

# Ploidy Level Affects Important Biomass Traits of Novel Shrub Willow (*Salix*) Hybrids

Michelle J. Serapiglia · Fred E. Gouker · J. Foster Hart ·  
Faride Unda · Shawn D. Mansfield ·  
Arthur J. Stipanovic · Lawrence B. Smart

Published online: 30 August 2014  
© Springer Science+Business Media New York 2014

**Abstract** Polyploidy is a common observation in the genus *Salix*, including some of the shrub willow species currently being bred as a potential bioenergy feedstock. Breeding of shrub willow has produced new species hybrids, among which a disproportionate number of high-yielding genotypes are triploid, produced from crosses between diploid and tetraploid parents. These novel hybrids display significant variation in biomass compositional quality, including differences according to ploidy. The triploid and tetraploid genotypes possess lower lignin content than diploid genotypes. Biomass composition was also significantly different across the 3-year growth cycle typical of bioenergy plantings. There were differences in syringyl/guaiacyl (S:G) lignin ratios among the 75 genotypes examined, in addition to significant correlations with willow growth traits, yield, and composition. These differences suggest that a long-term strategy of breeding for

triploid progeny will generate cultivars with improved growth traits and wood composition for conversion to biofuels.

**Keywords** Biofuel feedstock · Cellulose · Lignin · Polyploid · Wood composition

## Introduction

The *Salix* L. genus is a member of the Salicaceae family (also containing *Populus* L.) consisting of over 350 species ranging from dwarf, alpine willows to shrubs reaching over 10 m, to trees greater than 20 m in height [1]. Many of the shrub willow species are used for woody biomass production for renewable energy in Europe and the USA. Through controlled pollination and hybridization, many novel cultivars are being bred, selected, and planted commercially [2–6]. The shrub species of willow are ideal for bioenergy production, because they thrive in short-rotation coppicing systems and can maintain vigorous growth through many harvest cycles.

Willow breeding has traditionally exploited the high genetic diversity and heterozygosity inherent among natural populations [7]. The main breeding strategies for willow in the USA, UK, and Sweden have been to improve biomass yield by capturing heterosis through the production of interspecific hybrids [8, 9, 4]; however, breeding for pest and disease resistance has also been of high priority due to detrimental effects on growth and production [10–12]. It is well documented that heterosis is a common phenomenon observed among hybrid progeny in many plant species and also among polyploid individuals [13–16]. Recently, analysis of interspecific hybrid progeny of shrub willow found that triploids were disproportionately represented among genotypes with high yield and had better overall performance than diploid or tetraploid progeny [17]. Similar advances in yield and growth

**Electronic supplementary material** The online version of this article (doi:10.1007/s12155-014-9521-x) contains supplementary material, which is available to authorized users.

M. J. Serapiglia · F. E. Gouker · L. B. Smart (✉)  
Department of Horticulture, New York State Agricultural Experiment  
Station, Cornell University, 630 West North Street, Geneva,  
NY 14456, USA  
e-mail: lbs33@cornell.edu

J. F. Hart · F. Unda · S. D. Mansfield  
Department of Wood Science, Forest Sciences Centre, University of  
British Columbia, Vancouver, BC V6T 1Z4, Canada

A. J. Stipanovic  
Department of Chemistry, State University of New York College of  
Environmental Science and Forestry, Syracuse, NY 13210, USA

### Present Address:

M. J. Serapiglia  
Sustainable Biofuels and Coproducts Research Unit, Eastern  
Regional Research Center, USDA—ARS, 600 East Mermaid Lane,  
Wyndmoor, PA 19038, USA

in triploids compared to diploids have also been observed in poplar [18, 19].

While biomass yield is an important trait for bioenergy crop production, biomass quality traits need also to be considered. Wood quality has long been of high importance in the pulp and paper industry, and now, as interest in the production of bioproducts and biofuels increases, more research is focused on the genetic regulation of biomass quality traits. Woody traits that are important for industrial applications include polysaccharide and lignin content, specific gravity, fiber traits (microfibril angle, fiber length), and ash content. Wood properties and secondary cell wall deposition are controlled by complex gene regulatory networks [20, 21]. In addition to the underlying genetic control, wood formation is significantly impacted by environmental factors and plant growth habit (tension and compression wood), adding to the complexity of these highly polygenic traits and warranting further research. There is limited research on woody biomass traits and their relationship to ploidy levels and no known studies in willow. There has been some research on Chinese white poplar (*Populus tomentosa*) that has demonstrated that wood properties are more desirable in the triploids than in diploids [19, 22].

The purpose of this study was to demonstrate that ploidy level impacts shrub willow biomass traits and that these biomass traits are affected by harvest cycle year. We evaluated woody biomass traits among progeny and parental genotypes with different ploidy content as reported by Serapiglia et al. [17]. We compared traits among genotypes, different ploidy levels, and across multiple years of growth during a single, 3-year harvest cycle in a replicated trial at a single site, to understand how wood properties change in response to genome size and year of harvest. In addition, compositional analysis was performed on a subset of willow samples and used to further calibrate and refine a thermogravimetric analytical method developed by Serapiglia et al. [23].

## Materials and Methods

### Source Material and Tissue Collection

Stem segments were collected for analysis from 75 genotypes (Online Resource 1) in the 2008 Selection Trial at the Cornell University New York State Agricultural Experiment Station in Geneva, NY [17]. The trial was planted in a randomized complete block design, with three replicate blocks. Samples were collected from all three replicate blocks within the trial. Following each growing season of 2009, 2010, and 2011 during winter dormancy, a single stem representative of a typical canopy stem within each plot was selected from one plant and a 25-cm section collected at or near the centroid of the stem (~1.5 m above the soil). Biomass composition was

determined using high-resolution thermogravimetric analysis (HR-TGA) with a modified method from Serapiglia et al. [23]. Specific gravity of stem segments was determined by volumetric displacement [24]. These data were compared with growth measurements (height, stem number, stem diameter, and stem area) and final harvest yield [17].

In addition, stem segments from an expanded set of different field trials and different age classes were selected for analysis of biomass composition through wet chemical methods and to further calibrate and refine the thermogravimetric method in [23]. Data determined from these analyses included insoluble lignin, soluble lignin, high-performance liquid chromatography (HPLC) sugar profiles, and syringyl/guaiacyl (S:G) ratios.

### Wood Chemistry Analysis

Oven-dried stem segments (not debarked) were ground through a 0.5-mm screen using a MF 10 knife mill (IKA, Wilmington, NC). All samples were Soxhlet extracted overnight with hot 95 % acetone. Lignin and carbohydrate content was determined on extracted biomass according to Huntley et al. [25]. Following hydrolysis, acid-insoluble lignin was determined gravimetrically using pre-weighed oven-dried sintered glass crucibles. Acid-soluble lignin was determined spectrophotometrically at 205 nm. The components of the hydrolysate were quantified by HPLC using a Dionex DX-600 system (Dionex, Sunnyville, CA) with a SpectraAS3500 auto injector (Spectra-physics, Santa Clara, CA) and a Dionex CarboPac PA1 anion-exchange column. Monomeric sugars were quantified using a pulsed electrochemical detector with a gold electrode. The column was equilibrated with 250 mM NaOH, and the mobile phase was water maintained at a flow rate of 1 mL min<sup>-1</sup>. The concentrations of all monomers were calculated based on standards with known concentrations of arabinose, galactose, glucose, mannose, rhamnose, and xylose, with fucose as an internal standard.

