BioEnergy Research (2018) 11:351–363
https://doi.org/10.1007/s12155-018-9901-8

QTL Mapping of Wood FT-IR Chemotypes Shows Promise for Improving Biofuel Potential in Short Rotation Coppice Willow (*Salix* spp.)

Prashant Mohan-Anupama Pawar[1,2] · Anna Schnürer[3] · Ewa J. Mellerowicz[1] · Ann Christin Rönnberg-Wästljung[4]

Published online: 9 March 2018

Abstract

An increasing interest to convert lignocellulosic biomass into biofuels has highlighted the potential of using willows for this purpose, due to its fast growth in short rotation coppice systems. Here, we use a mapping population of 463 individuals of a cross between *Salix viminalis* and *S. viminalis* × *S. schwerinii* to investigate the genetic background of different wood chemical traits, information of importance for breeding towards different uses of wood. Furthermore, using a subset of the mapping population, the correlation between biogas production and chemical traits was investigated. The phenotyping of wood was carried by Furrier-transformed-Infrared spectrometry (FT-IR) and water content analysis. Quantitative trait loci (QTLs) analysis was used to identify regions in the genome of importance for the phenotypic variation of these chemical traits. We found 27 QTLs for various traits. On linkage group (LG) VI-1, QTLs for signals assigned to G-lignin, lignin, and the S/G ratio were collocated and on LG XIV we found a cluster of QTLs representing signals assigned to lignin, cellulose, hemicellulose, and water. The QTLs explained from 3.4 to 6.9% of the phenotypic variation indicating a quantitative genetic background where many genes influence the traits. For the biogas production, a positive and negative correlation was seen with the signals assigned to acetyl and lignin, respectively. This study represents a first step in the understanding of the genetic background of wood chemical traits for willows, information needed for complementary studies, mapping of important genes, and for breeding of varieties for biofuel production purposes.

Keywords *Salix* · Wood traits · QTL · Candidate genes · Biogas · Plant breeding · FT-IR spectroscopy

Introduction

The requirement to reduce the world's emission of greenhouse gases and to use more bio-based energy systems, not as strongly dependent on fossil fuels, has increased the interest for cultivation of bioenergy crops. Species from the *Saliceaceae* family (poplars and willows) are interesting alternatives of lignocellulosic bioenergy crops and they share many phenotypic and genetic properties such as fast growth, ease of vegetative propagation, similar genome size (~ 500 Mbp), and reasonably high levels of genetic diversity [1, 2]. These species have shown high biomass production in short rotation coppice (SRC) systems where plants are cut down in repeated cycles with a vigorous re-sprouting after each cutback [3]. Hybrid cultivars of willows have been available as bioenergy crop for heat and power since the late 1980s and put into practice in Sweden, other European countries [4] as well as in the USA [5]. Recently, there has also been an increasing interest of using lignocellulosic biomass from poplars and willows as raw material for conversion into biofuels such as biogas and bioethanol [3, 6, 7]. To improve the yields of biofuel as well as to understand the recalcitrance of the cell wall and pre-treatments to break the structure, intensive studies have been conducted in recent years (reviewed by [8, 9]). Natural variant screening has also been carried out in different species, revealing the genetic variation in biomass chemical traits that are thought to be important for recalcitrance, e.g., [10, 11].

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12155-018-9901-8) contains supplementary material, which is available to authorized users.

✉ Ann Christin Rönnberg-Wästljung
anki.wastljung@slu.se

[1] Department of Forest Genetics and Plant Physiology, Umeå Plant Science Center, Swedish University of Agricultural Sciences, SE-90183 Umeå, Sweden

[2] Present address: Department of Biochemistry, Purdue University, West Lafayette, Indiana 47907, USA

[3] Department of Molecular Sciences, Uppsala BioCenter, Linnean Center of Plant Biology in Uppsala, Swedish University of Agricultural Sciences, POB 7025, SE-750 07 Uppsala, Sweden

[4] Department of Plant Biology, Uppsala BioCenter, Linnean Center of Plant Biology in Uppsala, Swedish University of Agricultural Sciences, POB 7080, SE-750 07 Uppsala, Sweden

Until now, most of the plant breeding of poplar and willow cultivars for bioenergy use has focused on biomass growth and resistance traits, but to develop cultivars specifically for biofuel use there is a need for understanding the genetic background of traits connected to wood and cell walls. Genetic variation of different wood traits related to biofuel production potential, such as content of lignin, hemicellulose, and cellulose, has been used in clonal experiments of poplar families for QTL analysis and in experiments with unrelated plant material for association studies, e.g., [12–18]. Despite the domestication of poplars, there is still substantial natural variation in different wood traits [10, 16, 19]. Muchero et al. [17] detected a number of candidate genes for wood traits using a combined approach with QTL analysis as first step to identify important regions and then association mapping to identify genes. Association mapping has been conducted in several *Populus* species, both using a candidate gene approach with a priori known genes from different wood component pathways as well as a genome wide approach with thousands of SNP markers spread across the whole genome [16, 20, 21]. A set of interesting candidate genes have been identified and in several cases the effect of specific alleles could be demonstrated [16, 17].

In willows, fewer studies are available of the genetic background of wood traits. The early studies indicated genetic variation and moderate-sized heritabilities for wood traits [22, 23]. More recently, Ray et al. [24] and Serapiglia et al. [25] demonstrated variation between willow clones in cell wall traits and enzymatic saccharification. The only QTL study connected to cell wall traits in willows identified four QTLs for enzymatic saccharification [26].

The abovementioned studies indicate possibilities for recurrent selection to increase desirable traits towards more suitable plant material for biofuels. The emerging knowledge of specific regions in the genome and genes that influence wood traits could also be used in the breeding for selection of interesting individuals (marker assisted selection, MAS) at an early stage of the breeding process and thus increase the breeding efficiency [27, 28].

The close relationship between *Salix* and *Populus* and earlier identified macrosyntheny between the genomes of the two genera [29, 30] makes it possible to utilize the genomic resources available for *Populus* when studying genomic regions in *Salix*. In this study, we have used a mapping population that is based on a cross between *Salix viminalis* and the hybrid *S. schwerinii x S. viminalis*, two species that are commonly employed in plant breeding for development of high biomass-producing willow varieties in Europe. Our aims are to identify genomic regions and candidate genes important for cell wall characteristics to get a better understanding of the genetic background and of the possibilities for breeding towards willow varieties better adapted for production of biofuels. The biofuel possibilities were evaluated analyzing the biochemical methane potential (BMP) of a clonal subset of the mapping population.

Materials and Methods

Plant Material and Field Experiment

The population studied is the S_1 mapping population that consists of 463 F_1 progeny from the cross between the diploid female *S. viminalis* clone "78183" and the diploid hybrid male clone "Björn" (*Salix schwerinii* E. Wolf × *Salix viminalis* L.) [30]. The *S. schwerinii* parent (79069) to Björn originates from Siberia while the *S. viminalis* parent is an interspecific cross between the female clone 78195 originating from southern Sweden, and the male clone 78101 from Western Sweden. The S_1 population female parent, 78183, also originates from southern Sweden. Initially, the parental clones to S_1 were selected based on variation in phenology traits [31], but the offspring population has also shown variation in different growth traits and in drought response [32]. Based on this earlier documentation, we expect also a variation in different wood traits for the population.

In spring 2008, plants of 463 genotypes and the two parental clones were planted in an experimental field in Pustnäs, south of Uppsala (59° 48' N, 17° 39′ E, 25 m). The spacing was 130 × 50 cm (i.e., about 20,000 plants ha^{-1}), and the experiment followed a randomized complete block design comprising six blocks with one plant per genotype in each. Two border rows were planted around the experiment to reduce edge effects. The plants were obtained by 5-cm hardwood cuttings rooted in 0.2-L peat pots with Weibulls "Kron Mull" as growing medium. Before planting, the site was appropriately prepared [33], including plowing, harrowing, and repeated application of a systemic herbicide (Glyphomax, Dow AgroSciences, Indianapolis, IN). The plantation was irrigated in summer 2008 and weed controlled during the whole experimental period. The plants were cut back winter 2009 and 2011 and fertilized spring 2009, 2010, and 2011 with N P K (21-4-7) equal to an amount of 80 kg N ha^{-1} and year. For further details, see Ghelardini et al. [34].

Chemotyping and Biogas Experiments

In March 2012, wood samples from each individual plant in two randomly selected blocks were collected. The main stem from each plant was selected and cut off at the base, the weight equilibrium point of the shoot was determined, and a sample of 30 cm with this point in the center of the sample was taken. This sampling method will give a dry matter content representing the total shoot mean dry matter content [35]. The cuttings were stored in a freezer until FT-IR and biogas potential analyses.

