, originating from the driest habitat, whereas the lignin content of the Spain genotype was not different from that of the Italy genotype, which originated

<table>
<thead>
<tr>
<th>Bin ID</th>
<th>Bin name</th>
<th>Number of genes mapped</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>Cell</td>
<td>194</td>
<td>2.86E-08</td>
</tr>
<tr>
<td>31.4</td>
<td>Cell/vesicle transport</td>
<td>52</td>
<td>8.60E-07</td>
</tr>
<tr>
<td>10</td>
<td>cell wall</td>
<td>89</td>
<td>2.50E-03</td>
</tr>
<tr>
<td>26.10</td>
<td>Misc/cytochrome P450</td>
<td>25</td>
<td>2.83E-02</td>
</tr>
<tr>
<td>29.3.4</td>
<td>Protein targeting/secretory pathway</td>
<td>52</td>
<td>3.48E-02</td>
</tr>
</tbody>
</table>

P values as calculated by MapMan’s in-built Wilcoxon rank sum test and adjusted for multiple testing by Benjamini-Yekutieli correction.

Discussion

Drought-induced variations in wood anatomical traits and lignin content are not related to variation in saccharification potential

In contrast to our initial hypothesis that drought will cause an increase in lignification (Moore et al. 2008, Le Gall et al. 2015), the lignin content was not affected under drought. We had further hypothesized that along with an increased lignification, drought will cause a reduction in saccharification potential. Our data do not support these hypotheses, since we observed a significant increase in saccharification potential under drought across all genotypes (Figure 1) and no correlation with lignin or with wood anatomy. We are not aware of any previous studies on drought effects on saccharification potential in woody dicots, but van der Weijde et al. (2016) reported an increase in saccharification potential of Miscanthus under drought. These observations suggest that drought may cause alterations in cell wall chemistry apart from lignin that facilitate the enzymatic release of glucose.

Important bioenergy crops such as Miscanthus (Souza et al. 2015), Eucalyptus (Klash et al. 2010, Elissetche et al. 2011), and Populus (Wegrzyn et al. 2010, Studer et al. 2011, Zhou et al. 2011, Guerra et al. 2013, Porth et al. 2013, Muchero et al. 2015, Fahrenkrog et al. 2016) exhibit substantial natural genetic variation in lignin content. Likewise, saccharification potential of bioenergy crops showed genetic variation in earlier studies (Studer et al. 2011, Souza et al. 2015, van der Weijde et al. 2016). In agreement with these findings, P. nigra also showed genotypic variations in lignin content and in saccharification potential (Table 1, Figure 1). Our hypothesis that adaptation of poplars to low water availability in their native habitat correlates with high lignin content and low saccharification potential has to be refuted: saccharification potential was unexpectedly highest in genotype Spain, originating from the driest habitat, whereas the lignin content of the Spain genotype was not different from that of the Italy genotype, which originated...
Figure 5. Mapping of saccharification-related genes to Map Man functional categories (bins) related to cell wall metabolism (A, overview; B, details of bin 10.1, 'precursor synthesis'). Numbers in (A) refer to cell wall-related bins as follows: 10.1, 'precursor synthesis'; 10.2, 'cellulose synthesis'; 10.3, 'hemicellulose synthesis'; 10.4, 'pectin synthesis'; 10.5, 'cell wall proteins'; 10.6, 'degradation'; 10.7, 'modification', 10.8, 'pectin esterases' and in (B) to the bin 'cell wall/precursor synthesis' with sub-bins explained in Table S5 (available as Supplementary Data at Tree Physiology Online). Color code represents Pearson’s correlation coefficient of transcript abundance with saccharification potential.
Table 3. Saccharification-related bait genes present in the functional category ‘cell wall’ and prey transcription factors.