Syringyl and guaiacyl lignin monomer composition (S:G ratio) was determined according to Robinson and Mansfield [26]. Following thioacidolysis, samples were analyzed by gas chromatography using a Hewlett Packard 5890 series II system (Agilent Technologies, Santa Clara, CA) with a splitless injector, a flame ionizing detector, and a 30 m 5 % diphenyl-95 % dimethyl polysiloxane-coated RTX-5MS capillary column. For the genotypes in the 2008 Selection Trial, thioacidolysis was only performed on the third-year growth and from only one replicate within the trial.

### Compositional Analysis by HR-TGA

Willow biomass samples (10 mg of unextracted dry tissue for each sample) were analyzed using a Thermogravimetric

Analyzer 2950 (TA Instruments, New Castle, DE) with the TA Universal Analysis 2000 software. Samples from each of the three replicate blocks of the trial were analyzed once and means calculated for cellulose, lignin, and ash content as a proportion of the total biomass. The method used was “high-resolution dynamic” with a heating rate of  $20^{\circ}\text{C min}^{-1}$ , a final temperature of  $600^{\circ}\text{C}$ , a resolution of 4.0, and a sensitivity value of 1.0. The electrobalance was purged with nitrogen at a flow rate of  $40\text{ mL min}^{-1}$ , and the furnace was purged with compressed air at a flow rate of  $60\text{ mL min}^{-1}$ .

Data points from each individual thermogram were imported into Microsoft Excel to produce an X-matrix for statistical analysis. Using a macro created in Microsoft Excel, text files with TGA data were imported into Excel and each sample was corrected for water loss by identifying the temperature at which the derivative was closest to zero between  $100$  and  $150^{\circ}\text{C}$ . The weight at this temperature was established as the initial weight of that sample (100% relative weight). Following this, datasets were normalized for each thermogram, so each sample dataset had an equal number of data points within the temperature range from  $100$  to  $599^{\circ}\text{C}$ . The variation in temperature from run to run was corrected for by setting the point of maximum weight loss for cellulose to  $300^{\circ}\text{C}$  (highest point on the derivative curve) for all samples. After evaluation of the X-matrix loadings and weights, each thermogram was individually analyzed for cellulose and lignin content by identifying the thermal events associated with the combustion of each compound on the derivative curve. This method was modified from Serapiglia et al. [23] as follows: the weight loss from combustion of cellulose was that which occurred between  $275$  and  $330^{\circ}\text{C}$  ( $25^{\circ}$  less than the cellulose peak at  $300^{\circ}$  and  $30^{\circ}$  greater), and the weight loss associated with lignin combustion was that which occurred between  $370$  and  $490^{\circ}\text{C}$  ( $70^{\circ}$  and  $190^{\circ}$  greater than the cellulose peak). As the thermal degradation of hemicellulose and some amorphous cellulose polymers may be occurring simultaneously at certain temperatures, the quantification of cellulose by TGA potentially may include some content of hemicellulose. Ash content was determined by subtracting the final percent mass loss from 100.

### Statistical Analysis

Partial least square (PLS) models were derived from the corrected thermogram data using PROC PLS in SAS<sup>®</sup> version 9.3 [27]. Models were created for both the prediction of insoluble lignin and glucose/cellulose. Results of acid-soluble lignin analyses were not used in the model. Each model was validated by cross-validation and predictive capabilities were determined by the coefficient of determination  $r^2$

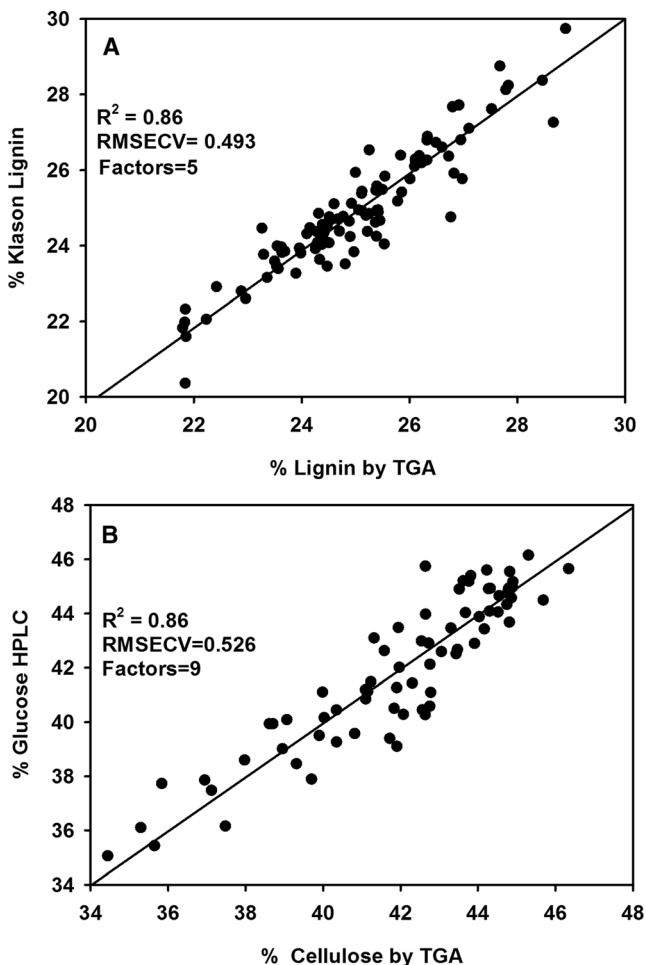
and root-mean standard error of cross-validation (RMSECV). Principle component analysis (PROC PCA) was also performed on the data sets to identify outliers. These outliers were removed prior to PLS model development.

Data collected across the three years of growth in the 2008 selection trial were analyzed in SAS<sup>®</sup> version 9.3 with a critical  $\alpha$  level of 0.05. Normality of distribution for each trait was checked by Kolmogorov D and Shapiro-Wilk's statistic W and graphically by normal-probability plots and histograms using PROC UNIVARIATE. No transformation of the data was required. PROC GLM was used to perform analysis of variance and to estimate significant effects of the factors and any interactions. A three-way factorial ANOVA was performed including three years of measurements, three field blocks, and three ploidy levels. In addition, another three-way factorial ANOVA was performed for genotype, with three years of measurements, three field blocks, and 75 genotypes. PROC GLM takes into account uneven sampling and uses type III sum of squares which tests for differences in the weighted means, which is necessary when comparing ploidy, since there was a different number of diploids, triploids, and tetraploids measured. This was also used for comparisons across families with multiple progeny. Analysis of genotype and ploidy was kept independent of each other because combining both genotype and ploidy into the same model led to nesting of genotype within ploidy, preventing comparison of all genotypes to one another. PROC CORR was used to identify any significant correlations among variables obtained in this study. Correlations were performed on each year with all 75 genotypes combined and separated by ploidy. Ordinary least squares regression was used to estimate the predictive ability of third-year composition from first-year measurements.