Preparation of Wood Powder

The stem samples with bark were weighed, freeze-dried for 48 h, and weighed again to determine their water content. Approximately 3.5 g of each sample was cut out from one end using secateurs. This portion was further cut into matchstick size pieces using a razor blade and ground to a rough powder using Ultra Centrifugal mill ZM 200 (Retsch, Haan, Germany) with a 0.5-mm size sieve. Two hundred milligrams of rough powder was milled to a fine powder using mixer mill (Retsch®MM400, Hann, Germany), and the rest was kept for the biogas analysis.

FT-IR

Ten milligrams of fine wood powder was mixed with 390 mg of KBr and ground using an amethyst mortar and pestle. Spectra were recorded using the Diffuse Reflectance method under vacuum, using a Bruker IFS 66v/S spectrometer (Bruker Optik GmbH, Ettlingen, Germany), at a spectral resolution of 4 cm^{-1} between 400 and 1900 cm^{-1}, as explained in [36]. Baseline correction and normalization using 64-point rubber band and offset- and vector-normalization were applied using OPUS (version 7.0.122; Bruker Optik). Data were analyzed using the SIMCA-P+ software (version 12.0.0.0, Umetrics AB, Sweden). Principal Component Analysis (PCA) was carried to examine variation between block 3 and block 5 and to remove the outliers [37]. For the QTL analysis, spectral range between 793 and 1834 cm^{-1} was considered [36].

Pyrolysis-GC-MS

Wood powder from the samples representing the genotypes with contrasting FT-IR signals at 1207 and 1605 cm^{-1}, assigned to G-lignin and lignin, respectively, was analyzed by pyrolysis-GC-MS according to [38] to verify the used FT-IR assignments.

Evaluation of Biogas Potential

The biochemical methane potential (BMP) of the different *Salix* clones was determined using an automatic methane potential test system (Bioprocess Control AB, Sweden). Rough dry powder from 50 randomly selected clones from two blocks was pooled and 2.1 g was transferred to triplicate test vials. The samples were mixed with 400 ml inoculum, taken from a full-scale sludge biogas digester operated at a wastewater treatment plant in Uppsala, Sweden. Before start, the inoculum was incubated for 4 days at 37 °C to reduce the endogenous methane production (from inoculum). Total solids (TS) and volatile solids (VS) of the inocula were measured according to international standard methods [39] and determined to be 2.6 and 1.6% of wet weight. Background methane from endogenous material in the inoculum was determined in triplicate using the same amount of inoculum but without addition of substrates. Cellulose in the same amount as the *Salix* powder (SIGMA, Cellulose fibres medium CAS 9004-34-6) was used as a standard in a separate set of triplicate bottles. The bottles were incubated during 45 days at 37 °C and obtained accumulated methane values were standardized to normal atmospheric pressure (atm) and 0 °C (273.15 K, 1 bar). The accumulated amount of methane was plotted over time and the methane produced from the inoculum only was withdrawn. The final value obtained after leveling off (40–43 days) was considered as the final methane potential (N mL CH_4 (g VS)$^{-1}$) [40].

Statistical Analysis and QTL Mapping

From the FT-IR spectra, wavelengths corresponding to main spectrum peaks (Online Resource 1) and regions representing interesting wood traits were selected (Table 1). Analysis of variance was conducted to estimate variance components (REML procedure) on water content and selected FT-IR wavenumbers data using the program JMP ® 10 [51] and applying the model:

$$Y_{ijk} = \mu + B_i + G_j + e_{ijk} \quad (1)$$

where Y_{ijk} is the phenotypic value for the *j*th genotype in the *i*th block, *u* is the overall mean, B_i is a fixed effect of the *i*th block, *G*j is the random the effect of genotype, and e_{ijk} is the residual error.

Broad sense heritability (H^2) was estimated for each trait using the formula:

$H^2 = \sigma^2_g/(\sigma^2_g + (\sigma^2_e/b))$, where σ^2_g and σ^2_e equals the genetic variance and error variance, respectively, *b* equals number of blocks.

Phenotypic correlations between all traits were estimated with Pearson product-moment correlations in JMP ® 10 [51] based on mean values for each trait and genotype.

The S_1 linkage map presented in [30] was used in the QTL analyses. The consensus linkage map consists of 495 markers covering the 19 linkage groups in willows. In a few cases, markers segregating in both parents were not available in part of a linkage group and thus the map for this part consists of a male or a female map [30]. QTL analyses were performed on mean values for each genotype and trait (water content, selected FT-IR wavenumbers, biogas potential, and number of days to reach 50 and 80% of the final biogas potential (the biogas data were from a subset of the mapping population)) using MapQTL 6.0 [52]. In a stepwise analysis starting with interval mapping (IM) and a regression model, the genome was scanned at 1 cM intervals to determine putative QTLs involved in the variation of each trait. A significance threshold value, estimated as the logarithm of odds ratio (LOD), to determine significant QTL was estimated with a permutation test

Table 1 Assignment of different compounds measured by FT-IR

FT-IR wavenumber (cm^{-1})	Assigned compound	Description[a)]	Reference
1128	S-lignin	Aromatic C-H in-plane deformation	[41]
1169	Acetyl groups in esterified hemicelluloses	C-O-C stretching	[42]
1207	G-lignin	C-O stretching	[43]
1229	G-lignin	Aromatic ring breathing with C-O stretching	[41]
1254	Acetyl groups in esterified hemicelluloses	C-O-C stretching	[42]
1271	G-lignin	Aromatic ring and C=O stretch	[44]
1360	Cellulose/hemicellulose/acetyl groups	CH-bending	[45]
1429	Cellulose/hemicellulose	C-H deformation, (CH_2) symmetrical bending	[45, 46]
1466	Lignin	CH_3 asymmetric bending	[45]
1510	Lignin	Aromatic C=C stretching (skeletal vibration)	[41, 44, 45, 47–50]
1593 and 1605	Lignin	Aromatic C=C stretching (skeletal vibration)	[41, 45, 47–49]
1666	Water	O-H deformation	[44, 46]
1709		Unconjugated carbonyl/carboxyl stretch	[44]
1734	Acetyl groups in esterified hemicelluloses	C=O	[45]
1128/1207		S/G ratio	
1510/1593		Lignin cross-linking (G/S ratio)	[49]

[a)] The description is based on information provided in [36]

of 1000 repetitions. A genome-wide threshold for a significant QTL was set at $p = 0.05$. In a second step, to get a more precise QTL estimate, the markers closest to the significant QTLs based on the IM analysis were used as cofactors in a multiple QTL model (MQM) analysis with a regression model. One and two LOD confidence interval for each QTL were estimated using the LOD drop-off method [53] based on the LOD-value at the peak position of the QTL. The proportion of the phenotypic variation explained by each significant QTL was estimated (% Expl.). The difference between maternal alleles were estimated as the absolute effect of $A_m = ((\mu_{ac} + \mu_{ad}) - (\mu_{bc} + \mu_{bd}))/2$, the difference in paternal alleles were estimated as the absolute effect of $A_p = ((\mu_{ac} + \mu_{bc}) - (\mu_{ad} + \mu_{bd}))/2$ and the paternal-maternal interaction effect was estimated as the absolute effect of $A_i = ((\mu_{ac} + \mu_{bd}) - (\mu_{ad} + \mu_{bc}))/2$ where μ_{ac}, μ_{bc}, μ_{ad}, and μ_{bd} are the estimated phenotypic means of the four genotypic classes ac, bc, ad, and bd obtained from an ab × cd cross [52].

Candidate Genes from QTL-Regions

The DNA-sequence of the markers closest to the flanking 2-LOD QTL-regions was BLAST searched towards the *Populus trichocarpa* genome v. 3.0 (http://www.phytozome.net/poplar.php) to identify their positions in the *Populus* genome. If the marker was an AFLP-marker, no DNA-sequence was available and thus the second closest marker was used. Using these positions, gene models within the QTL-regions (Fig. 2) were subtracted with the Biomart tool from Phytozome v 9.0 database. Candidate genes defined as "genes related to cell wall biosynthesis" were identified based on the predicted gene function available from the annotation, and from the CAZyme classification [54].

Results

Variation of Wood Traits

The phenotypic distribution of FT-IR traits as well as water content are presented in Fig. 1. For all analyzed traits, a considerable phenotypic variation existed in the offspring population that in many cases exhibited mean values above or below the parental mean values. The water content of the wood samples for offspring population varied between 37 and 54% of the total sample weight (Fig. 1).

