

Article

## Comparison of Pyrolysis Mass Spectrometry and Near Infrared Spectroscopy for Genetic Analysis of Lignocellulose Chemical Composition in *Populus*

Jianxing Zhang <sup>1</sup>, Evandro Novaes <sup>2</sup>, Matias Kirst <sup>1,3</sup> and Gary F. Peter <sup>1,3,\*</sup>

<sup>1</sup> School of Forest Resources & Conservation, University of Florida, P.O. Box 110410, Gainesville, FL 32611-0410, USA; E-Mails: zhangjx04@ufl.edu (J.Z.); mkirst@ufl.edu (M.K.)

<sup>2</sup> Escola de Agronomia, Universidade Federal de Goiás, P.O. Box 131, Goiânia, GO 74690-900, Brazil; E-Mail: novaes@ufg.br

<sup>3</sup> Genetics Institute and Plant Molecular and Cellular Biology Program, University of Florida, P.O. Box 110410, Gainesville, FL 32611-0410, USA

\* Author to whom correspondence should be addressed; E-Mail: gfpeter@ufl.edu; Tel.: +1-352-846-0896; Fax: +1-352-846-1277.

Received: 15 January 2014; in revised form: 7 March 2014 / Accepted: 12 March 2014 /

Published: 21 March 2014

---

**Abstract:** Genetic analysis of wood chemical composition is often limited by the cost and throughput of direct analytical methods. The speed and low cost of Fourier transform near infrared (FT-NIR) overcomes many of these limitations, but it is an indirect method relying on calibration models that are typically developed and validated with small sample sets. In this study, we used >1500 young greenhouse grown trees from a clonally propagated single *Populus* family, grown at low and high nitrogen, and compared FT-NIR calibration sample sizes of 150, 250, 500 and 750 on calibration and prediction model statistics, and heritability estimates developed with pyrolysis molecular beam mass spectrometry (pyMBMS) wood chemical composition. As calibration sample size increased from 150 to 750, predictive model statistics improved slightly. Overall, stronger calibration and prediction statistics were obtained with lignin, S-lignin, S/G ratio, and  $m/z$  144 (an ion from cellulose), than with C5 and C6 carbohydrates, and  $m/z$  114 (an ion from xylan). Although small differences in model statistics were observed between the 250 and 500 sample calibration sets, when predicted values were used for calculating genetic control, the 500 sample set gave substantially more similar results to those obtained with the pyMBMS data. With the 500 sample calibration models, genetic correlations obtained with FT-NIR and

pyMBMS methods were similar. Quantitative trait loci (QTL) analysis with pyMBMS and FT-NIR predictions identified only three common loci for lignin traits. FT-NIR identified four QTLs that were not found with pyMBMS data, and these QTLs were for the less well predicted carbohydrate traits.

**Keywords:** FT-NIR; quantitative trait loci (QTL); *Populus*; pyMBMS; lignin; S-lignin; G-lignin; S/G ratio; C5; C6; clonal repeatability; genotypic correlations

---

## 1. Introduction

Forests trees capture greenhouse gases [1] and provide a renewable supply of wood for pulp, paper, construction and bioenergy. Genetic improvement of forest tree species, such as *Eucalyptus*, *Populus*, and *Pinus* species has increased growth, improved stem form and disease resistance [2], leading to significant gains in yield at shorter rotations [3,4]. However, despite the importance of wood chemical composition for yields of chemical pulp and biofuel and our knowledge of genes that code for enzymes involved in synthesis of cellulose, hemicellulose and lignin from model herbaceous and tree species, forest tree breeders have only recently begun dissecting the genetic architecture of these traits [5–11].

A variety of methods for measuring the chemical composition of lignocellulosic biomass have been developed and some have been applied to understand genetic control and environmental effects [12,13]. Although classical wet chemical methods have been used widely, their application for genetic analyses is limited by their high cost and low throughput. Composition data from miniaturized wet chemical methods were used to calculate genetic parameters for wood chemical composition in loblolly pine [14] and *Eucalyptus* [13,15]. Pyrolysis molecular beam mass spectrometry (pyMBMS) is faster than wet chemical methods and has been used to characterize lignocellulosic components in herbaceous and woody biomass [16–18], analyze genetic trials aimed at characterizing trait genetic architecture [19], map quantitative trait loci (QTL) [10,20] and identify genes that affect cellulose, hemicellulose and lignin content in pine and poplar [19,21]. Despite these advances in our knowledge of the genetic mechanisms controlling wood lignocellulose composition, analysis of the very large number of samples needed for traditional and advanced molecular marker based breeding in commercially important species requires faster and lower cost methods [22].