<table>
<thead>
<tr>
<th>Ath locus</th>
<th>Ath hit description</th>
<th>Ptr locus</th>
<th>Locus</th>
<th>Description</th>
<th>$r^2_{\text{coexp}}$</th>
<th>$P(r^2_{\text{coexp}})$</th>
<th>Ptr locus</th>
<th>$r_{\text{sacc}}$</th>
<th>$P(r_{\text{sacc}})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT2G42800</td>
<td>Receptor like protein 29</td>
<td>Potri.008G211800</td>
<td>−0.48</td>
<td>1.63E−02</td>
<td>AT1G23380</td>
<td>KNAT6</td>
<td>0.87</td>
<td>9.88E−52</td>
<td>Potri.008G188700</td>
</tr>
<tr>
<td>AT2G23130</td>
<td>Arabinogalactan protein 17</td>
<td>Potri.007G051600</td>
<td>0.56</td>
<td>4.63E−03</td>
<td>AT1G29160</td>
<td>Dof-type zinc finger domain-containing protein</td>
<td>0.81</td>
<td>1.54E−43</td>
<td>Potri.011G065900</td>
</tr>
<tr>
<td>AT2G42800</td>
<td>Receptor like protein 29</td>
<td>Potri.008G211800</td>
<td>−0.48</td>
<td>1.63E−02</td>
<td>AT1G53160</td>
<td>SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 4</td>
<td>0.86</td>
<td>3.35E−50</td>
<td>Potri.007G138800</td>
</tr>
<tr>
<td>AT2G42800</td>
<td>Receptor like protein 29</td>
<td>Potri.008G211800</td>
<td>−0.48</td>
<td>1.63E−02</td>
<td>AT1G75430</td>
<td>BEL1-LIKE HOMEODOMAIN 11 (BLH11)</td>
<td>0.83</td>
<td>9.01E−46</td>
<td>Potri.002G030900</td>
</tr>
<tr>
<td>AT2G36870</td>
<td>Xyloglucan endotransglucosylase/hydrolase 32</td>
<td>Potri.009G006600</td>
<td>0.51</td>
<td>1.08E−02</td>
<td>AT1G77920</td>
<td>bZIP family transcription factor</td>
<td>0.85</td>
<td>7.85E−50</td>
<td>Potri.002G097000</td>
</tr>
<tr>
<td>AT2G40610</td>
<td>Expansin A8</td>
<td>Potri.013G154700</td>
<td>0.42</td>
<td>4.00E−02</td>
<td>AT2G18300</td>
<td>Basic helix-loop-helix (bHLH) family protein</td>
<td>0.83</td>
<td>7.23E−46</td>
<td>Potri.007G023600</td>
</tr>
<tr>
<td>AT1G02810</td>
<td>Plant invertase/pectin methylesterase inhibitor superfamily</td>
<td>Potri.014G127000</td>
<td>−0.54</td>
<td>6.82E−03</td>
<td>AT2G34830</td>
<td>WRKY DNA-binding protein 35 (WRKY35)</td>
<td>0.86</td>
<td>5.49E−51</td>
<td>Potri.002G195300</td>
</tr>
<tr>
<td>AT2G42800</td>
<td>Receptor like protein 29</td>
<td>Potri.008G211800</td>
<td>−0.48</td>
<td>1.63E−02</td>
<td>AT2G42280</td>
<td>bHLH family protein</td>
<td>0.83</td>
<td>1.86E−45</td>
<td>Potri.006G057600</td>
</tr>
<tr>
<td>AT2G47930</td>
<td>Arabinogalactan protein 26</td>
<td>Potri.002G207500</td>
<td>−0.57</td>
<td>3.45E−03</td>
<td>AT2G47260</td>
<td>WRKY23</td>
<td>0.81</td>
<td>7.34E−43</td>
<td>Potri.014G18200</td>
</tr>
<tr>
<td>AT2G42800</td>
<td>Receptor like protein 29</td>
<td>Potri.008G211800</td>
<td>−0.48</td>
<td>1.63E−02</td>
<td>AT3G13960</td>
<td>GROWTH-REGULATING FACTOR 5 (AtGRF5)</td>
<td>0.82</td>
<td>3.80E−45</td>
<td>Potri.019G042300</td>
</tr>
<tr>
<td>AT2G42800</td>
<td>Receptor like protein 29</td>
<td>Potri.008G211800</td>
<td>−0.48</td>
<td>1.63E−02</td>
<td>AT3G15270</td>
<td>SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 5 (SPL5)</td>
<td>0.88</td>
<td>2.99E−54</td>
<td>Potri.001G398200</td>
</tr>
<tr>
<td>AT2G42800</td>
<td>Receptor like protein 29</td>