The heritability (H^2) estimates for the traits were low to medium (Table 2). The highest heritabilities (30–41%) were

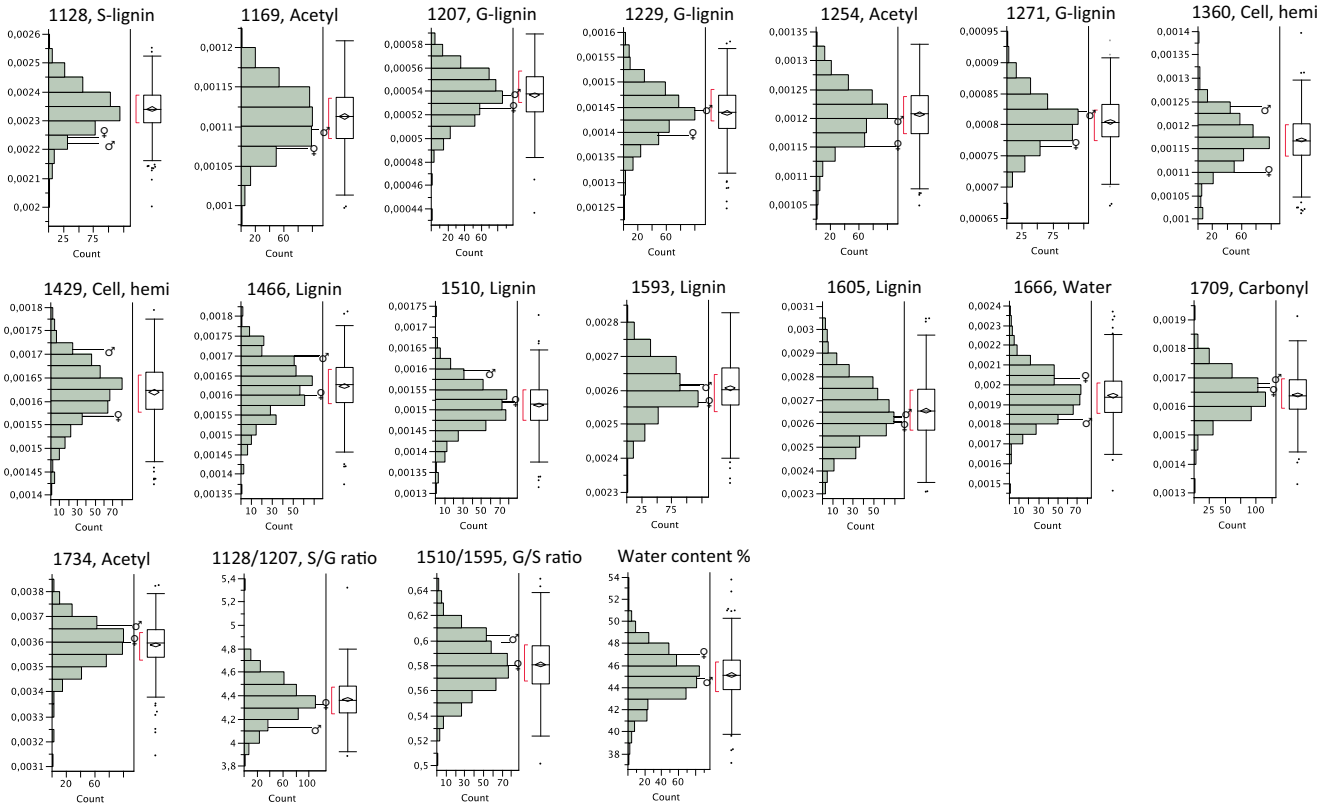

Fig. 1 Phenotypic distributions and box plots of selected FT-IR wavenumbers and water content for the offspring population S_1, based on mean values for each genotype. Mean values for the female parental genotype (78183) and the male parental genotype (Björn) are indicated in the figure. Abbreviations: *Cell, hemi* - cellulose and/or hemicellulose

seen for FT-IR signals assigned to cellulose, hemicellulose (1429 cm^{-1}), G/S ratio, G-lignin (1271 cm^{-1}), and lignin (1510 cm^{-1}), and for the water content (Table 2). The phenotypic variation seen in Fig. 1 for the low heritability traits is mainly due to environmental causes. A small difference between the environments (blocks) in the FT-IR study is illustrated by the principal component analysis (Online Resource 2).

Phenotypic correlations between all traits are reported in Table 3. Correlations between selected FT-IR wavenumbers (in cm^{-1}) and the water content are low and mostly nonsignificant (Table 3). Significant positive correlations are found between different FT-IR signals, for example between wavenumbers assigned to acetyl (1169) and S-lignin (1128), between acetyl (1169) and G-lignin (1207, 1229, 1271), between wavenumbers assigned to cellulose and hemicellulose (1360 and 1429), lignin (1510, 1466), G-lignin (1271), and G/S ratio (1510/1593) Significant negative correlations were found between wavenumbers assigned to cellulose and hemicellulose (1360) and water (1666), and expectedly between G-lignin (1207) and the S/G ratio (1128/1207). Similarly, between lignin (1510) and S/G ratio (1128/1207) as well as between lignin (1605) and G/S ratio (1510/1593). Expectedly, signals assigned to acetyl at 1169 and 1254 cm^{-1} were positively correlated. However, the vibrations at 1734 cm^{-1}, which were also assigned to acetyl, were not significantly correlated with those signals but they were positively correlated with signals assigned to carboxyl/carbonyl (1709) and lignin (1510). This suggests subtle difference exists either in different molecular forms or in molecular context, or in contribution of other compounds between 1169 and 1254 versus 1734 cm^{-1} signals. Similarly, the lignin forms corresponding to signals at 1128, 1207, 1229, 1271, and 1466 appear not to be correlated with those represented by signals at 1593 and 1605 cm^{-1}.

QTL Mapping

In total, 27 QTLs were identified for FT-IR wavenumbers assigned to different compounds, and for water content (Table 4, Fig. 2). The phenotypic variation explained by each of the different QTLs varied from 3.4 to 6.9%, showing a quantitative genetic background where each of the QTL explains a small part of the total variation. Several QTLs were identified for many of the traits, e.g., for acetyl (1169 cm^{-1}) three QTLs at linkage groups (LGs) III, VII, and IX were found, for lignin (1510) two QTLs at the LGs VI-1 and XII were found. No QTL were identified for FT-IR wavenumber 1743 cm^{-1} assigned to the acetyl esters. A grouping of the assigned FT-IR wavenumbers by the compounds such as S-lignin, acetyl, G-lignin, cellulose or hemicellulose, and lignin identified one, five, four, four, and five QTLs, respectively

Table 2 Mean values, variance components (using model 1) and heritabilites (H^2) for FT-IR signals, water content and acetic acid estimated for the S_1 population

FT-IR wavenumber (cm^{-1})	Mean values × 10^{-5}	Var $_{genotype}$ (%)	Var $_{error}$ (%)	H^2_{mean} [a] (%)
1128	234.0	1	99	3
1169	111.2	12	88	22
1207	53.7	11	89	20
1229	143.9	13	87	24
1254	120.7	12	88	21
1271	80.5	26	74	41
1360	116.9	16	84	28
1429	161.9	18	82	30
1466	162.3	16	84	28
1510	151.2	28	72	43
1593	260.6	11	89	20
1605	265.5	14	86	25
1666	194.4	16	84	27
1709	164.1	5	95	10
1734	358.7	5	95	10
1128/1207 [b]	4.369	12	88	22
1510/1593 [b]	0.581	21	79	35
Water content [b]	45.15	24	76	38

[a] $H^2_{mean} = \sigma^2_g/(\sigma^2_g + (\sigma^2_e/b))$, where b equals number of blocks, see also the "Materials and Methods" section

[b] Not multiplied with 10^{-5}

(Table 4, Fig. 2). At six regions, co-location of different QTLs were found. Especially at LG XIV, a cluster of QTLs representing lignin (1466, 1605 cm^{-1}), cellulose or hemicellulose (1360, 1429 cm^{-1}), and water (1666 cm^{-1}) were co-locating and in a region on LG VI-1 QTLs for G-lignin (1207 cm^{-1}), lignin (1510 cm^{-1}), and the S/G ratio (1128/1207 cm^{-1}) were positioned within the same region (Table 4, Fig. 2). For ten regions, all at different LGs, unique QTLs were positioned. For the biogas data, no QTL could be identified, most probably due to the limited number of clones that were phenotyped for biogas potential.