One such method is near infrared (NIR) spectroscopy, which is an indirect method that relates absorbance differences in NIR wavelengths to differences in anatomical, chemical and mechanical properties. NIR has been used for qualitative and quantitative chemical analyses in many fields and industrial applications, including agricultural products [23], food [24], and forestry [25–27]. NIR relies on multivariate models to predict properties of new samples, and has been used to estimate genetic parameters for wood chemical content in different tree species. For example, in *Eucalyptus globulus* [12,28–31] NIR was utilized for prediction of cellulose, pulp yield, lignin and extractives, and *Eucalyptus nitens* [32,33] for cellulose, and in *Pinus pinaster Ait* [34,35] for lignin, extractives, cellulose and monosaccharides. Nevertheless, no direct comparison of genetic parameter estimates of wood chemical composition obtained with NIR and a direct method have been reported. Nor have the

utility of NIR predictions for quantitative trait loci (QTL) mapping, association genetics, and genomic selection been investigated.

The goals of this research were to assess the importance of calibration sample size on FT-NIR model predictions and compare genetic parameter estimates and significant QTLs from the indirect FT-NIR predictions to pyMBMS within a single *Populus* family.

## 2. Materials and Methods

### 2.1. Samples and FT-NIR

For this research, wood samples from a single pseudo-backcross family of *Populus* were used. The experimental design, sample processing and collection of pyMBMS data are described in detail by Novaes *et al.* [10]. Briefly, ground samples of wood were available from 2376 plants from 396 genotypes grown under two different nitrogen treatments (with ~3 clonal replicates) and then harvested after 10 weeks. The basal 5 cm section of the debarked stem were dried and ground to 40 mesh. Pyrolysis MBMS data were available for wood samples from two replicates (1515), and 1505 of these samples were scanned with FT-NIR. The spectra were obtained with a Perkin-Elmer Spectrum 400 FTIR/FTNIR (PerkinElmer Ltd., Beaconsfield, UK) equipped with an X-Y stage autosampler to increase the efficiency of scanning. About 4 mg of powdered samples were loaded into wells of an X-Y plate (96 wells) autosampler with diffuse reflectance from Pike technologies (Pike Tech., Madison, WI, USA). Aliquots from each wood sample were loaded into three wells and each well was scanned 32 times, at a resolution of 8 cm<sup>-1</sup> (10,000–4000 cm<sup>-1</sup>). Spectra from the three wells were averaged prior to analysis. The Spectrum Quant+ software (PerkinElmer Ltd., Beaconsfield, UK) was used for calibration and prediction. The second derivatives of the NIR spectra were applied to adjust the baseline. In all calibrations with the software package, a partial least square (PLS) algorithm was used to develop the calibration and prediction models from FT-NIR spectra data. The 1505 samples were randomly separated into calibration and prediction sets (150 vs. 1355, 250 vs. 1255, 500 vs. 1005, and 750 vs. 755). Calibration models were also developed using full cross validation.

### 2.2. Calibration Statistics

The standard error of estimate (SEE) was used to determine how fit the calibration models were, and standard error of prediction (SEP) was used to determine how fit is the prediction performed by this calibration model [27], as described below:

$$SEE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (1)$$

where  $\hat{y}_i$  is the estimated value of sample  $i$  by the calibration model,  $y_i$  is the known value (from pyMBMS) of sample  $i$ ,  $n$  is the number of samples in the calibration model.

$$SEP = \sqrt{\frac{\sum_{i=1}^m (y_{pred,i} - y_{ref,i})^2}{m}} \quad (2)$$

where  $y_{pred,i}$  is the predicted value of sample  $i$  by the calibration model,  $y_{ref,i}$  is the known value (from pyMBMS) of sample  $i$ ,  $m$  is the number of samples in the prediction set.