## Results

### Prediction of Insoluble Lignin and Cellulose Content by HR-TGA

Correlation coefficients for results of wet chemical analysis with HR-TGA results were 0.86 for both lignin and cellulose (Fig. 1). The predictive PLS model for lignin had an RMSECV of 0.49 with 5 PLS factors and that for cellulose had a RMSECV of 0.53 with 9 PLS factors. For lignin, 98 samples were used in the calibration model, and for cellulose, 76 samples were used. Through PCA, more outliers were identified for cellulose content than for lignin. Predictive models for hemicellulose (predicted from all sugars other than



**Fig. 1** Thermogravimetric analysis (TGA) models from partial least square (PLS) analysis. **a** Predicted lignin content from TGA vs. Klason lignin content. **b** Predicted cellulose content from TGA vs. glucose content determined by HPLC. Root-mean standard error of cross-validation (RMSECV) and the number of factors in each optimized model are indicated

glucose) were poor and are not presented. The range in variation of the non-glucose sugars across all the samples was low making it difficult to develop a good PLS model (Online Resource 2).

### Biomass Composition

Biomass traits from the 75 genotypes in the 2008 Selection Trial were analyzed at the end of each growing season over the 3-year growth cycle. All traits were significantly different by year of growth and genotype (Table 1). There were significant interactions between year and genotype for cellulose, lignin, and ash content. Lignin and ash content were significantly different by ploidy, but cellulose content was not. Ploidy of each genotype is described in Serapiglia et al. [17]. There was a significant interaction by year and ploidy for cellulose and ash, but not for lignin. Analysis by year showed that cellulose content was only significantly different by ploidy in the first

year of growth. Lignin was significantly different by ploidy regardless of year. Ash content was significantly different by ploidy in the first and third year of growth. At the time of harvest after 3 years of growth, cellulose content was greatest in the diploid, *S. × dasyclados* 'SV1' (Fig. 2a). The cellulose content in the diploid genotypes ranged from 40.3 to 46.4% and in the triploids was from 39.9 to 45.8%. The lowest cellulose content was observed in the triploid, *S. miyabeana* × (*S. viminalis* × (*S. schwerinii* × *S. viminalis*)) clone ID 02X-326-015 (Fig. 2b). The tetraploid genotypes had the smallest range in cellulose content, from 41.7 to 44.7% (Fig. 2c), which may simply reflect that fewer tetraploid genotypes were characterized. An examination of the progeny within each family showed that on average, the triploid progeny in family 05X-295 and the diploid progeny in family 05X-298 had the greatest cellulose content (44.0 and 43.9%, respectively). Two tetraploid progeny in family 01X-264 had the lowest cellulose content (41.8%).

On average, the diploids had the highest lignin content across all 3 years (Figs. 3a and 4b). However, the genotype with the greatest lignin content in the third year of growth was a triploid, *S. viminalis* × (*S. viminalis* × *S. miyabeana*) 'Sheridan'. It also had among the lowest cellulose content. The lowest lignin content was observed in a tetraploid, *S. miyabeana* 05X-287-013. On average, the triploid progeny in family 05X-291 had the lowest lignin content (24.3%). Lignin content was greatest in the first year of growth and lowest in the second year of growth (Fig. 4b). An opposite pattern was observed for cellulose content across the years (Fig. 4a). Cellulose content in the diploids was the lowest in the first year of growth (Fig. 4a). As with lignin, ash content was lowest in the second year of growth, but tetraploids displayed the greatest ash content in the third year of growth (Figs. 4c and 5c). The diploid progeny in family 05X-299 (2.1%) and the tetraploid progeny in family 05X-287 (2.3%) had the greatest ash content (Fig. 5).

S:G ratios were determined for third-year biomass from samples from only one block of the field trial (Fig. 6). The other replicates were not analyzed due to cost and time of analysis, precluding any statistical analyses by genotype. For the 75 genotypes analyzed, there was a wide range in the S:G ratio from 1.6 to 2.9. The range and average S:G ratio across genotypes of different ploidy were fairly similar, indicating that ploidy may not have an impact on S:G ratio. The mean S:G ratio of the diploids was 2.1 and for triploids and tetraploids was 2.3. The greatest S:G ratio was for a diploid, *S. × dasyclados* 'SV1', and the lowest S:G ratio was found in a diploid, *S. purpurea* clone ID 00-01-088.

### Correlations with Growth Traits

All the traits we measured were significantly correlated with third-year biomass yield, except for stem number when

**Table 1** ANOVA results for biomass traits measured across the 3 years of growth

| Source effect     | DF  | F value     |          |        | p value           |                   |                   |
|-------------------|-----|-------------|----------|--------|-------------------|-------------------|-------------------|
|                   |     | % Cellulose | % Lignin | % Ash  | % Cellulose       | % Lignin          | % Ash             |
| Year              | 2   | 434.32      | 1167.55  | 116.13 | <i>&lt;0.0001</i> | <i>&lt;0.0001</i> | <i>&lt;0.0001</i> |
| Block             | 2   | 4.60        | 1.61     | 0.46   | <i>0.01</i>       | 0.20              | 0.60              |
| Genotype          | 74  | 4.16        | 3.96     | 3.63   | <i>&lt;0.0001</i> | <i>&lt;0.0001</i> | <i>&lt;0.0001</i> |
| Genotype×block    | 148 | 1.63        | 1.35     | 1.40   | 0.53              | 0.50              | 0.57              |
| Genotype×year     | 148 | 0.99        | 1.01     | 0.99   | <i>0.0002</i>     | <i>0.01</i>       | <i>0.01</i>       |
| Year              | 2   | 233.48      | 664.40   | 78.35  | <i>&lt;0.0001</i> | <i>&lt;0.0001</i> | <i>&lt;0.0001</i> |
| Block             | 2   | 2.77        | 0.94     | 0.41   | 0.06              | 0.33              | 0.66              |
| Ploidy            | 2   | 1.20        | 10.53    | 11.32  | 0.22              | <i>&lt;0.0001</i> | <i>&lt;0.0001</i> |
| Ploidy×block      | 4   | 0.72        | 0.28     | 0.36   | 0.59              | 0.89              | 0.83              |
| Year×ploidy       | 4   | 2.38        | 0.35     | 3.82   | <i>0.03</i>       | 0.82              | <i>0.003</i>      |
| Ploidy for year 1 | 2   | 6.64        | 3.36     | 6.10   | <i>0.001</i>      | <i>0.04</i>       | <i>0.003</i>      |
| Ploidy for year 2 | 2   | 0.17        | 4.61     | 1.44   | 0.84              | <i>0.01</i>       | 0.24              |
| Ploidy for year 3 | 2   | 1.15        | 4.96     | 17.72  | 0.32              | <i>0.01</i>       | <i>&lt;0.0001</i> |