The allelic effects of the different QTLs are presented in Table 4. In most cases, the effects corresponded to a few percentage of the mean value of the trait. For 14 QTLs, the

Table 3 Pearson product-moment correlations based on mean values for each genotype between FT-IR wavenumbers and water content

Trait	*1128* S-lignin	*1169* Acetyl	*1207* G-lignin	*1229* G-lignin	*1254* Acetyl	*1271* G-lignin	*1360* Cell/hemi	*1429* Cell/hemi	*1466* Lignin	*1510* Lignin	*1593* Lignin	*1605* Lignin	*1666* Water	*1709* Carbonyl	*1734* Acetyl	*1128/1207* S/G ratio	*1510/1593* G/S ratio	*Water content %*
1128		**0.76**	0.42	0.47	0.49	0.34	0.16	0.07	0.04	-0.12	-0.12	-0.11	-0.01	-0.24	-0.15	0.42	-0.02	0.00
1169			0.27	0.40	**0.51**	0.42	0.25	0.21	0.13	-0.07	-0.26	-0.28	-0.11	-0.29	-0.12	0.36	0.15	-0.04
1207				**0.81**	**0.68**	**0.61**	0.28	0.20	0.29	0.42	0.06	-0.15	-0.09	0.17	0.18	**-0.64**	0.36	-0.08
1229					**0.86**	**0.69**	**0.52**	0.32	0.40	0.33	-0.13	-0.40	-0.29	-0.04	0.08	-0.41	0.43	-0.10
1254						**0.88**	**0.57**	0.41	0.38	0.22	-0.26	-0.45	-0.39	-0.15	0.04	-0.26	0.44	-0.06
1271							**0.58**	0.46	0.42	0.39	-0.22	-0.37	-0.42	-0.06	0.04	-0.32	**0.57**	-0.05
1360								**0.76**	**0.76**	0.37	-0.18	-0.47	**-0.63**	-0.02	0.14	-0.13	**0.52**	-0.05
1429									**0.92**	0.47	0.14	-0.17	-0.46	0.15	0.33	-0.13	0.35	-0.12
1466										**0.61**	0.18	-0.21	-0.45	0.24	0.40	-0.24	0.46	-0.09
1510											0.42	0.13	-0.05	0.34	**0.53**	**-0.51**	**0.64**	-0.06
1593												**0.87**	**0.51**	**0.50**	0.61	-0.16	-0.42	0.00
1605													**0.68**	0.33	0.37	0.06	**-0.60**	0.07
1666														0.31	0.26	0.08	-0.48	0.13
1709															**0.73**	-0.39	-0.09	-0.05
1734																-0.32	0.01	-0.08
1128/1207																	-0.36	0.08
1510/1593																		-0.06

Gray font indicates nonsignificant correlations, correlations where $0.01 < p < 0.05$ are in black, correlations where $0.001 < p < 0.01$ are in blue, and correlations at $p < 0.001$ are in red. Numbers in bold are correlations above 0.5 or below − 0.5

Table 4 QTLs identified for wood compounds measured by FT-IR, acetic acid, and for water content. Mean values for each genotype is used in the analysis

FT-IR wave number (cm^{-1})	Compound[a] for FT-IR analysis	Linkage group	Peak position (cM)	Marker closest to peak position	LOD[b]	LOD thresh GW[c]	PVE%	Diff. between maternal alleles[d] × 10^{-5}	Diff. between paternal alleles[d] × 10^{-5}	Interaction effect of alleles[d] × 10^{-5}
1128	S-lignin	IX	62.2	IX_7om_sa	5.3	4.3	5.3	2.1	2.4	3.1
1169	Acetyl	III	78.1	SB1048	4.7	4.4	4.4	0.08	1.7	0.02
1169	Acetyl	VII	95.3	VII_12om_sa	4.7	4.4	4.3	1.1	1.3	0.45
1169	Acetyl	IX	64.3	L16b.159	5.9	4.4	5.4	0.92	1.4	0.98
1207	G-lignin	VI-1	46.0	SB868	6.1	4.5	6.1	0.82	0.70	0.01
1229	G-lignin	III	130.1	III-17_sa	4.4	4.4	4.5	1.1	1.3	1.5
1254	Acetyl	III	129.3	III-17_sa	4.4	4.5	4.2	0.70	1.6	1.3
1271	G-lignin	Ib	261.0	SB226	4.5	4.4	4.4	1.7	0.54	0.02
1271	G-lignin	III	109.6	L31gr.177	5.4	4.4	5.4	1.3	1.8	0.68
1360	Cell/hemi/acetyl	IX	44.9	L4b.148	4.5	4.4	4.2	0.03	2.1	0.81
1360	Cell/hemi/acetyl	XIV_male	8.0	XIV-1_sa	4.9	4.4	4.5	0.45	2.0	1.6
1429	Cell/hemi	IX	67.4	IX_16_sa	4.9	4.4	4.5	0.12	2.1	1.5
1429	Cell/hemi	XIV_male	11.0	L8gr.140	7.1	4.4	6.5	1.6	2.7	1.9
1466	Lignin	XIV_male	5.5	XIV-1_sa	5.5	4.5	5.6	1.8	2.6	2.2
1510	Lignin	VI-1	46.0	SB868	7.0	4.5	6.6	1.8	2.4	0.34
1510	Lignin	XII	29.9	R_54_sa	4.9	4.5	4.5	0.22	2.6	1.4
1593[e]	Lignin	II	110.2	II_f70	4.9	4.4	5.0	3.2	0.52	2.0
1605	Lignin	XIV_male	9.0	XIV-1_sa	5.1	4.4	5.2	0.002	6.1	0.58
1666	Water	XIV_male	3.0	L14b.198f	6.2	4.4	5.7	2.0	6.2	2.3
1666	Water	XIV_male	84.3	XIV_18_sa_pIII	4.9	4.4	4.5	6.0	5.0	2.1
1709	Carbonyl/carboxyl	XVIII	1.3	XVIII_7_sa_pIII	5.7	4.4	5.7	0.66	3.7	0.95
1734	Acetyl	–	–	–	–	–	–	–	–	–
1128/1207[f]	S/G	VI-1	46.0	SB868	5.0	4.5	5.1	0.06	0.06	0.0009
1128/1207[f]	S/G	XV	5.0	B33.gr2	5.9	4.5	5.6	0.03	0.08	0.019
1510/1593[f]	G/S, lignin cross-linking	V_male	51.7	V_8om_sa_pI	6.8	4.4	6.9	0.002	0.01	0.003
Water content[f]	–	II	102.6	II-12_sa	4.7	4.3	4.3	0.10	0.91	0.16
Water content[f]	–	XVII	47.4	R_36_sa	6.3	4.3	6.1	1.13	0.41	0.11

[a] For description, see Table 1

[b] LOD = the logarithm of odds

[c] Genome-wide LOD threshold value

[d] For formulas for estimation of differential effects, see the "Materials and Methods" section

[e] Analysis only with Interval mapping

[f] Effects estimates should not be multiplied with 10^{-5}

difference between paternal alleles were considerably larger than the difference between the maternal alleles, e.g., for lignin (1605 cm^{-1}) at LG XIV_male (Table 4). Maternal effects were considerably larger than the paternal effect only for three QTLs (e.g., lignin QTL (1593 cm^{-1}) at LG II; Table 4). In one region (LG XIV), all co-locating QTLs had a large difference between paternal alleles compared to between the maternal alleles, indicating that the same locus/loci are involved in the genetic background of the traits (Table 4). For other clustered QTLs (LG II, LG III, LG VI-1, LGIX), the individual QTL did not show the same pattern in allelic effects, possibly indicating that the individual QTL in the same cluster had different genetic background (Fig. 2, Table 4).

To validate the approach, we used pyrolysis-GC-MS for the samples representing the contrasting genotypes corresponding to QTL for FT-IR signal at 1207 cm^{-1} (assigned to G-lignin) and 1605 cm^{-1} (assigned to lignin). The genotypes showed expected trends in the pyrolysis signals corresponding to G-lignin and to lignin (Online Resource 3), validating the proposed assignments of these FT-IR signals. It is however possible that the FT-IR signals also reflect changes in cell wall architecture among the genotypes, which are difficult to define and which are not simply correlated with the content of the assigned compounds.

Candidate Genes in QTL Regions

For each QTL region marked in black in Fig. 2, the corresponding genes lists in *Populus* were prepared and analyzed for the presence of cell wall-related genes. Approximately 200 such genes were identified (Online Resource 4). The most interesting finding was the presence of clusters of genes functioning in similar biochemical processes at some loci. For example, clusters of genes responsible for homogalacturonan degradation were found in the vicinities of two QTLs related to lignin, on LG II, and on LG VI-1. There was also a cluster