The coefficient of determination ( $R^2$ ) was utilized to evaluate the calibration and prediction performances.

### 2.3. Statistical Analysis of Phenotypic Data

The distribution of residuals was checked by PROC INSIGHT (SAS Institute Inc. 9.2<sup>®</sup> 2002–2008, Cary, NC, USA). Data cleaning was conducted by removing outliers and recording errors as reported by Novaes *et al.* [10]. The mixed model used in this paper was the same as before [10]:

$$y_{ijklmno} = \mu + r_i + T_k + rt_{ik} + b_{j(i)} + c_l + rc_{il} + tc_{kl} + p_{m(i)} + q_{n(i)} + e_{ijklmno} \quad (3)$$

where  $y_{ijklmno}$  is the response of the  $o$ th ramet of the  $l$ th clone in the  $k$ th treatment of the  $j$ th bench within the  $i$ th replication;  $\mu$  is population mean;  $r_i$  is the random effect of replication, which is normally and independently distributed (NID) as  $N(0, \sigma_r^2)$ ;  $T_k$  is the treatment effect of nitrogen;  $rt_{ik}$  is the interaction of replication by treatment,  $\sim NID(0, \sigma_{rt}^2)$ ;  $b_{j(i)}$  is the random effect of bench (incomplete block) within replication,  $\sim NID(0, \sigma_{rbc}^2)$ ;  $c_l$  is the random effect of clone,  $\sim NID(0, \sigma_c^2)$ ;  $rc_{il}$  is the random effect of replication by clone interaction,  $\sim NID(0, \sigma_{rc}^2)$ ;  $tc_{kl}$  is the random effect of treatment by clone interaction,  $\sim NID(0, \sigma_{tc}^2)$ ;  $p_{m(i)}$  is the random effect of row within replication,  $\sim NID(0, \sigma_p^2)$ ;  $q_{n(i)}$  is the random effect of column with replication,  $\sim NID(0, \sigma_q^2)$ ;  $e_{ijklmno}$  is the random error effect within the experiment,

$$e_{ijklmno} \sim NID(0, R), \text{ and } R = \begin{pmatrix} \sigma_{e_1}^2 & 0 & 0 \\ 0 & \sigma_{e_2}^2 & 0 \\ 0 & 0 & \sigma_{e_3}^2 \end{pmatrix} \quad (4)$$

The treatment effect for each trait was obtained using the SAS<sup>®</sup> System for mixed models. Restricted maximum likelihood with ASReml was utilized to obtain the variance components and genetic parameters. For least-square means in QTL analysis, we took both clone effect and its interaction with treatment as fixed effects in the model.

ASReml was used to calculate the clonal repeatability for each trait in the univariate analysis with estimates of variance components as follows:

$$H^2 = \sigma_c^2 / \left[ \sigma_c^2 + \sigma_{rc}^2 + \sigma_{tc}^2 + \left( \sum_1^3 \sigma_e^2 / 3 \right) \right] \quad (5)$$

where  $\sigma_c$ ,  $\sigma_{rc}$ ,  $\sigma_{tc}$ ,  $\sigma_{e_1}^2$ ,  $\sigma_{e_2}^2$ , and  $\sigma_{e_3}^2$  were defined as previously. Pair-wise genetic correlations between wood chemical traits were estimated as described before [10].

### 2.4. QTL Analysis

A previously published genetic map [10,36] was utilized to test whether the FT-NIR wood chemical predictions could be used to detect genomic loci (QTL) controlling the quantitative traits. Briefly, the genetic map contains 163 microsatellite and 18 microarray-based markers covering all the 19 linkage groups of poplar, at an average density of one marker for every 16 cM (Kosambi's map function). The genetic map was constructed on MapMaker 3.0 [37] and the QTL analysis was performed with composite interval mapping based on maximum likelihood estimation [38] using Windows QTL Cartographer v.2.5 [39]. The presence of QTLs along the linkage groups is tested with a likelihood ratio

(LR) test. The LR compares the likelihood of having a QTL (full model) at any single position of the map against the likelihood of not having it (reduced model). The LR threshold for definition of a QTL was  $\alpha = 0.05$ , determined with genome-wide analysis of 1000 permutations [40]. For this article we only tested the performance of the FT-NIR prediction with the 500 sample set on the high N fertilization treatment.