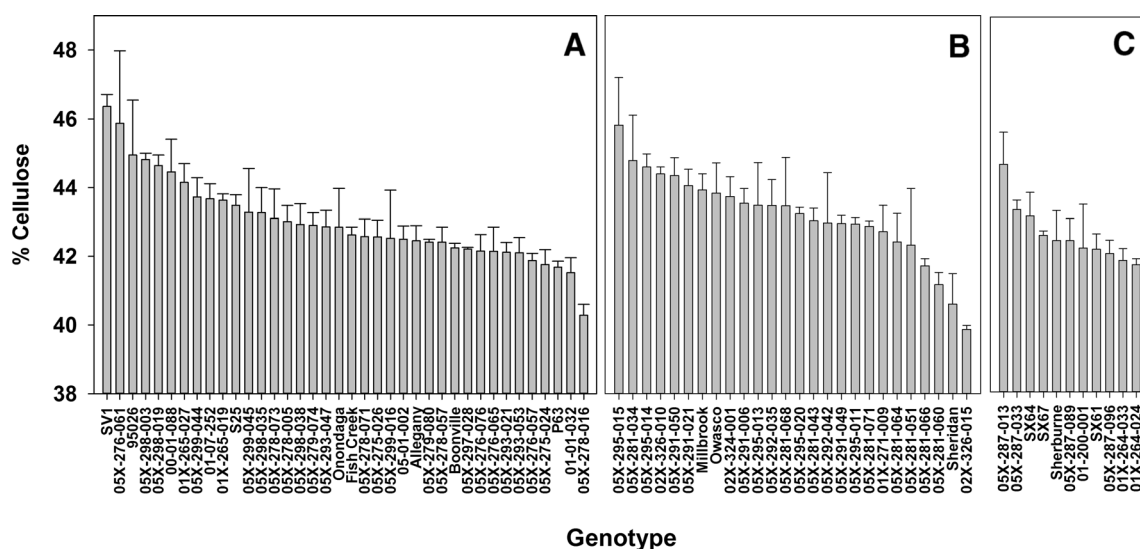
Significant effects are in italic

examining all genotypes as a whole (Table 2). Second-year measurements of cellulose and ash content did not correlate with yield. There was a strong inverse relationship between cellulose content and lignin content across all 3 years of growth. Lignin content was negatively correlated with yield regardless of year of growth and was also negatively correlated with height (strongest in the first year). The S:G ratio measured in year 3 was significantly correlated with all growth measurements, all compositional components, and yield. Correlations by ploidy in the third year of growth indicate that many of the relationships with compositional traits observed are in the diploids. Height correlated with all compositional traits in the

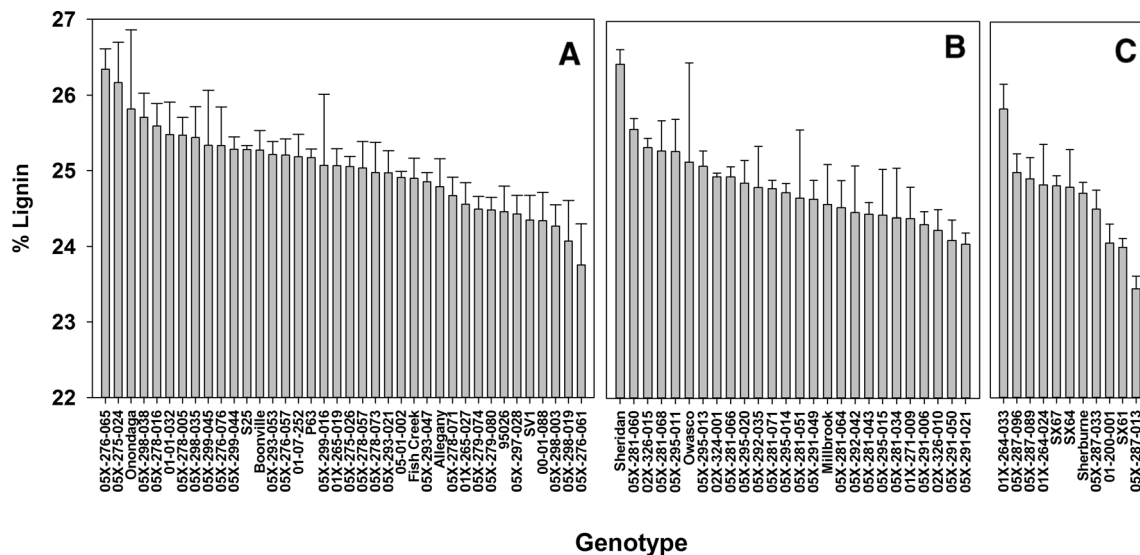
diploids only. Specific gravity did not correlate with cellulose content in the tetraploids. Cellulose content was negatively correlated with lignin and ash across all ploidy levels. The S:G ratio correlated with cellulose and lignin in the diploids and tetraploids, but not in the triploids.

## Discussion

While it has long been suspected that triploids were more vigorous than their diploid or tetraploid parents [13], data



**Fig. 2** Cellulose content in shrub willow biomass at harvest after three seasons of growth. **a** Diploid genotypes. **b** Triploid genotypes. **c** Tetraploid genotypes. Data are the mean of three field replicates±standard error



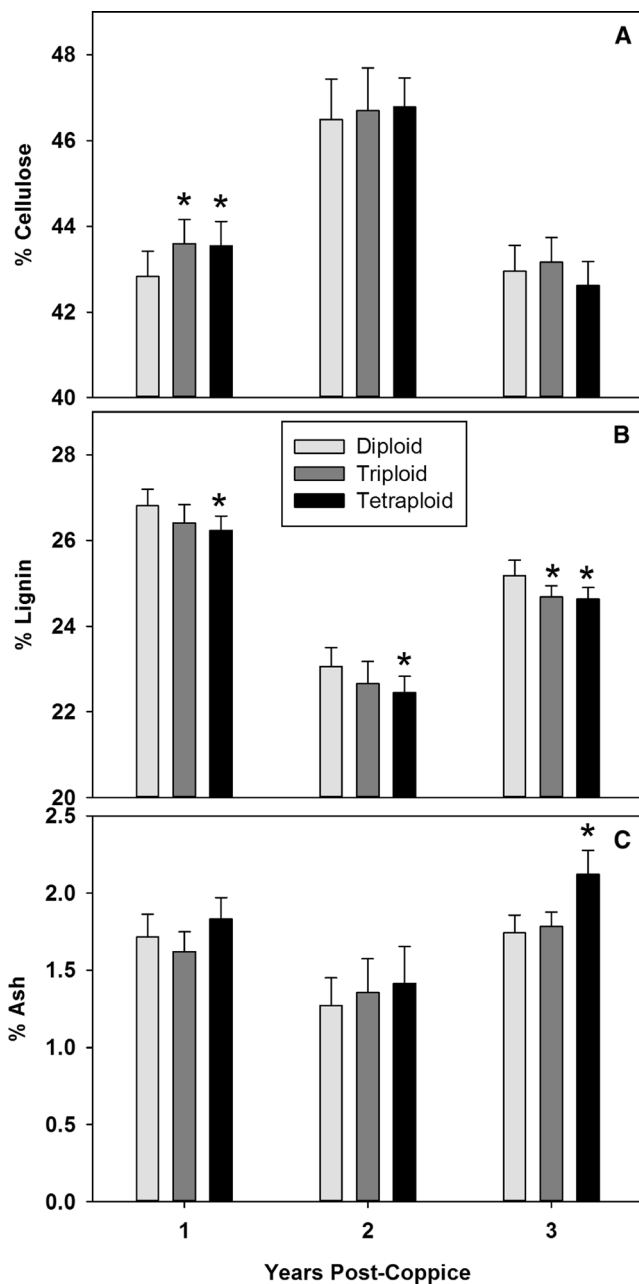
**Fig. 3** Lignin content in shrub willow biomass at harvest after three seasons of growth. **a** Diploid genotypes. **b** Triploid genotypes. **c** Tetraploid genotypes. Data are the mean of three field replicates  $\pm$  standard error