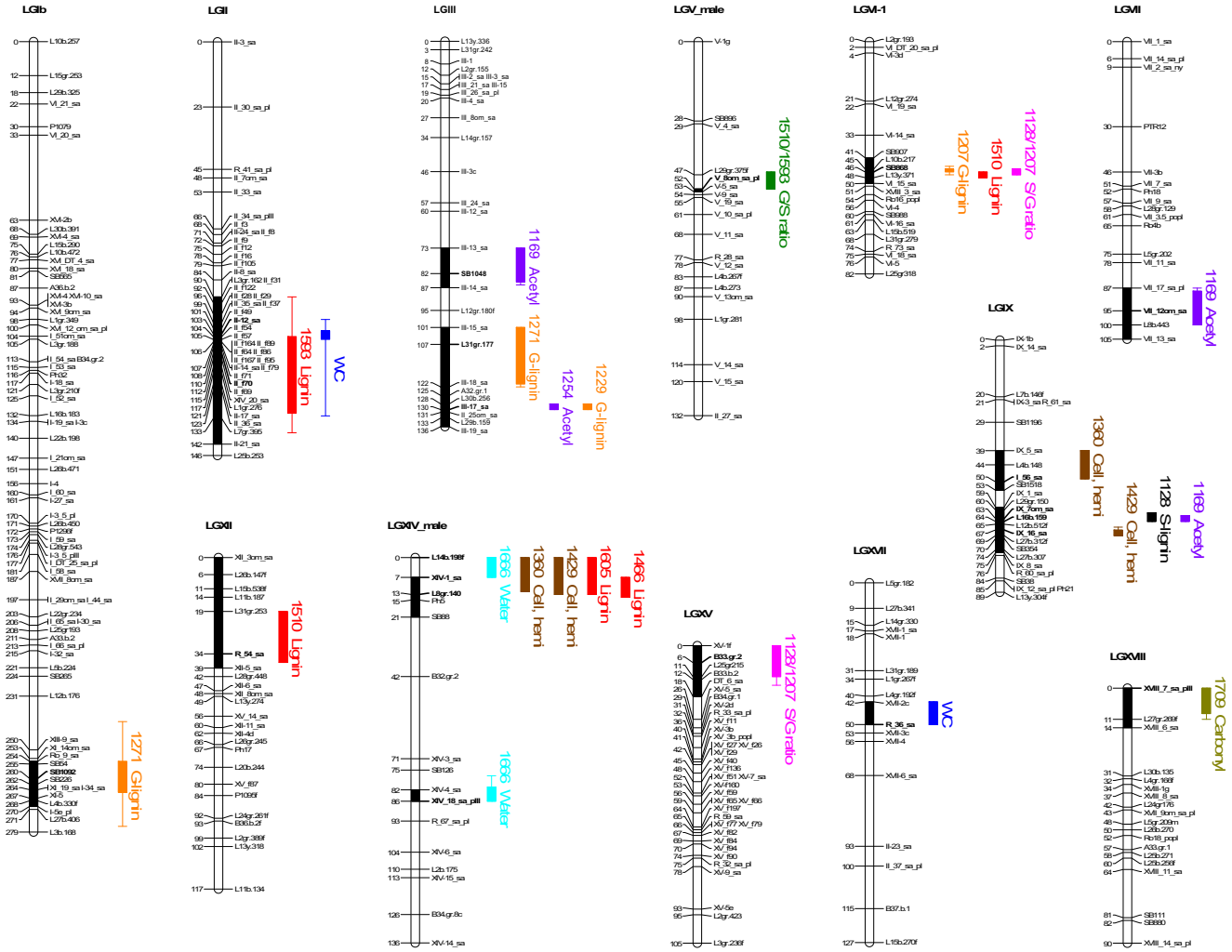

Fig. 2 QTLs for selected FT-IR wavenumbers, water content (WC) and acetic acid content for a F_1 population between *Salix viminalis* and *S. viminalis* × *S. schwerinii*. Each QTL is indicated to the right of the linkage group with a bar for 1-LOD and a line for 2-LOD confidence intervals, respectively, if no line the 1 and 2-LOD intervals overlap. Markers closest to the peak position of the QTL are in bold. Black parts of the linkage groups represent the corresponding regions in the *Populus trichocarpa* genome from where gene lists have been subtracted. For FT-IR wavenumber description, see Table 1. Linkage group names and markers as in [30] but markers have been removed where dense positions for a clear figure. For abbreviations, see Fig. 1

of xylanase genes associated with QTL for S-lignin coinciding with a QTL for cellulose or hemicellulose on LG IX (Fig. 2, Online Resource 4), and a group of several GT2 processive glycosyl transferases along with two cellulases in the vicinity of the QTL for acetyl on LG III.

Cellulose biosynthetic genes were found in several analyzed loci. The secondary wall—CesA7A—was associated with the QTL for G-lignin and lignin on LG VI-1, and primary wall CesA genes were associated with QTLs for S-lignin, acetyl, and cellulose or hemicellulose on LG IX, and G-lignin on LG I-b (Online Resource 4).

Several lignin biosynthetic genes were also identified within the examined regions. Two genes encoding cinnamyl-alcohol dehydrogenase (CAD) enzyme were found within QTLs for G-lignin on LG Ib and G-lignin/lignin on LG IV-1 (Online Resource 4). A gene encoding 4-coumarate-CoA ligase (4CL) was located within a QTL for G-lignin on LG III. A cluster of cinnamoyl-CoA reductase (CCR) encoding genes was found to be associated with cellulose or hemicellulose QTL on LG IX.

Biogas Potential and Correlation with FT-IR

The different *Salix* clones showed large variation in final biochemical methane potential and values varied between 43.1 (SD 9.9) and 242.2 (SD 21.6) N mL CH_4/g TS, with a mean value of 149.7 (SD 45.2) (Online Resource 5, Fig. 3). The cellulose control reached 327.5 (SD 9.9), an expected value for a well-functioning inoculum. The degradation rates were estimated by determining the time needed to reach 50 and 80% of the final potential and here the values varied between 7–20 and 20–30 days, respectively (Online Resource 5, Fig. 3). The time to reach the full potential ranged from 39

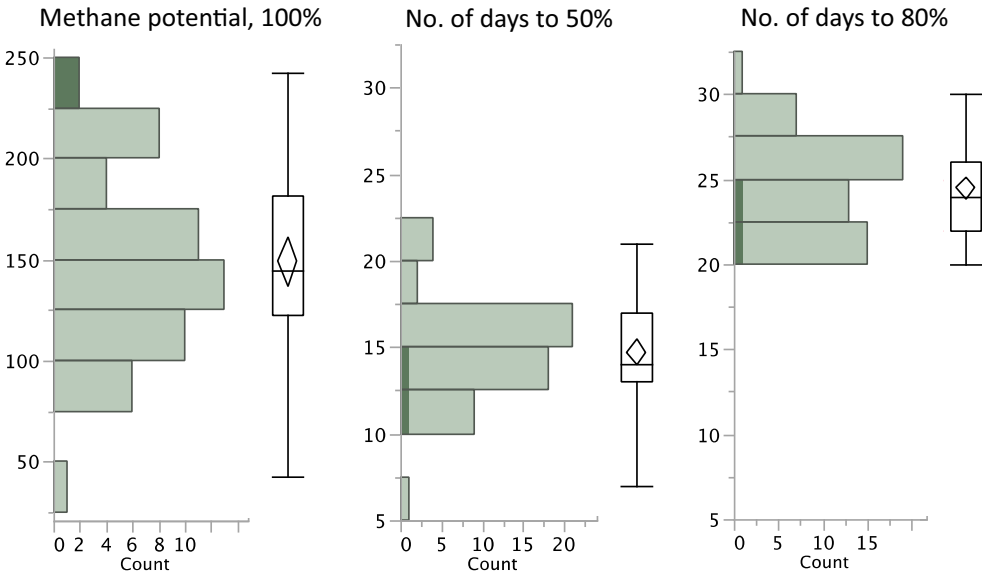

Fig. 3 Phenotypic distributions of final methane potential (methane potential, 100%), number of days to reach 50% of the final methane potential (no. of days to 50%), and number of days to reach 80% of the final methane potential (no. of days to 80%) for a subset of clones in the mapping population and the parents. The clones with the highest final methane potential are shaded in dark gray

to 47 days. The distribution of the BMP values, i.e., 100% of the final value, and the days to reach 50 and 80% of the potential, illustrated that the clones showing highest (> 225 ml CH_4/g TS) or lowest (< 100 ml CH_4/g) yields did not have the highest and lowest production rate, respectively (Fig. 3). Still, clones with high yields had rate similar or higher than the average. Results from the FT-IR analysis were correlated to the BMP values, both the final values and the time needed to reach 50, 80, and 100% of the final BMP value. The results showed the same pattern but for the 50% values, the significance was slightly better and thus this result is presented in Fig. 4. A positive correlation ($P \leq 1\%$) was seen between the methane production and FT-IR wavenumbers 1220–1265, 1165, and 905 in centimeter^{-1}. These wavenumbers include signals from acetyl (1254, 1165), G-lignin (1229), and cellulose (905). A negative correlation was seen for wavenumbers 1545, 1530, and 1190–1200 cm^{-1}, most of them including signals corresponding to lignin (1510, 1207). For the days to reach 50% of the BMP values, a weak negative correlation ($P \leq 10\%$) was seen for signals corresponding to hemicellulose and cellulose (1300–1375 cm^{-1}), i.e., a higher content of hemicellulose and cellulose resulted in a faster methane production.