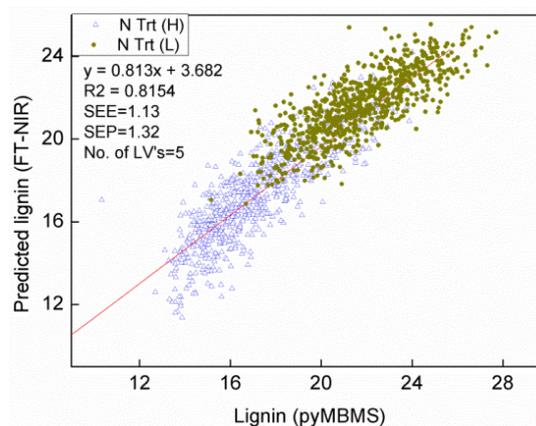
### 3. Results

#### 3.1. Effects of Sample Size on Calibration and Prediction

To investigate the impact of the number of samples needed for good calibration and prediction using pyMBMS data, we compared the correlation coefficients when using 150, 250, 500 and 750 samples in the calibration set (Table 1). All sets of samples were chosen randomly, and the remaining samples were used as a prediction set. The four calibration sets had similar means and ranges of chemical components (data not shown). In general, similar calibration  $R^2$  were obtained with all sample sets for lignin, G-lignin, and S-lignin (Table 1); however, the calibration  $R^2$  for C5, C6 and  $m/z$  114, an ion from xylan, were overall lower than for lignin and more variable across different sample sizes (Table 1). The prediction correlation coefficients strengthened slightly for all chemical components with increasing sample size. For example, lignin prediction increased from  $R^2 = 75.50\%$  (150) to  $R^2 = 82.39\%$  (750) (Table 1). The lignin prediction results were better than the sugar (C5, C6) components, except for  $m/z$  144, an ion arising from cellulose, which had a prediction  $R^2$  of 70.55% with the 500 calibration sample set.

We previously reported that when grown at high nitrogen, wood lignin content is significantly lower than when grown at low nitrogen [10]. The FT-NIR predicted lignin contents also differed significantly between low and high nitrogen environments, for calibration models developed with all sample sizes (Table 2). For the 500 sample calibration set, FT-NIR prediction has a slope of 0.813 relative to the pyMBMS; the predicted FT-NIR values were slightly higher at high nitrogen, and slightly lower at low nitrogen than the pyMBMS data (Figure 1) [41]. With both FT-NIR and pyMBMS all carbohydrate components did not differ between nitrogen levels. These results validate the ability of calibration models to detect relatively large differences in poplar wood lignin content induced by nitrogen availability.

**Figure 1.** Pyrolysis molecular beam mass spectrometry (pyMBMS) vs. Fourier transform near infrared (FT-NIR) predicted lignin for low and high nitrogen treatments.



**Table 1.** Different sample size calibration and prediction results.

	Trait	Factors	Calibration % $R^2$	SEE	SEP	Prediction % $R^2$
150 set	Lignin	3	85.00	1.29	1.55	75.50
	G-lignin	4	77.71	0.80	1.19	56.00
	S-lignin	3	87.53	0.92	1.10	72.25
	C5	4	67.63	1.39	2.31	29.91
	C6	2	50.49	3.04	3.48	37.85
	$m/z$ 114	4	71.2	0.28	0.44	49.37
	$m/z$ 144	3	76.67	0.22	0.28	62.93
250 set	Lignin	4	87.39	1.17	1.45	79.96
	G-lignin	4	76.16	0.79	1.02	59.94
	S-lignin	4	89.5	0.81	1.03	78.71
	C5	4	66.42	1.36	1.81	36.06
	C6	4	74.61	2.25	2.95	47.89
	$m/z$ 114	4	72.45	0.28	0.36	50.53
	$m/z$ 144	4	84.74	0.18	0.24	64.91
500 set	Lignin	5	86.68	1.13	1.32	81.54
	G-lignin	5	77.46	0.76	0.94	62.24
	S-lignin	4	81.87	0.98	1.08	80.14
	C5	4	56.27	1.50	1.74	36.89
	C6	4	64.42	2.47	2.83	49.45
	$m/z$ 114	4	64.92	0.30	0.34	52.37
	$m/z$ 144	5	80.86	0.18	0.22	70.55
750 set	Lignin	5	84.79	1.26	1.39	82.39
	G-lignin	7	78.21	0.76	0.95	69.23
	S-lignin	4	83.78	0.99	1.08	80.41
	C5	4	47.53	1.66	1.82	41.27
	C6	4	57.62	2.75	2.99	49.44
	$m/z$ 114	4	59.75	0.32	0.35	55.46
	$m/z$ 144	5	76.72	0.20	0.23	69.97