from recent trials with species hybrids of shrub willow confirms that a disproportionate number of top performing genotypes are triploid [17]. We further characterized these novel hybrid progeny for biomass quality traits important to the biofuel industry using HR-TGA. Overall, triploid and tetraploid genotypes had significantly lower lignin content than diploids at harvest. This has also been observed among triploid genotypes of *P. tomentosa* compared to wild-type diploid genotypes. In *P. tomentosa*, triploids not only had faster overall growth (height, diameter at breast height, volume, and biomass), but fiber length and width were greater in triploids than in a diploid clone [19, 22]. Polyploids of white ash (*Fraxinus americana*) have also been identified as having longer fibers and vessel elements [28]. These studies have shown that in addition to improving overall growth, triploids also display improvements in biomass quality traits. The biomass of triploid willow in this trial also had the highest specific gravity as reported in Serapiglia et al. [17], similar to what was seen in triploid aspen [29–31].

Hu et al. [32] recently found that genes associated with sugar transport and carbon metabolism were upregulated in leaves of triploid hybrid poplar (*Populus*  $\times$  *euramericana*) compared to diploids, although they also observed that genes involved in lignin synthesis were upregulated too. They suggested that increased cellulose and lignin metabolism could play an important role in overall growth and vigor of triploids, but they did not analyze wood composition in their triploids. Our results suggest that lignin synthesis should be lower in developing stems, but Hu et al. studied leaf RNA, not

stem RNA [32]. Heterosis in triploid willow is likely not due to heterozygosity masking the effects of deleterious alleles [17], but more probably a result of gene dosage effect [33, 34]. We are currently conducting transcriptome analysis to compare patterns of gene expression in polyploids with those in diploids. The results of these studies should determine if expression of critical genes associated with these key biomass traits are related to gene dosage. These results could provide further insight into yield and biomass quality improvements in shrub willow.

Across all genotypes, there was a negative relationship between lignin content and yield, as well as a negative correlation with height across all 3 years. This relationship has been observed in poplar and other tree species [35]. Kirst et al. [36] identified overlapping genomic regions (QTLs) regulating both growth and lignin content in Eucalyptus. Due to the high energy content of lignin, the biosynthesis of this polymer requires more carbon and energy input to make, which could reduce energy and carbon allocated to other cellular growth processes impacting the overall growth of the plant. It is well known that lignin has a higher energy content and higher heat value than cellulose [37, 38]. As suggested by Novaes et al. [35], lignin requires high levels of reduced carbon for synthesis compared to that of cellulose; therefore, in theory, reducing lignin content should increase wood production. In the current selection of willow genotypes, triploid and tetraploid genotypes had lower lignin content and improved yield compared to diploids. While ash content was greater in the tetraploids, triploids maintained lower



**Fig. 4** Shrub willow biomass composition across the 3-year growth cycle. **a** % cellulose, **b** % lignin, **c** % ash. Data points are the mean of the genotypes of each ploidy: diploid, triploid, and tetraploid. \* indicates significant difference from diploids

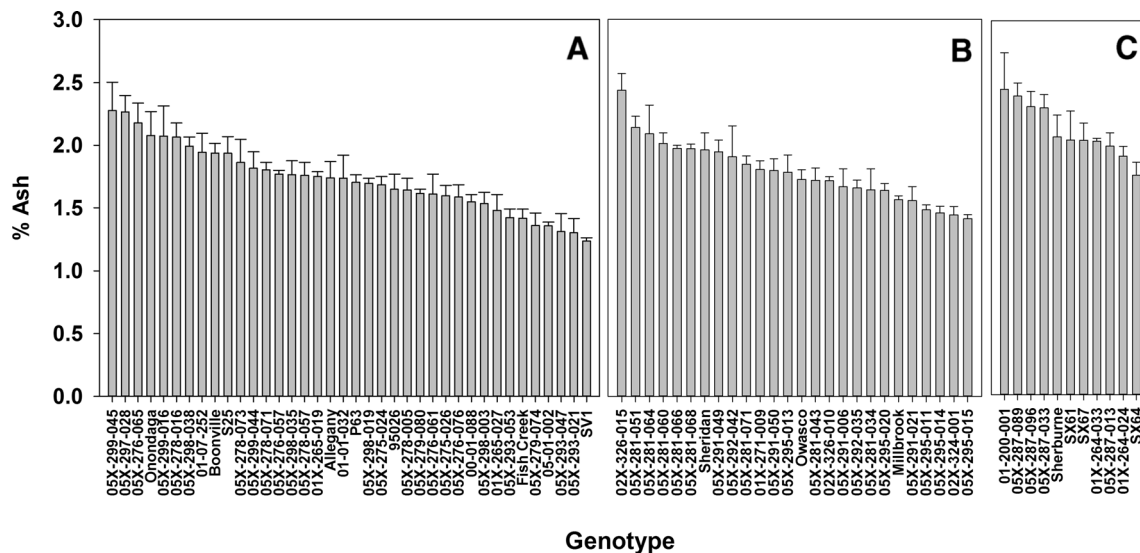
ash similar to those of the diploids, which would be an added benefit for conversion technologies. Ash is a critical issue in both biochemical and thermochemical conversion methodologies, and results here show that through breeding, differences in ash content can be selected for.

Overall, cellulose was positively correlated with yield and strongly negatively correlated with lignin and ash.

With these relationships, it is difficult to explain why cellulose was not significantly greater in the triploids, which produced greater yield and lower lignin content. One key observation with the correlations between biomass traits, growth traits, and yield is that many of these relationships break down in the polyploids compared with the diploids. In year 3, the growth traits correlated with compositional traits in the diploids, but most of these relationships were lost in the triploids. Serapiglia et al. [17] described significant differences in growth patterns and plant structure among diploids, triploids, and tetraploids, with diploids being shorter with more stems, tetraploids being taller with fewer stems per plant, while triploids were the tallest with intermediate numbers of stems. Biomass production and overall wood deposition, in theory, should be directly associated with carbon allocation into the cell wall, which would explain why there are relationships between composition and growth traits. But the genetic differences in these triploid hybrids are potentially having a strong impact on this relationship.