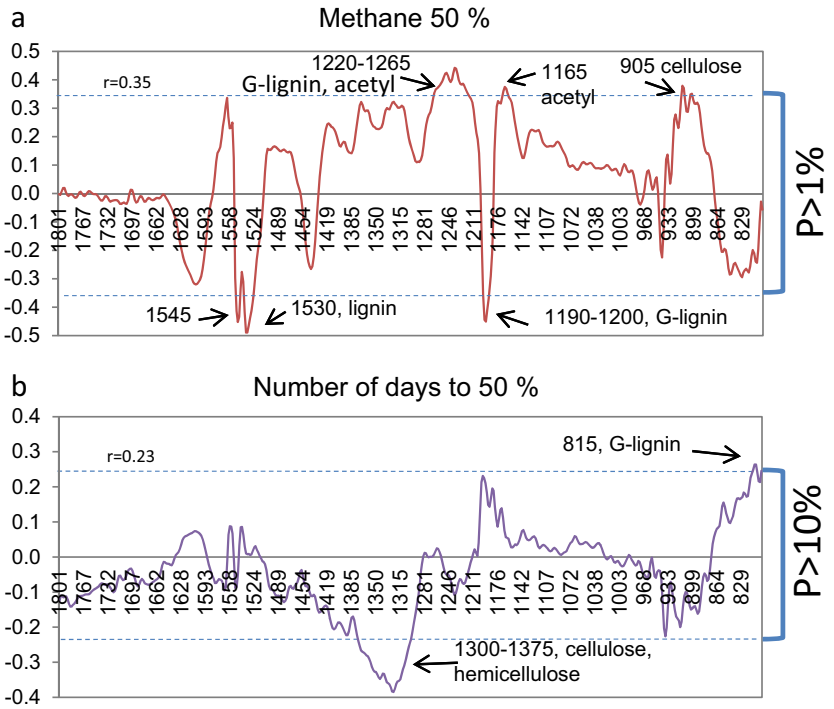

Fig. 4 Correlations between the FT-IR signals and **a** methane production at 50% of final production level, and **b** number of days to reach the 50% methane production level, in analyzed *Salix* clones. Coefficients of correlations (*Y*-axis) for different wave numbers in centimeter^{-1} (*X*-axis) are shown. Data are based on average data obtained for clones selected for the biogas measurements, $N = 55$. Broken lines indicate boundaries for the significance levels shown. The peaks crossing the broken lines have *P* values < than indicated and are considered significant. Suggested assignment of the significant wavenumbers discussed in the text is displayed

Discussion

In this study, we have identified QTLs for different FT-IR signals representing wood traits important for different applications of woody biomass as, e.g., production of biogas. Knowledge of the genetic variation and the genetic background of the phenotypic variation of traits is a prerequisite for understanding the possibilities to direct breeding towards specific goals connected to wood characteristics. FT-IR is a sensitive method to detect chemical composition of cell walls, although the interpretation of the spectra is usually very complex. We have applied an assignment of the peaks that are typically used for cell wall studies in different lignocellulosic materials, derived mostly from woody species [36]. Expected correlations between signals from different wavenumbers assigned to the same or similar compounds, for example positive correlations for G-lignin and lignin assigned signals (1207, 1229, 1271 cm^{-1}) and their negative correlation with S/G ratio assigned signals (1128/1207 cm^{-1}), or a high correlation between different signals for acetyl (1169 and 1254 cm^{-1}; Table 3) provide confidence for the proposed individual assignments. An independent verification of the wood composition of the contrasting genotypes from QTL 1207 and 1605 cm^{-1} by pyrolysis-GC-MS showed the expected change in lignin contents in these genotypes (Online Resource 3). However, low correlations between signals assigned to the same compound were also recorded, most likely reflecting their different molecular form, context, or contributions from other compounds. The FT-IR spectra should thus be treated as chemical "fingerprints" of samples rather than specific indicators of concentrations of different compounds.

We have found broad sense heritabilities for the different traits of low (3–20%) to moderate (21–43%) size that indicate that we have genetic variation that could be utilized in breeding. Comparing these values with clonal studies of *Salix* and *Populus* is difficult since these estimates are based on the population and the environment in which they are measured, but earlier clonal studies of *Salix* have identified variation between clones for wood traits as lignin, cellulose, and hemicellulose [24, 25], and for enzymatic saccharification [24] even though no heritabilities have been estimated. An early study of wood traits in *Salix* estimated narrow sense heritabilities to 17–28% for wood density and to 46% for dry matter content [23] also showing possibilities for a directional breeding for these wood traits. The highly significant and positive correlations between FT-IR signals assigned to acetyl and S- or G-lignin, and between cellulose/hemicellulose/acetyl and lignin observed in this study (Table 3) are expected since they reflect the proportion of different cell wall layers in the willow stem biomass. The chemistry of this type of biomass was shown to be primarily affected by the content of tension wood in woody stems [55, 56]. Low-tension wood individuals have high lignin and acetylated xylan contents, which are characteristic for the S-wall layers, while individuals with high proportion of tension wood have low lignin and acetylated xylan contents and high crystalline cellulose and galactan contents typically found in the G-layers [57].

QTLs for Different Wood Traits

We identified 27 QTLs representing different wood chemical traits on 12 of the *Salix* LGs. Four QTLs of different wavenumbers were identified for each of lignin, G-lignin, and acetyl. No other study in *Salix* has identified QTLs for wood components. In the study by Brereton et al. [26] where they also used a population with a *S. viminalis* × *S. schwerinii* background, they identified four QTLs for enzymatic saccharification on the LGs V, X, XI, and XVI and a collocating QTL on LG V for glucose. We also found one QTL on LG V for the S/G ratio but if this is in a similar genomic region is not possible to judge since there are no common markers between the two linkage maps. In several *Populus* studies, there have been reported QTLs for different wood traits as lignin, cellulose, and hemicellulose and S/G ratio [13–15, 17]. In the study by Muchero et al. [17], several QTLs were found on linkage group XIV and several of the traits had QTL peaks around the position of 7 Mbp. The QTL cluster that we identified on LG XIV was between 1.5 and 2.1 Mbp on the corresponding region of the *Populus* chromosome and thus not overlapping with the collocated QTLs found by Muchero et al. [17]. On the other hand, Ranjan et al. [14] found QTLs for root lignin and stem S/G content in the region around 2 Mbp on chromosome XIV.

The higher number of QTLs with larger difference between paternal alleles compared to maternal alleles is probably an effect of the hybrid origin of the male parent and an allelic difference between the two species. This is also what could be expected and seen earlier for other traits for the same population [58].

The markers closest to the peak position of the different QTLs in the cluster at LG XIV-male all have high differential paternal effects for cellulose, hemicellulose, and for lignin which could indicate a pleiotrophic effect of a locus influencing all the traits. Also, for the collocating QTLs at LGs VI-1 and IX, the individual QTLs in the clusters have a similar pattern regarding differential effects of maternal and paternal alleles, indicating a similar genetic background of the traits in the cluster. Interesting candidate genes related to cell wall properties were also identified in the clusters of QTLs.

Candidate Genes in QTL Regions

Although as many as 200 candidate genes were identified within the QTL regions, based on gene annotation, we note that most genes remained unannotated or without any known

biological function so this analysis is not complete. Any of these genes could be responsible for the detected variability. Therefore, additional studies are required to draw any conclusions about the genes responsible for the observed variation. It is also intriguing if the presence of gene clusters in the identified regions has any functional meaning for the observed effects, for example by affecting cross-over and thus gene segregation.

Biogas Production Potential

The biogas potential of willow has been reported around 50–310 ml CH_4/g VS, depending on pre-treatment and variety [7], values in line with results from the present study. For an economically feasible biogas production, values around 200 ml CH_4/biomass have been suggested, which were observed for many analyzed progenies in the current analyses, altogether illustrating that willow in this regard represent a possible substrate for biogas production [59].

A high conversion rate is essential for the process as a material with slow degradation rate will not be fully converted, and the BMP will not be reached, within the hydraulic retention time of the biogas process, typically ca 20–40 days. In this study, the clones had different degradation rates but still all reached their BMP values within 37–47 days. A positive correlation was seen for the degradation rate and the content of cellulose and hemicellulose while a negative correlation was seen between lignin content and the final BMP value. These results were in line with a recent study showing a positive correlation between BMP and the sum of non-lignin fractions versus lignin for various lignocellulosic material, including forest residues [60].

Anaerobic digestion to biogas, in contrast to ethanol production, proceeds through a series of reactions involving different microbial groups, including microorganism having hydrolytic activity [61]. Thus, lignocellulosic materials such as willow can be degraded without a preceding sacharification, as required for ethanol production. Still, to achieve rates and yields sufficiently high for an industrial process, a pre-treatment is required to break up lignocelluose, otherwise difficult to reach for microbial degradation [62]. Pre-treatment investigated for willow to be used for biogas include both mechanical, thermal, and chemical pre-treatments [7]. In this study, the Salix clones were milled to a rough powder before evaluated, to be comparable to a mechanical pre-treatment.

For optimal utilization of willow in a biogas system, co-digestion with complementary material would be required [63]. Willow is characterized by a high C/N quota as well as low levels of buffering components and possibly also trace metals, essential for microbial activity, and thus without complementary material the biogas process would fail due to nutrient deficiency [61]. Still, willow has a high energy content per mass and can thus be used as a complement to less carbon-rich and more diluted material, such as manure [63, 64]. Addition of willow to a manure-based biogas system would allow a higher organic load with only marginal effects on retention time and, consequently, result in increased volumetric gas yields and a more efficient use of digester volume. All are of critical importance for the economy of a farm-based biogas plant [65]. Moreover, addition of willow to the biogas process would generate a digestion residue with a high proportion of recalcitrant carbon not degraded in the process. When using this residue as a fertilizer, a positive effect on climate change mitigation can be obtained as stable carbon is returned to the soil [66].