<td>Potri.008G211800</td>
<td>−0.48</td>
<td>1.63E−02</td>
<td>AT3G19184</td>
<td>DNA binding</td>
<td>0.88</td>
<td>1.24E−53</td>
<td>Potri.005G137600</td>
</tr>
<tr>
<td>AT2G36870</td>
<td>Xyloglucan endotransglucosylase/hydrolase 32</td>
<td>Potri.009G006600</td>
<td>0.51</td>
<td>1.08E−02</td>
<td>AT3G28920</td>
<td>ARABIDOPSIS THALIANA HOMEBOX PROTEIN 34</td>
<td>0.81</td>
<td>4.86E−43</td>
<td>Potri.015G032700</td>
</tr>
<tr>
<td>AT1G02810</td>
<td>Plant invertase/pectin methylesterase inhibitor superfamily</td>
<td>Potri.014G127000</td>
<td>−0.54</td>
<td>6.82E−03</td>
<td>AT3G49760</td>
<td>Arabidopsis thaliana basic leucine-zipper 5 (AtbZIP5)</td>
<td>0.85</td>
<td>1.51E−48</td>
<td>Potri.007G006900</td>
</tr>
<tr>
<td>AT2G42800</td>
<td>Receptor like protein 29</td>
<td>Potri.008G211800</td>
<td>−0.48</td>
<td>1.63E−02</td>
<td>AT4G13640</td>
<td>Unfertilized embryo sac 16 (UNE16)</td>
<td>0.80</td>
<td>2.56E−42</td>
<td>Potri.001G314800</td>
</tr>
<tr>
<td>AT2G42800</td>
<td>Receptor like protein 29</td>
<td>Potri.008G211800</td>
<td>−0.48</td>
<td>1.63E−02</td>
<td>AT4G37740</td>
<td>GROWTH REGULATING FACTOR 2 (AtGRF2)</td>
<td>0.81</td>
<td>3.85E−43</td>
<td>Potri.003G108800</td>
</tr>
<tr>
<td>AT2G42800</td>
<td>Receptor like protein 29</td>
<td>Potri.008G211800</td>
<td>−0.48</td>
<td>1.63E−02</td>
<td>AT4G37750</td>
<td>AINTEGUMENTA (ANT)</td>
<td>0.93</td>
<td>3.75E−69</td>
<td>Potri.014G012200</td>
</tr>
<tr>
<td>AT1G32170</td>
<td>Xyloglucan endotransglucosylase/hydrolase 30</td>
<td>Potri.003G097300</td>
<td>−0.61</td>
<td>1.51E−03</td>
<td>AT5G02030</td>
<td>REPLUMLESS (RPL)</td>
<td>0.86</td>
<td>5.29E−51</td>
<td>Potri.010G197300</td>
</tr>
<tr>
<td>AT1G11580</td>
<td>Methylesterase PCA</td>
<td>Potri.011G135000</td>
<td>−0.68</td>
<td>2.91E−04</td>
<td>AT5G02030</td>
<td>REPLUMLESS (RPL)</td>
<td>0.90</td>
<td>1.69E−59</td>
<td>Potri.010G197300</td>
</tr>
<tr>
<td>AT2G42800</td>
<td>Receptor like protein 29</td>
<td>Potri.008G211800</td>
<td>−0.48</td>
<td>1.63E−02</td>
<td>AT5G37970</td>
<td>HBS1</td>
<td>0.84</td>
<td>1.17E−46</td>
<td>Potri.006G117700</td>
</tr>
<tr>
<td>AT2G28760</td>
<td>UDP-XYL synthase 6</td>
<td>Potri.010G207200</td>
<td>−0.62</td>
<td>1.36E−03</td>
<td>AT5G05790</td>
<td>myb family transcription factor</td>
<td>0.86</td>
<td>6.09E−50</td>
<td>Potri.010G193000</td>
</tr>
<tr>
<td>AT1G11580</td>
<td>Methylesterase PCA</td>
<td>Potri.011G135000</td>
<td>−0.68</td>
<td>2.91E−04</td>
<td>AT5G05790</td>
<td>myb family transcription factor</td>
<td>0.84</td>
<td>1.09E−47</td>
<td>Potri.010G193000</td>
</tr>
<tr>
<td>AT1G11580</td>
<td>Methylesterase PCA</td>
<td>Potri.011G135000</td>
<td>−0.68</td>
<td>2.91E−04</td>
<td>AT5G06800</td>
<td>myb family transcription factor</td>
<td>0.83</td>
<td>3.00E−45</td>
<td>Potri.006G191000</td>
</tr>
<tr>
<td>AT1G02810</td>
<td>Plant invertase/pectin methylesterase inhibitor superfamily</td>
<td>Potri.014G127000</td>
<td>−0.54</td>
<td>6.82E−03</td>
<td>AT5G10280 MYB DOMAIN PROTEIN 92 (ATMYB92)</td>
<td>0.83</td>