Cellulose content of third-year biomass had a range of 6% across all genotypes, while lignin content had a range of 3%. One question to address is whether or not these differences can have a significant impact on downstream conversion processes. Serapiglia et al. [39] demonstrated that an increase in cellulose content of 6% in willow can increase sugar release from enzymatic hydrolysis of pretreated biomass by up to 50%. Van Acker et al. [40] showed that transgenic poplar with a 5% reduction in lignin content can increase ethanol yield by 161% compared to wild-type poplar. It has also been shown that sugar release from willow biomass can increase 5-fold without a significant change in total lignin content [41]. Studer et al. [42] demonstrated that an increase in saccharification efficiency and ethanol yield was more related to the S:G ratio in the lignin (S:G ratio greater than 2), rather than the reduction in total lignin. Similar results were obtained with poplar engineered to have altered lignin monomer compositions [43]. With the S:G ratio in the lignin of shrub willow ranging from 1.6 to 2.9, it is promising that this variation can be exploited in breeding to improve progeny, but also to gain more understanding about lignin and its complex interaction with the other polymers in the cell wall.

Across the 3-year growth cycle of this willow trial, cellulose content was greatest and lignin and ash content were the lowest in the second year. Specific gravity follows a similar trend to cellulose across the years as reported in Serapiglia et al. [17]. This is the first experimental field trial that has examined biomass

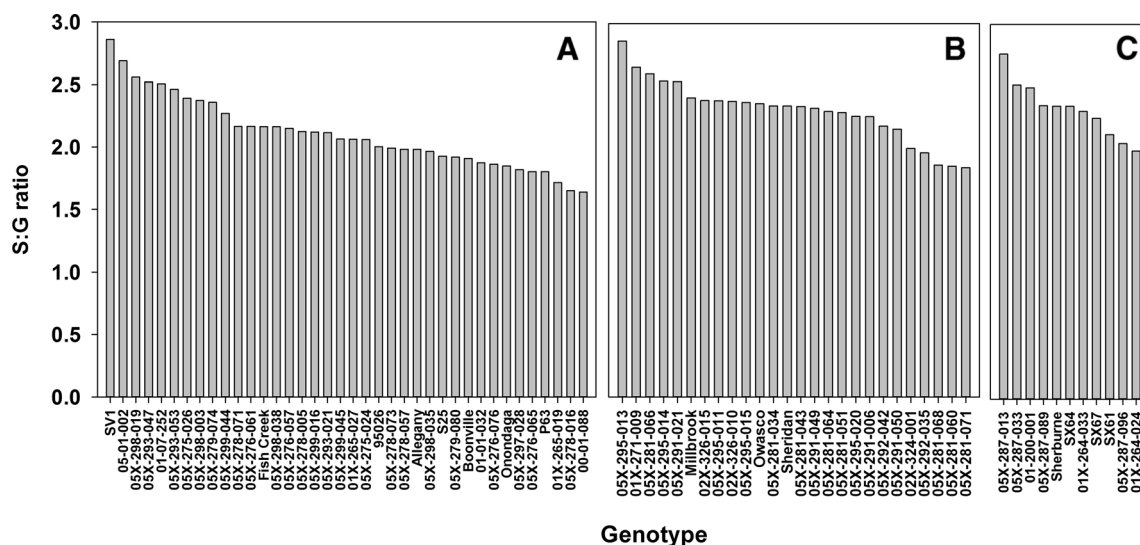


**Fig. 5** Ash content in shrub willow biomass at harvest after three seasons of growth. **a** Diploid genotypes. **b** Triploid genotypes. **c** Tetraploid genotypes. Data are the mean of three field replicates  $\pm$  standard error

composition over the 3-year harvest cycle commonly practiced for shrub willow. Depending on growing conditions, market prices, and weather, growers may elect to harvest on a 2-year instead of 3-year cycle, which could potentially impact the composition of the harvested biomass. The greater cellulose content and lower ash content observed in the second year may result in feedstock that is more suitable for biochemical conversion applications. There is also the question of how well first-year growth sampling can predict biomass composition at harvest after three seasons. Ordinary least squares regressions with year 1 composition and year

3 composition are significant, but weak ( $R^2 < 0.21$ ), indicating that year 1 compositional data is only slightly indicative of year 3 composition (Online Resource 3). Some caution is warranted when performing studies on younger willow plants, since there are significant changes in composition and biomass traits as the plants age.

As the genomic resources for shrub willow expand and functional genomics and genomic selection advance the shrub willow breeding programs, rapid, low-cost methods for phenotypic analysis should continue to prove more vital. As a result, the further refinement of the HR-TGA method has become a critical step in the genetic improvement of shrub



**Fig. 6** Syringyl/guaiacyl (S:G) ratio in shrub willow biomass at harvest after three seasons of growth. **a** Diploid genotypes. **b** Triploid genotypes. **c** Tetraploid genotypes. Data represent analysis of a single field replicate

**Table 2** Correlations of yearly measurements with yield harvested after three seasons

|                                       | %<br>Cellulose | % Lignin | % Ash    | Yield   | S:G      |
|---------------------------------------|----------------|----------|----------|---------|----------|
| First year post-coppice               |                |          |          |         |          |
| Height                                | 0.421**        | −0.369** | −0.262** | 0.676** | –        |
| Stem diameter                         | −0.121*        | −0.019   | −0.023   | 0.290** | –        |
| Stem area                             | 0.156*         | −0.209*  | −0.072   | 0.637** | –        |
| Stem number                           | −0.281**       | 0.137*   | 0.031    | −0.029  | –        |
| Specific gravity                      | 0.119          | 0.058    | −0.182*  | 0.134*  | –        |
| % Cellulose                           | –              | −0.538** | −0.614** | 0.303** | –        |
| % Lignin                              |                | –        | 0.290**  | −0.230* | –        |
| % Ash                                 |                |          | –        | −0.242* | –        |
| Second year post-coppice              |                |          |          |         |          |
| Height                                | 0.041          | −0.137*  | −0.004   | 0.780** | –        |
| Stem diameter                         | −0.074         | 0.015    | −0.009   | 0.284** | –        |
| Stem area                             | −0.042         | −0.048   | 0.053    | 0.614** | –        |
| Stem number                           | −0.082         | 0.076    | −0.033   | −0.034  | –        |
| Specific gravity                      | 0.016          | 0.193*   | −0.187*  | 0.237*  | –        |
| % Cellulose                           | –              | −0.702** | −0.349** | 0.005   | –        |
| % Lignin                              |                | –        | 0.09175  | −0.147* | –        |
| % Ash                                 |                |          | –        | 0.042   | –        |
| Third year post-coppice—all genotypes |                |          |          |         |          |
| Height                                | 0.137*         | −0.292** | −0.167*  | 0.774** | 0.282**  |
| Stem diameter                         | 0.037          | −0.006   | −0.163*  | 0.315** | −0.160*  |
| Stem area                             | 0.044          | −0.147*  | 0.013    | 0.704** | 0.127*   |
| Stem number                           | −0.014         | 0.099    | −0.151*  | −0.114  | −0.290** |
| Specific gravity                      | 0.240*         | 0.100    | −0.430** | 0.132*  | 0.260**  |
| % Cellulose                           | –              | −0.653** | −0.498** | 0.214*  | 0.214*   |
| % Lignin                              |                | –        | 0.328**  | −0.328* | −0.232*  |
| % Ash                                 |                |          | –        | −0.145* | −0.125*  |
| Yield                                 |                |          |          | –       | 0.268**  |
| Third-year diploids                   |                |          |          |         |          |
| Height                                | 0.251*         | −0.349*  | −0.513** | 0.722** | 0.207*   |
| Stem diameter                         | 0.102          | −0.102   | −0.148   | 0.454** | −0.153   |
| Stem area                             | 0.216*         | −0.187*  | −0.266*  | 0.635** | 0.001    |
| Stem number                           | −0.022         | 0.011    | 0.005    | 0.208*  | −0.209*  |
| Specific gravity                      | 0.237          | −0.069   | −0.394** | 0.311*  | 0.387**  |
| % Cellulose                           | –              | −0.678** | −0.417** | 0.152   | 0.291*   |
| % Lignin                              |                | –        | 0.600**  | −0.184* | −0.180*  |
| % Ash                                 |                |          | –        | −0.460* | −0.372** |
| Yield                                 |                |          |          | –       | 0.130    |