Breeding for Different Wood Applications

This study demonstrates genetic variation for different wood components and also identify many genome-wide significant QTL regions for the traits. This is an important step towards understanding the genetic background in *Salix* wood traits but further studies are required to verify candidate genes and possibly causative SNPs in the genes. In *Populus*, such attempts have been conducted using a combination of QTL and association mapping approaches [17] or by using candidate genes or genome-wide marker approaches in association mapping [16, 20, 21]. Breeding towards different application of *Salix* wood as, e.g., biofuels seems as a realistic alternative since the phenotypic variation in most of the traits has a genetic component. Even though there were too few clones analyzed for biogas potential to identify QTLs, we found a lot of variation between the different clones in the biogas traits as well as a negative and positive correlation to lignin and acetyl signals, respectively, which indirectly indicate that selection for high acetyl or for a lower content of lignin would be positive for biogas production. For example, the QTL-markers for FT-IR signals 1254 and 1207 assigned to acetyl and lignin could possibly be used in selection for higher biogas production due to the positive and negative correlation, respectively. An implementation of molecular techniques in the selection and breeding of these traits would be of high value since wood traits as well as biogas potential are difficult, costly, and time consuming to phenotype.

Acknowledgements We thank Marie Melander for help with collection of plant material from the field, Tao Yang for sample preparation and running FT-IR measurements, and Maria Eriksson for analysis of biomethane potential. We acknowledge the assistance of Dr. András Gorzsás and the Vibrational Spectroscopy Platform at the Chemical-Biological Centre (KBC) with the FT-IR spectroscopic analysis, and the UPSC Cell Wall Lab funded by Bio4Energy and TC4F for the assistance with sample preparation. PM-AP's studentship was funded by the FORMAS program BioImprove. The work was also supported by the STandUP for energy program.

References

1. Ingvarsson PK (2005) Nucleotide polymorphism and linkage disequilibrium within and among natural populations of European aspen (*Populus tremula* L., Salicaceae). Genetics 169:945–953
2. Berlin S, Fogelqvist J, Lascoux M, Lagercrantz U, Rönnberg-Wästljung A-C (2011) Polymorphism and divergence in two willow species, *Salix viminalis* L. and *Salix schwerinii* E. Wolf. G3 Genes Genomes Genetics 1:387–400
3. Karp A, Hanley SJ, Trybush SO, Macalpine W, Pei M, Shield I (2011) Genetic improvement of willow for bioenergy and biofuels. J Integr Plant Biol 53:151–165. https://doi.org/10.1111/j.1744-7909.2010.01015
4. Karp A, Shield I (2008) Bioenergy from plants and the sustainable yield challenge. New Phytol 179:15–32
5. Smart LB, Cameron KD (2008) Genetic improvement of willow (*Salix* spp.) as a dedicated bioenergy crop. In: Vermerris W (ed) Genetic improvement of bioenergy crops. Springer, New York, pp 377–396
6. Phitsuwan P, Sakka K, Ratanakhanokchia K (2013) Improvement of lignocellulosic biomass in planta: a review of feed stocks, biomass recalcitrance, and strategic manipulation and ideal plants designed for ethanol production and processabilty. Biomass Bionenergy 58:390–405
7. Matsakas L, Rova U, Christakopulus P (2016) Strategies for enhanced biogas generation through anaerobic biogas generation of forest material—an overview. Bioresources 11(2):5482–5499
8. Pu Y, Hu F, Huang F, Davison BH, Ragauskas AJ (2013) Assessing the molecular structure basis for biomass recalcitrance during dilute acid and hydrothermal pretreatments. Biotechnol Biofuels 6:15
9. Chen H, Liu J, Chang X, Chen D, Xue Y, Liu P, Lin H, Han S (2017) A review on the pretreatment of lignocellulose for high-value chemicals. Fuel Process Technol 160:196–206
10. Davison BH, Drescher SR, Tuskan GA, Davis MF, Nghiem NP (2006) Variation of S/G ratio and lignin content in a *Populus* family influences the release of xylose by dilute acid hydrolysis. Appl Biochem Biotechnol 130:427–435
11. da Costa RMF, Lee SJ, Allison GG, Hazen SP, Winters A, Bosch M (2014) Genotype, development and tissue-derived variation of cell-wall properties in the lignocellulosic energy crop *Miscanthus*. Ann Bot 114:1265–1277
12. Zhang D, Zhang Z, Yang K (2006) QTL analysis of growth and wood chemical content traits in an interspecific backcross family of white poplar (*Populus tomentosa* × *P. bolleana*) × *P. tomentosa*. Can J Forest Res 36:2015–2023
13. Novaes E, Osorio L, Drost DR, Miles BL, Boaventura-Novaes CRD, Benedict C, Dervinis C, Yu Q, Sykes R, Davis M, Martin TA, Peter GF, Kirst M (2009) Quantitative genetic analysis of biomass and wood chemistry of *Populus* under different nitrogen levels. New Phytol 182:878–890. https://doi.org/10.1111/j.1469-8137.2009.02785.x
14. Ranjan P, Yin T, Zhang X, Kalluri UC, Yang X, Jawdy S (2009) Bioinformatics-based identification of candidate genes from QTLs associated with cell wall traits in *Populus*. Bioenergy Res 3:172–182
15. Yin T, Zhang X, Gunter L, Priya R, Sykes R, Davis M, Wullschleger SD, Tuskan GA (2010) Differential detection of genetic loci underlying stem and root lignin content in *Populus*. PLoS One 5:e14021–e14027. https://doi.org/10.1371/journal.pone.0014021
16. Porth I, Klápště J, Skyba O, Lai BSK, Geraldes A, Muchero W, Tuskan GA, Douglas CJ, El-Kassaby YA, Mansfield SD (2013) *Populus trichocarpa* cell wall chemistry and ultrastructure trait variation, genetic control and genetic correlations. New Phytol 197:777–790. https://doi.org/10.1111/nph.12014
17. Muchero W, Guo J, DiFazio SP, Chen J-G, Ranjan P, Slavov GT, Gunter LE, Jawdy S, Bryan AC, Sykes R, Ziebel A, Klápste J, Porth I, Skyba O, Unda F, El-Kassaby YA, Douglas C, Mansfield SD, Martin J, Schackwitz W, Evans LM, Czarnecki O, Tuskan GA (2015) High-resolution genetic mapping of allelic variants associated with cell wall chemistry in *Populus*. BMC Genomics 16:24
18. Farkenkrog AM, Neves LG, Resende MFR Jr, Vazquez AI, de los Campos G, Dervinis C, Sykes R, Davis M, Davenport R, Barbazuk WB, Kirst M (2016) Genome-wide association study reveals putative regulators of bioenergy traits in *Populus deltoides*. New Phytol 213:799–811. https://doi.org/10.1111/nph.14154
19. 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. Proc Natl Acad Sci U S A 108:6300–6305. https://doi.org/10.1073/pnas.1009252108
20. Wegrzyn JL, Eckert AJ, Choi M, Lee JM, Stanton BJ, Sykes R, Davis MF, Chung-Jui T, Neale DB (2010) Association genetics of traits controlling lignin and cellulose biosynthesis in black cottonwood (*Populus trichocarpa*, Salicaceae) secondary xylem. New Phytol 188:515–532
21. Guerra FP, Wegrzyn JL, Sykes R, Davis MF, Stanton BJ, Neale DB (2013) Association genetics of chemical wood properties in black poplar (*Populus nigra*). New Phytol 197:162–176. https://doi.org/10.1111/nph.12003
22. Sennerby-Forsse L (1985) Clonal variation of wood specific gravity, moisture content, and stem bark percentage in 1-year-old shoots of 20 fast-growing *Salix*. Can J For Res 15:531–534
23. Rönnberg-Wästljung A-C, Gullberg U (1999) Genetics of breeding characters with possible effects on biomass productionin *Salix viminalis* (L.) Theor Appl Genet 98:531–540
24. Ray MJ, Brereton NJB, Shield I, Karp A, Murphy RJ (2012) Variation in cell wall composition and accessibility in relation to biofuel potential of short rotation coppice willows. Bioenergy Res 5:685–698
25. Serapiglia MJ, Humiston MC, Xu H, Hogsett DA, de Orduña RM, Stipanovic AJ, Smart LB (2013) Enzymatic saccharification of shrub willow genotypes with differing biomass composition for biofuel production. Front Plant Sci 4:57. https://doi.org/10.3389/fpls.2013.00057
26. Brereton NJ, Pitre FE, Hanley SJ, Ray MJ, Karp A (2010) QTL mapping of enzymatic saccharification in short rotation coppice willow and its independence from biomass yield. Bioenergy Res 3:251–261
27. Francia E, Tacconi G, Crosatti C, Barabaschi D, Bulgarelli D, Dall'Aglio E, Vale` G (2005) Marker assisted selection in crop plants. Plant Cell Tissue Organ Cult 82:317–342
28. Grattapaglia D, Plomion C, Kirst M, Sederoff RR (2009) Genomics of growth traits in forest trees. Curr Opin Plant Biol 12:148–156
29. Hanley SJ, Mallot MD, Karp A (2006) Alignment of a *Salix* linkage map to the *Populus* genomic sequence reveals macrosynteny between willow and poplar genomes. Tree Genet Genomes 3:35–34
30. Berlin S, Lagercrantz U, von Arnold S, Öst T, Rönnberg-Wästljung A-C (2010) High-density linkage mapping and evolution of paralogs and orthologs in *Salix* and *Populus*. BMC Genomics 11:129