<td>5.55E−46</td>
<td>Potri.007G093900</td>
<td>−0.43</td>
</tr>
<tr>
<td>AT2G21350</td>
<td>Arabinogalactan protein 17</td>
<td>Potri.007G051600</td>
<td>0.56</td>
<td>4.63E−03</td>
<td>AT5G12330 LATERAL ROOT PRIMORDIUM 1 (LRP1)</td>
<td>0.81</td>
<td>2.59E−43</td>
<td>Potri.003G196100</td>
<td>0.65</td>
</tr>
<tr>
<td>AT3G22800</td>
<td>Leucine-rich repeat (LRR) family protein</td>
<td>Potri.010G083000</td>
<td>−0.50</td>
<td>1.31E−02</td>
<td>AT5G14750 MYB DOMAIN PROTEIN 66 (ATMYB66)</td>
<td>0.81</td>
<td>1.11E−42</td>
<td>Potri.003G064600</td>
<td>0.53</td>
</tr>
<tr>
<td>AT1G02810</td>
<td>Plant invertase/pectin methylesterase inhibitor superfamily</td>
<td>Potri.014G127000</td>
<td>−0.54</td>
<td>6.82E−03</td>
<td>AT5G10280 MYB DOMAIN PROTEIN 92 (ATMYB92)</td>
<td>0.83</td>
<td>5.55E−46</td>
<td>Potri.007G093900</td>
<td>−0.43</td>
</tr>
<tr>
<td>AT2G36870</td>
<td>Xyloglucan endotransglycosylase/hydrolase 32</td>
<td>Potri.009G006600</td>
<td>0.51</td>
<td>1.08E−02</td>
<td>AT5G15210 ARABIDOPSIS THALIANA HOMEOBOX PROTEIN 30</td>
<td>0.82</td>
<td>8.67E−44</td>
<td>Potri.004G126500</td>
<td>−0.48</td>
</tr>
<tr>
<td>AT2G42800</td>
<td>Receptor like protein 29</td>
<td>Potri.008G211800</td>
<td>−0.48</td>
<td>1.63E−02</td>
<td>AT5G37020 AUXIN RESPONSE FACTOR 8 (WRKY27)</td>
<td>0.94</td>
<td>4.72E−70</td>
<td>Potri.001G358500</td>
<td>−0.53</td>
</tr>
<tr>
<td>AT1G11580</td>
<td>Methylesterase PCR A</td>
<td>Potri.011G135000</td>
<td>−0.68</td>
<td>2.91E−04</td>
<td>AT5G52830 myb domain protein 36 (MYB36)</td>
<td>0.81</td>
<td>1.06E−42</td>
<td>Potri.017G149000</td>
<td>0.51</td>
</tr>
<tr>
<td>AT1G11580</td>
<td>Methylesterase PCR A</td>
<td>Potri.011G135000</td>
<td>−0.68</td>
<td>2.91E−04</td>
<td>AT5G52830 myb domain protein 36 (MYB36)</td>
<td>0.81</td>
<td>1.06E−42</td>
<td>Potri.017G149000</td>
<td>0.51</td>
</tr>
<tr>
<td>AT1G11580</td>
<td>Methylesterase PCR A</td>
<td>Potri.011G135000</td>
<td>−0.68</td>
<td>2.91E−04</td>
<td>AT5G52830 myb domain protein 36 (MYB36)</td>
<td>0.81</td>
<td>1.06E−42</td>
<td>Potri.017G149000</td>
<td>0.51</td>
</tr>
<tr>
<td>AT1G11580</td>
<td>Methylesterase PCR A</td>
<td>Potri.011G135000</td>
<td>−0.68</td>
<td>2.91E−04</td>
<td>AT5G52830 myb domain protein 36 (MYB36)</td>
<td>0.81</td>
<td>1.06E−42</td>
<td>Potri.017G149000</td>
<td>0.51</td>
</tr>
<tr>
<td>AT1G11580</td>
<td>Methylesterase PCR A</td>
<td>Potri.011G135000</td>
<td>−0.68</td>
<td>2.91E−04</td>
<td>AT5G52830 myb domain protein 36 (MYB36)</td>
<td>0.81</td>
<td>1.06E−42</td>
<td>Potri.017G149000</td>
<td>0.51</td>
</tr>
<tr>
<td>AT1G11580</td>
<td>Methylesterase PCR A</td>
<td>Potri.011G135000</td>
<td>−0.68</td>
<td>2.91E−04</td>
<td>AT5G52830 myb domain protein 36 (MYB36)</td>
<td>0.81</td>
<td>1.06E−42</td>
<td>Potri.017G149000</td>
<td>0.51</td>
</tr>
<tr>
<td>AT1G11580</td>
<td>Methylesterase PCR A</td>
<td>Potri.011G135000</td>
<td>−0.68</td>
<td>2.91E−04</td>
<td>AT5G52830 myb domain protein 36 (MYB36)</td>
<td>0.81</td>
<td>1.06E−42</td>
<td>Potri.017G149000</td>
<td>0.51</td>
</tr>
</tbody>
</table>