**Table 2** (continued)

|                        | %<br>Cellulose | % Lignin | % Ash    | Yield   | S:G     |
|------------------------|----------------|----------|----------|---------|---------|
| Third-year triploids   |                |          |          |         |         |
| Height                 | 0.044          | −0.058   | −0.054   | 0.581** | 0.007   |
| Stem diameter          | −0.129         | 0.058    | 0.139    | 0.438** | −0.136  |
| Stem area              | −0.164         | 0.049    | 0.226*   | 0.505** | −0.138  |
| Stem number            | −0.072         | 0.062    | −0.001   | 0.155   | −0.174  |
| Specific gravity       | 0.259*         | 0.178    | −0.421** | 0.115   | 0.300*  |
| % Cellulose            | –              | −0.704** | −0.651** | 0.106   | 0.082   |
| % Lignin               |                | –        | 0.313*   | −0.106  | −0.143  |
| % Ash                  |                |          | –        | −0.105  | −0.089  |
| Yield                  |                |          |          | –       | 0.002   |
| Third-year tetraploids |                |          |          |         |         |
| Height                 | 0.137          | −0.224   | −0.160   | 0.673** | −0.280  |
| Stem diameter          | −0.065         | 0.140    | −0.332*  | 0.275   | −0.141  |
| Stem area              | 0.078          | −0.088   | −0.063   | 0.678** | −0.064  |
| Stem number            | −0.109         | 0.169    | −0.396*  | −0.139  | −0.141  |
| Specific gravity       | 0.079          | 0.420*   | −0.203   | −0.002  | −0.127  |
| % Cellulose            | –              | −0.557** | −0.543** | −0.289  | 0.531** |
| % Lignin               |                | –        | 0.077    | −0.350* | −0.409* |
| % Ash                  |                |          | –        | −0.204  | 0.092   |
| Yield                  |                |          |          | –       | 0.046   |

\* $p < 0.05$  (significant correlation); \*\* $p < 0.0001$  (significant correlation)

willow. The predictive models for lignin and cellulose improved upon the correlation coefficients presented in Serapiglia et al. [23]. The current models for lignin and cellulose and the ash content from the TGA results demonstrated the wide range of compositional differences and the high variability in these traits in the measured willow genotypes. In addition, we were able to demonstrate that polyploidy in shrub willow can not only impact changes in growth and specific gravity, but also biomass composition. The lower lignin content and higher yield observed in the triploids and tetraploid progeny are promising for future breeding efforts to produce genotypes suited for conversion to bioenergy. Ultimately, improvements in total biomass yield, biomass quality, and biomass energy content must be matched with the conversion process for which the willow is being grown as a feedstock. The variation in biomass composition documented for shrub willow provides optimism that selection for appropriate feedstock characteristics can readily be accomplished.

**Acknowledgments** This work was funded by grants from the Northeast Sun Grant Initiative from the US Department of Transportation and the US Department of Agriculture. The authors would like to thank Kayla Relyea, Michelle Von Loessl, Berdine Coetzee, and Rebecca Chase for their assistance on this project. Steve Gordner and Matt Christiansen provided expert assistance with field trial establishment, maintenance, and harvesting. Additional technical support was provided by Kimberly Cameron, Michael Rosato, Kayleigh Hogan, Jane Petzoldt, Cody Lafler, and James Ballerstein.