31. Tsarouhas V, Gullberg U, Lagercrantz U (2003) Mapping of quantitative trait loci controlling timing of bud flush in *Salix*. Hereditas 138:172–178
32. Berlin S, Ghelardini L, Bonosi L, Weih M, Rönnberg-Wästljung A-C (2014) QTL mapping of biomass and nitrogen economy traits in willows (*Salix* spp.) grown under contrasting water and nutrient conditions. Mol Breed 34:1987–2003
33. Danfors B, Ledin S, Rosenqvist H (1997) Short-rotation willow coppice. Growers' Manual. Uppsala, Sweden: Swedish Institute of Agricultural and Environmental Engineering
34. Ghelardini L, Berlin S, Weih M, Lagercrantz U, Gyllenstrand N, Rönnberg-Wästljung A-C (2014) Genetic architecture of spring and autumn phenology in *Salix*. BMC Plant Biol 14:31
35. Telenius BF (1997) Implications of vertical distribution and within-stand variation in moisture content for biomass estimation of some willow and hybrid poplar clones. Scand J For Res 12:336–339
36. Gorzsas A, Stenlund H, Persson P, Trygg J, Sundberg B (2011) Cell-specific chemotyping and multivariate imaging by combined FT-IR microspectroscopy and orthogonal projections to latent structures (OPLS) analysis reveals the chemical landscape of secondary xylem. Plant J 66:903–914
37. Trygg J, Wold S (2003) O2-PLS, a two-block (X-Y) latent variable regression (LVR) method with an integral OSC filter. J Chemom 17(1):53–64
38. Gerber L, Eliasson M, Trygg J, Moritz T, Sundberg B (2012) Multivariate curve resolution provides a high-throughput data processing pipeline for pyrolysis gas chromatography/mass spectrometry. J Anal Appl Pyrolysis 95:95–100
39. APHA A (1998) Standard methods for water and wastewater examination. American Public Health Association 20th Edition, Washington DC
40. Schnürer A, Bonn I, Moesteth J (2016) Protocol for start-up and operation of CSTR biogas processes. Hydrocarbon and lipid microbiology protocols, Springer Protocol Handbooks (Spinger Berlin). doi https://doi.org/10.1007/8623_2016_214, Protocol for Start-Up and Operation of CSTR Biogas Processes
41. Sun RC, Tomkinson J, Sun XF, Wang NJ (2000) Fractional isolation and physico-chemical characterization of alkali-soluble lignins from fast-growing poplar wood. Polymer 41(23):8409–8417
42. Sun RC, Fang JM, Tomkinson J, Geng ZC, Liu JC (2001) Fractional isolation, physico-chemical characterization and homogeneous esterification of hemicelluloses from fast-growing poplar wood. Carbohydr Polym 44(1):29–39
43. Xiao B, Sun XF, Sun RC (2001) The chemical modification of lignins with succinic anhydride in aqueous systems. Polym Degrad Stab 71(2):223–231
44. Boeriu CG, Bravo D, Gosselink RJA, van Dam JEG (2004) Characterisation of structure-dependent functional properties of lignin with infrared spectroscopy. Ind Crop Prod 20:205–218
45. Labbé N, Rials TG, Kelley SS, Cheng ZM, Kim JY, Li Y (2005) FT-IR imaging and pyrolysis-molecular beam mass spectrometry: new tools to investigate wood tissues. Wood Sci Technol 39(1):61–U19
46. Naumann A, Navarro-Gonzalez M, Peddireddi S, Kues U, Polle A (2005) Fourier transform infrared microscopy and imaging: detection of fungi in wood. Fungal Genet Biol 42(10):829–835
47. Faix O (1991) Classification of lignins from different botanical origins by FT-IR spectroscopy. Holzforschung 45:21–27
48. Faix O, Bremer J, Schmidt O, Stevanovic T (1991) Monitoring of chemical-changes in white-rot degraded beech wood by pyrolysis-gas chromatography and fourier-transform infrared-spectroscopy. J Anal Appl Pyrolysis 21(1–2):147–162
49. Zhong RQ, Morrison WH, Himmelsbach DS, Poole FL, Ye ZH (2000) Essential role of caffeoyl coenzyme a O-methyltransferase in lignin biosynthesis in woody poplar plants. Plant Physiol 124(2):563–577
50. Gou JY, Park S, Yu XH, Miller LM, Liu CJ (2008) Compositional characterization and imaging of "wall-bound" acylesters of *Populus trichocarpa* reveal differential accumulation of acyl molecules in normal and reactive woods. Planta 229:15–24
51. SAS Institute Inc (2012) Using JMP 10. SAS Institute Inc., Cary
52. Van Ooijen JW (2009) MapQTL 6, software for the mapping of quantitative trait loci in experimental populations, Wageningen Netherlands
53. Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121: 185–199
54. Geisler-Lee J, Geisler M, Coutinho PM, Segerman B, Nishikubo N, Takahashi J, Aspeborg H, Djerbi S, Master E, Andersson-Gunneråas S, Sundberg B, Karpinski S, Teeri TT, Kleczkowski LA, Henrissat B, Mellerowicz EJ (2006) Poplar carbohydrate-active enzymes (CAZymes). Gene identification and expression analyses. Plant Physiol 140:946–962
55. Brereton NJ, Pitre FE, Ray MJ, Karp A, Murphy J (2011) Investigation of tension wood formation and 2,6-dichlorbenzonitrile application in short rotation coppice willow composition and enzymatic saccharification. Biotechnol Biofuels 4:13
56. Brereton NJ, Ray MJ, Shield I, Martin P, Karp A, Murphy RJ (2012) Reaction wood—a key cause of variation in cell wall recalcitrance in willow. Biotechnol Biofuels 5(1):83. https://doi.org/10.1186/1754-6834-5-83
57. Mellerowicz EJ, Gorshkova T (2012) Tensional stress generation in gelatinous fibres: a review and possible mechanism based on cell wall structure and composition. J Exp Bot 63:551–565
58. Berlin S, Hallingbäck HR, Beyer F, Nordh N-E, Weih M, Rönnberg-Wästljung A-C (2017) Genetics of phenotypic plasticity and biomass traits in hybrid willows across contrasting environments and years. Ann Bot 120(1):87–100
59. Angelidaki I, Ellegaard L (2003) Codigestion of manure and organic wastes in centralized biogas plants. Appl Biochem Biotechnol 109:95–105
60. Liu X, Bayard R, Benbelkacem H, Buffiere P, Gourdon R (2015) Evaluation of the correlations between biodegradability of lignocellulosic feedstocks in anaerobic digestion process and their biochemical characteristics. Biomass Bioenergy 81:534e543
61. Schnürer A (2016) Biogas production: microbiology and technology. Adv.Biochem.Eng Biotechnol. Doi:https://doi.org/10.1007/10_2016_5
62. Sawatdeenarunat C, Surendra KC, Takara D, Oechsner H, Khanal SK (2015) Anaerobic digestion of lignocellulosic biomass: challenges and opportunities. Bioresour Technol 178:178–186
63. Estevez MM, Zengin Z, Linjordet R, Schnürer A, Morken J (2014) Semi-continuous anaerobic co-digestion of cow manure and steam-exploded Salix with recirculation of liquid digestate. J Environ Manag 136:9–15
64. Neshat A, Mohammadi M, Najafpour GD, Lahijani P (2017) Anaerobic co-digestion of animal manures and lignocellulosic residues as a potent approach for sustainable biogas production. Renew Sust Energ Rev 79:308–322
65. Ahlberg-Eliasson K, Nadeau E, Levén L, Schnürer A (2017) Production efficeiency of Swedish farm-scale biogas plants. Biomass Bioenergy 97:27–37
66. Ericsson N, Norberg Å, Sundberg C, Ahlgren S, Hansson P-A (2014) Climate impact and energy efficiency from electricity generation through anaerobic digestion or combustion of short rotation coppice willow. Appl Energy 132:86–98