**Co-expression of Arabidopsis prey genes with bait genes was analyzed across multiple microarray experiments using CressExpress ([http://cressexpress.org/](http://cressexpress.org/)).** Ath Locus: *Arabidopsis thaliana* gene locus; Ptr Locus: gene locus IDs in the Phytozome *Populus trichocarpa* v3.0 database; \( r_{\text{Sacc}} \): Pearson’s coefficient for the correlation of *Populus nigra* best hits’ gene transcript abundance with saccharification potential; \( P(r_{\text{Sacc}}) \): P-values for the respective correlations. \( r_{\text{coexp}} ^2 \) and \( P(r_{\text{coexp}} ^2) \): squared Pearson’s correlation coefficient and P values for the correlation of the transcript abundances of a bait with co-expressed Arabidopsis gene.
from the wettest habitat (Table 1, Figure 1). Previous studies also reported inconsistent results concerning the relationship between lignin content and saccharification potential. Some studies with transgenic poplars with low lignin content reported negative correlations between lignin and saccharification potential (Min et al. 2012, Acker et al. 2014), while Voelker et al. (2010) found unaffected saccharification potential when lignin decrease was modest, and decreasing saccharification in transgenic poplar lines with strongly reduced lignin content. Studer et al. (2011) did not find a general negative relationship between natural variation in lignin content of P. trichocarpa and sugar release, but reported that the effect of lignin on saccharification depends on the lignin composition and pretreatment methods. In Miscanthus natural genetic variation in lignin content was not consistently correlated with saccharification potential (Souza et al. 2015). In a study with maize, QTLs for lignin content were different from those detected for saccharification (Penning et al. 2014). Altogether, these studies suggest that lignin is not a good predictor for the saccharification potential.