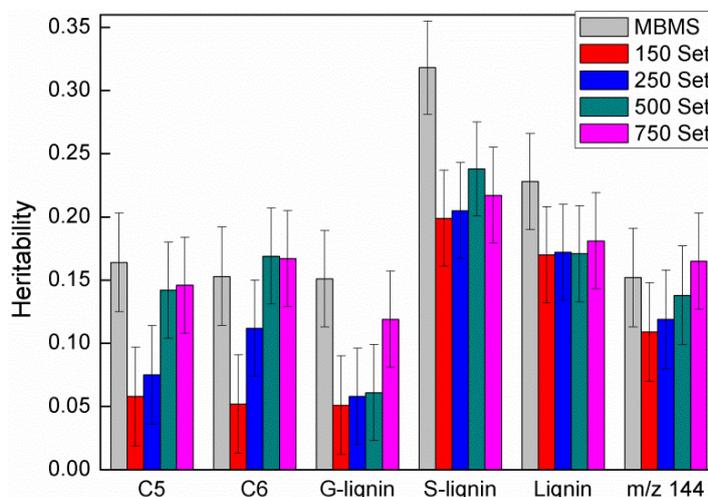
Note: 150, 250, 500 and 750 sets are randomly selected samples from all samples for calibrations. All pyMBMS chemical data were peak intensities. Lignin: (G-lignin + S-lignin + 120 + 152 + 180 + 181)  $\times$  correction factor, G-lignin: Guaiacyl lignin monomer  $m/z$  peaks sum of 124 + 137 + 138 + 150 + 164 + 178, S-lignin: Syringyl lignin  $m/z$  peaks sum of 154 + 167 + 168 + 182 + 194 + 208 + 210, C5: Five-carbon hemicelluloses sugar  $m/z$  peaks sum of 57 + 73 + 85 + 96 + 114, C6: Six-carbon cellulose sugar  $m/z$  peaks sum of 57 + 60 + 73 + 98 + 126 + 144,  $m/z$  114: peak 114 intensity associated with xylose(3-hydroxy-2-penteno-1,5-lacto,  $C_5H_6O_3$ ),  $m/z$  144: peak 144 intensity associated with glucose (methylbenzodioxole,  $C_8H_8O_2$ , due to levoglucosan) [16,42].

**Table 2.** Estimates of clonal repeatability and average trait value for seven wood chemistry phenotypes in two nitrogen treatments with pyMBMS and FT-NIR of 500 set.