## References

- Argus GW (1997) Infrageneric classification of *Salix* (*Salicaceae*) in the New World. The American Society of Plant Taxonomists, Mich, Ann Arbor
- Larsson S (1998) Genetic improvement of willow for short-rotation coppice. *Biomass Bioenerg* 15:23–26
- Smart LB, Cameron KD (2008) Genetic improvement of willow (*Salix* spp.) as a dedicated energy crop. In: Vermerris WE (ed) Genetic improvement of Bioenergy crops. Springer Science, New York, pp 347–376
- Karp A, Hanley S, Trybush S, Macalpine W, Pei MH, Shield I (2011) Genetic improvement of willow for bioenergy and biofuels. *J Integr Plant Biol* 53:151–165
- Volk TA, Abrahamson LP, Cameron KD, Castellano P, Corbin T, Fabio E, Johnson G, Kuzovkina-Eischen J, Labrecque M, Miller R, Sidders D, Smart LB, Staver K, Stanosz GR, Van Rees K (2011) Yields of willow biomass crops across a range of sites in North America. *Aspects Appl Biol* 112:67–74
- Serapiglia MJ, Cameron KD, Stipanovic AJ, Abrahamson LP, Volk TA, Smart LB (2012) Yield and woody biomass traits of novel shrub willow hybrids at two contrasting sites. *BioEnergy Res* 6:533–546
- Lin J, Gibbs JP, Smart LB (2009) Population genetic structure of native versus naturalized sympatric shrub willows (*Salix*; *Salicaceae*). *Am J Bot* 96:771–785
- Weih M, Rönnerberg-Wästljung A-C, Glynn C (2006) Genetic basis of phenotypic correlations among growth traits in hybrid willow (*Salix dasyclados* × *S. viminalis*) grown under two water regimes. *New Phytol* 170:467–477
- Cameron KD, Phillips IS, Kopp RF, Volk TA, Maynard CA, Abrahamson LP, Smart LB (2008) Quantitative genetics of traits indicative of biomass production and heterosis in 34 full-sib *F<sub>1</sub>* *Salix eriocephala* families. *BioEnergy Res* 1:80–90
- Orians CM, Huang C, Wild A, Zee P, Dao MTT, Fritz RS (1997) Willow hybridization differentially affects preference and performance of herbivorous beetles. *Entomol Exp Appl* 83:285–294
- Pei MH, Lindegaard K, Ruiz C, Bayon C (2008) Rust resistance of some varieties and recently bred genotypes of biomass willows. *Biomass Bioenerg* 32:453–459
- Pei MH, Shield I, Macalpine W, Lindegaard KN, Bayon C, Karp A (2010) Mendelian inheritance of rust resistance to *Melampsora larici-epitea* in crosses between *Salix sachalinensis* and *S. viminalis*. *Plant Pathol* 59:862–872
- Zsuffa L, Mosseler A, Raj Y (1984) Prospects for interspecific hybridization in willow for biomass production. In: Perttu K (ed) Ecology and management of forest biomass production systems, vol 15. Swed. Univ. Agric. Sci, Uppsala, pp 261–281
- Riedelsheimer C, Czedik-Eysenberg A, Grieder C, Lisek J, Technow F, Sulpire R, Altmann T, Stitt M, Willmitzer L, Melchinger AE (2012) Genomic and metabolic prediction of complex heterotic traits in hybrid maize. *Nature Genet* 44:217–220
- Chen ZJ (2013) Genomic and epigenetic insights into the molecular bases of heterosis. *Nat Rev Genet* 14:471–482
- Goff SA, Zhang Q (2013) Heterosis in elite hybrid rice: speculation on the genetic and biochemical mechanism. *Curr Opin Plant Biol* 16: 221–227
- Serapiglia MJ, Gouker FG, Smart LB (2014) Early selection of novel triploid hybrids of shrub willow with improved biomass yield relative to diploids. *BMC Plant Biol* 14:74. doi:10.1186/1471-2229-1114-1174
- Nilsson-Ehle H (1938) Über eine in der naturgefundene gigasform von *Populus tremula*. *Hereditas* 21:379–383
- Zhang P, Wu F, Kang X (2012) Genotypic variation in wood properties and growth traits of triploid hybrid clones of *Populus tomentosa* at three clonal trials. *Tree Genet Genomes* 8:1041–1050
- Denis M, Favreau B, Ueno S, Camus-Kulandaivelu L, Chaix G, Gion JM, Nourrisier-Mountou S, Polidori J, Bouvet JM (2013) Genetic variation of wood chemical traits and association with underlying genes in *Eucalyptus urophylla*. *Tree Genet Genomes* 9:927–942
- Porth I, Klápště J, Skyba O, Friedmann MC, Hannemann J, Ehling J, El-Kassaby YA, Mansfield SD, Douglas CJ (2013) Network analysis reveals the relationship among wood properties, gene expression levels and genotypes of natural *Populus trichocarpa* accessions. *New Phytol* 200:727–742
- Zhang P, Wu F, Kang X (2013) Genetic control of fiber properties and growth in triploid hybrid clones of *Populus tomentosa*. *Scand J Forest Res* 28:621–630
- Serapiglia MJ, Cameron KD, Stipanovic AJ, Smart LB (2009) Analysis of biomass composition using high-resolution thermogravimetric analysis and percent bark content as tools for the selection of shrub willow bioenergy crop varieties. *BioEnergy Res* 2:1–9
- TAPPI Standard T 258 om-06 (2006) Basic density and moisture content of pulpwood. In TAPPI Test Methods 2006. TAPPI Press, Technology Park, Atlanta
- Huntley SK, Ellis D, Gilbert M, Chapple C, Mansfield S (2003) Significant increases in pulping efficiency in C4H-F5H transformed poplars: improved chemical savings and reduced environmental toxins. *J Agr Food Chem* 51:6178–6183
- Robinson AR, Mansfield SD (2009) Rapid analysis of poplar lignin monomer composition by a streamlined thioacidolysis procedure and near-infrared reflectance-based prediction modeling. *Plant J* 58:706–714
- SAS Institute Inc. SAS 9.1.3 Help and documentation. In: Cary, NC: SAS Institute Inc., 2000–2004
- Armstrong J (1982) Polyploidy and wood anatomy of mature white ash, *Fraxinus americana*. *Wood Fiber Sci* 14:331–339
- Van Buijtenen JP, Joranson PN, Einspahr DW (1958) Diploid versus triploid aspen as pulpwood sources, with reference to growth, chemical, physical, and pulping differences. *TAPPI* 41:170–175
- Einspahr DW, van Buijtenen JP, Peckman JR (1963) Natural variation and heritability in triploid aspen. *Silvae Genet* 12
- Benson MK, Einspahr DW (1967) Early growth of diploid, triploid and triploid hybrid aspen. *Forest Sci* 13:150–155
- Hu B, Wang B, Wang C, Song W, Chen C (2012) Microarray analysis of gene expression in triploid black poplar. *Silvae Genet* 61:148–157
- Birchler JA, Veitia RA (2012) Gene balance hypothesis: connecting issues of dosage sensitivity across biological disciplines. *Proc Natl Acad Sci U S A* 109:14746–14753
- Yao H, Gray AD, Auger DL, Birchler JA (2013) Genomic dosage effects on heterosis in triploid maize. *Proc Natl Acad Sci U S A* 110: 2665–2669
- Novaes E, Kirst M, Winter-Sederoff H, Sederoff R (2010) Lignin and biomass: a negative correlation for wood formation and lignin content in trees. *Plant Physiol* 154:555–561
- Kirst M, Myburg AA, De Leon JPG, Kirst ME, Scott J, Sederoff R (2004) Coordinated genetic regulation of growth and lignin revealed by quantitative trait locus analysis of cDNA microarray data in an interspecific backcross of *Eucalyptus*. *Plant Physiol* 135:2368–2378

37. Shafizadeh F, Chin PPS (1977) Thermal deterioration of wood. In IS Goldstein, ed, Wood Technology: Chemical Aspects. American Chemical Society Symposium Series 43, pp 57–81
38. Demirbaş A (2001) Relationships between lignin contents and heating values of biomass. Energy Convers Manag 42:183–188
39. Serapiglia MJ, Humiston MC, Xu H, Hogsett DA, de Orduna RM, Stipanovic AJ, Smart LB (2013) Enzymatic saccharification of shrub willow genotypes with differing biomass composition for biofuel production. Front Plant Sci 4:57. doi:[10.3389/fpls.2013.00057](https://doi.org/10.3389/fpls.2013.00057)
40. Van Acker R, Leplé J-C, Aerts D, Storme V, Goeminne G, Ivens B, Légée F, Lapierre C, Piens K, Van Montagu MCE, Santoro N, Foster CE, Ralph J, Soetaert W, Pilate G, Boerjan W (2014) Improved saccharification and ethanol yield from field-grown transgenic poplar deficient in cinnamoyl-CoA reductase. Proc Natl Acad Sci U S A 111:845–850
41. Brereton NJB, Ray MJ, Shield I, Martin P, Karp A, Murphy RJ (2012) Reaction wood—a key cause of variation in cell wall recalcitrance in willow. Biotech Biofuel 5:83. doi:[10.1186/1754-6834-1185-1183](https://doi.org/10.1186/1754-6834-1185-1183)
42. 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
43. Mansfield SD, Kang K-Y, Chapple C (2012) Designed for deconstruction—poplar trees altered in cell wall lignification improve the efficacy of bioethanol production. New Phytol 194:91–101