Transcriptional regulation of genes involved in cell wall polysaccharide biosynthesis and modification is related to genotypic variation in saccharification potential

To uncover the molecular basis for differences in saccharification potential we re-constructed transcriptional networks and found that co-expression modules correlated with this trait were enriched in distinct sets of GO terms (Table S2 available as Supplementary Data at Tree Physiology Online). This finding underpins the assumption that modules summarize functionally related genes. The functional categories pointed to an involvement of cell wall metabolism, especially cell wall polysaccharides (bin 10), and vesicle-associated secretory processes (bins 31.4 and 29.3.4) in the control of saccharification potential. In agreement with lacking correlation between lignin content and saccharification potential, categories related to lignin biosynthesis were not enriched, suggesting that variation in expression of genes involved in cell wall polysaccharide biosynthesis is more important for saccharification potential than moderate variations in lignin. However, among the top saccharification-correlated genes, none of the biosynthetic genes for cell wall polysaccharides was detected. Instead three putative kinases (Potri.002G132900, Potri.007G039800, Potri.010G155600) and one gene involved in brassinosteroid metabolism (Potri.011G155600) were present among the top 10 correlated genes (Figure 4), suggesting that signaling and hormone regulation are important check points for the molecular control of saccharification potential.

Surprisingly, most GDEGs assigned to cellulose biosynthesis (bins 10.2, 10.2.1, 10.2.2) were negatively correlated with saccharification potential (Figure 5A, Table S5 available as Supplementary Data at Tree Physiology Online). Saccharification potential depends not only on cellulose content, but also on the level of cellulose aggregation and thus accessibility for enzymatic degradation (Nookaraju et al. 2013). Primary cell walls contain cellulose microfibrils with low aggregation, whereas during the synthesis of secondary cell walls, highly aggregated microfibrils are formed (Li et al. 2016). Synthesis of cellulose in primary and secondary cell walls involves distinct sets of genes (Li et al. 2016). It is thus possible that the GDEGs that were negatively correlated with saccharification potential (Table S5 available as Supplementary Data at Tree Physiology Online) represent poplar cellulose synthase genes involved in the synthesis of highly aggregated cellulose microfibrils.

Although pectins comprise only a small fraction of secondary cell walls (Harholt et al. 2010) there is accumulating evidence that they play a role in determining saccharification potential, by mediating cell adhesion, masking cellulose and hemicelluloses and thus restricting their accessibility by enzymes (Xiao and Anderson 2013, Tavares et al. 2015). The effect of pectins on saccharification is dependent on the degree of pectin acetylation and methyl-esterification (Xiao and Anderson 2013). Once secreted into the apoplast, pectins can be de-methylesterified by pectin methyltransferases, enabling cross-linking and possibly increasing cell wall stiffness (Xiao and Anderson 2013). A reduction of de-methylesterified pectins was found to increase saccharification potential in Arabidopsis, wheat and tobacco (Lionetti et al. 2010). In agreement with these findings, putative pectin methyltransferases (bins 10.8.1 and 10.8.2) were mostly negatively correlated with saccharification potential (Figure 5B, Table S5 available as Supplementary Data at Tree Physiology Online).

The major compound of hemicellulose and second most abundant polysaccharide in woody dicots is glucuronoxylan (Awano et al. 1998, Lee et al. 2009). Here the transcript abundances of two genes related to glucuronoxylan biosynthesis (Potri.002G132900 and Potri.016G086400, of bin 10.3.2) were negatively correlated with saccharification potential (Figure 5A, Table S5 available as Supplementary Data at Tree Physiology Online). There is experimental evidence that the P. × canescens orthologs of Potri.002G132900 and Potri.016G086400 are functional homologs of Arabidopsis PARVUS (Lee et al. 2009) and IRX9 (Zhou et al. 2007, Lee et al. 2011), respectively. PARVUS (At1g19300) and IRX9 (At2g37090) are required for the synthesis of the tetrasaccharide reducing end sequence of glucuronoxylans, and the xylosyl backbone of glucuronoxylans, respectively (Brown et al. 2005, 2007, Lee et al. 2007a, 2007b, Peña et al. 2007). Lee et al. (2011) found that RNAi down-regulation of the P. × canescens ortholog of Potri.016G086400 resulted in an increased glucose release from wood. Remarkably, UDP-glucuronic acid decarboxylase (UXS) transcript abundances were negatively correlated with saccharification potential (GDEGs in bin 10.1.5, Figure 5B, Table S5 available as Supplementary Data at Tree Physiology Online). UXS enzymes catalyze the irreversible synthesis of UDP-xylene from UDP-glucuronic acid (Kuang et al. 2016), and are key enzymes for nucleotide sugar partitioning for cell wall
polysaccharide biosynthesis (Du et al. 2013). In transgenic tobacco, down-regulation of UXS genes led to a significant increase in saccharification efficiency (Cook et al. 2012). Taken together, these findings illustrate the power of our experimental and analytical approach to identify candidate genes related to the regulation of wood traits.

**Putative transcriptional regulators of saccharification-related genes involved in cell wall metabolism**

To identify putative transcriptional regulators of saccharification-related genes involved in cell wall metabolism, we used the identified cell wall genes as baits in an in silico co-expression analysis. Among the genes with a significant correlation with saccharification, 31 potential transcription factors were identified that showed significant co-expression with cell wall biosynthesis genes in *Arabidopsis* and were also present among the poplar genes, significantly correlated with saccharification (Table 3). Among these genes, MYB domain, WRKY and homeobox domain transcription factors that belong to families known to be involved in wood formation were retrieved (Hussey et al. 2013, Yang and Wang 2016). For example, two putative pectin esterases (Potri.014G127000, Potri.011G135000) were co-expressed with WRKY, MYB and bZIP transcription factors. Interestingly, a gene encoding a putative UXS protein (Potri.010G207200), which appears to be important for the saccharification potential as discussed above, was co-expressed with a MYB family (Potri.010G193000) and TGA1 (Potri.007G085700) transcription factor. MYB transcription factors are known to play a role in the regulation of xylan biosynthesis (Rennie and Scheller 2014), but to our knowledge, no published results are available for the specific MYB family members highlighted by our study.

In addition, the gene set of co-expressed transcription factors contained novel candidate regulators such as TGACG SEQUENCE-SPECIFIC BINDING PROTEIN 1 (TGA1), SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 4 (SPL4), AGAMOUS-LIKE 8 (AGL8), SCARECROW, AUXIN RESPONSE FACTOR 8 (ARF8), ANTEGUMENTA (ANT), etc. These transcription factors play roles in redox regulation (TGA) and development. Our approach suggests that developmental changes brought about by drought or genetic differentiation among the genotypes also control cell wall composition that affects saccharification. Therefore, the present correlative results provide an excellent platform for future targeted approaches to improve the saccharification potential of poplar.

**Conclusions**

The efficient conversion of lignocellulosic biomass to biofuels varies with wood composition, which itself is subject to natural genetic variation and is dependent on environmental conditions, such as water availability. Previous research on wood traits that may affect saccharification mainly focused on lignin, revealing opportunities for improving biomass quality by reducing lignin content. However, there is also growing awareness that this approach has limitations (Voelker et al. 2010, Tavares et al. 2015). By studying variation in wood traits and saccharification potential in three genotypes of *P. nigra* exposed to a moderate drought treatment, we show that glucose release was not impaired, but moderately improved, by gradually decreasing water availability. Interestingly, the saccharification potential was not related to lignin content or expression of genes related to lignin biosynthesis. Instead, our research identified transcriptional regulation of biosynthesis of hemicelluloses and modification of pectins as potential targets for improving wood quality for the production of biofuels. Further experiments are needed in which the expression of the identified candidate genes is modified by overexpression or suppression to test the causal relationship of the identified genes and pathways with saccharification potential.

**Supplementary Data**

Supplementary Data for this article are available at Tree Physiology Online.

**Conflict of interest**

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Author contributions**


**Funding**

This work was conducted in the frame of the WATBIO (Development of improved perennial biomass crops for water-stressed environments), which is a collaborative research project funded from the European Union’s Seventh Programme for research, technological development and demonstration under grant agreement No. 311929. The research has received funding from the French National Research Agency through the Laboratory of Excellence ARBRE (ANR-12-LABXARBRE-01). S.P. received a Ph.D. scholarship from the program Erasmus Mundus, India4EUII.
Wildhagen et al.

Acknowledgments

We thank C. Kettner (Forest Botany and Tree Physiology, Göttingen, Germany) for plant maintenance and plant production and T. Klein (Laboratory for Radio-Isotopes, Göttingen, Germany) for RNA isolation. We acknowledge the providers of the original P. nigra genotypes ‘France 61-29’ (INRA, Paris, France represented by G. Pilate) and ‘Spain RIN2-new’ (CITA, Zaragoza, Spain, represented by JV Lacasa Azlor) and C. Bastien (INRA, Orleans, France) for providing the stock cuttings.

References


Lee C, Teng Q, Zhong R, Richardson EA, O


O’Brien TP, Feder N, McCully ME (1964) Polymeric staining of plant cell walls by toluidine blue O. Protoplasma 59:368–373.