Trait	$H^2 \pm SE$		N (L) Mean $\pm$ SE		N (H) Mean $\pm$ SE		Prob >  t	
	FT-NIR	pyMBMS	FT-NIR	pyMBMS	FT-NIR	pyMBMS	FT-NIR	pyMBMS
Lignin	0.17 $\pm$ 0.04	0.23 $\pm$ 0.04	21.6 $\pm$ 0.1	21.7 $\pm$ 0.1	17.5 $\pm$ 0.1	17.4 $\pm$ 0.1	<0.0001	<0.0001
G-lignin	0.06 $\pm$ 0.04	0.15 $\pm$ 0.04	13.2 $\pm$ 0.1	13.2 $\pm$ 0.1	11.2 $\pm$ 0.1	11.2 $\pm$ 0.1	<0.0001	<0.0001
S-lignin	0.24 $\pm$ 0.04	0.32 $\pm$ 0.04	13.0 $\pm$ 0.1	13.0 $\pm$ 0.1	10.2 $\pm$ 0.1	10.2 $\pm$ 0.1	<0.0001	<0.0001
C5	0.14 $\pm$ 0.04	0.16 $\pm$ 0.04	25.6 $\pm$ 0.1	25.7 $\pm$ 0.1	27.1 $\pm$ 0.1	27.0 $\pm$ 0.1	<0.0001	<0.0001
C6	0.17 $\pm$ 0.04	0.15 $\pm$ 0.04	32.7 $\pm$ 0.1	32.9 $\pm$ 0.1	35.7 $\pm$ 0.1	35.5 $\pm$ 0.2	<0.0001	<0.0001
<i>m/z</i> 144	0.14 $\pm$ 0.04	0.15 $\pm$ 0.04	0.89 $\pm$ 0.01	0.90 $\pm$ 0.01	1.35 $\pm$ 0.01	1.34 $\pm$ 0.02	<0.0001	<0.0001
<i>m/z</i> 114	0.04 $\pm$ 0.04	0.17 $\pm$ 0.03	1.80 $\pm$ 0.01	1.83 $\pm$ 0.01	2.47 $\pm$ 0.01	2.43 $\pm$ 0.02	<0.0001	<0.0001

Note: Clonal repeatability ( $H^2$ ), Standard error (SE), N (L): Low nitrogen treatment, N (H): High nitrogen treatment. The last column is the p-values for testing nitrogen treatment for each phenotype.

An important standard measure of genetic control of a phenotype is heritability, the ratio of genetic to phenotypic variation. The best measure of genetic control for clonally propagated populations is clonal repeatability, a measure of broad sense heritability or total genetic control. With the FT-NIR predicted values, the clonal repeatabilities were all lower than those obtained with the pyMBMS data (Figure 2). For C5, C6, *m/z* 144, and G-lignin, the FT-NIR heritability estimates increased with larger calibration sample size (Figure 2). For *m/z* 114, an ion from xylan, the heritability was zero for every calibration set, and thus was dropped from the genetic analyses described below. For this population a random set of 500 samples gave strong calibration models with the smallest standard errors of prediction and for all chemical components gave the highest estimates of genetic control for most traits. Thus, the 500 sample calibration model was used to predict the chemical composition of the 1505 wood samples for the genetic correlation and QTL analyses described below.

**Figure 2.** Clonal repeatability and standard errors for different FT-NIR size calibration sets compared with pyMBMS; Error bars correspond to standard errors of the mean.

### 3.2. Comparison of Heritability and Genetic Correlations among Traits

Compared with pyMBMS, the clonal repeatability estimates with FT-NIR predictions were about 25% lower for total lignin and S-lignin and 60% lower for G-lignin. For both pyMBMS and FT-NIR, the genetic control of lignin was higher than for carbohydrates [10]. However, the clonal repeatability was very similar for C5 and cellulose  $m/z$  144 ions, and a little better for C6 estimated with FT-NIR when compared with pyMBMS (Table 2). Interestingly, with FT-NIR the genetic control of  $m/z$  144, a cellulose ion, was almost as high as the sum of the C6 sugar ions [16].

Pair-wise genetic correlations between traits are important for understanding whether common genetic pathways are involved in the control of two or more traits and for applying the appropriate breeding and selection strategies. Genetic correlations of chemical traits within and across pyMBMS and FT-NIR were quite similar, with most of the pair-wise correlation estimates being stronger than 0.60 (absolute value) (Table 3). The strongest genetic correlations between pyMBMS and FT-NIR predictions were for lignin (1.00), S-lignin (0.99), G-lignin (0.74), S/G (0.92), C6 (0.87), C5 (0.73), C6/C5 (1.00),  $m/z$  144 (0.87), and C6/lignin (0.87). Genetic correlations among chemical components were similar between pyMBMS and FT-NIR data for most of the traits, except for 9 pairs where correlations differed by more than 0.2.

Pair-wise genetic correlations between pyMBMS and FT-NIR predictions indicate lignin, S-lignin, S/G, C6, C6/C5,  $m/z$  144 and C6/Lignin were highly correlated while G-lignin and C5 had weaker correlations (Table 4).

### 3.3. QTL Analysis

For all six wood chemistry traits under high nitrogen treatment, QTL analyses were performed using values from all 1505 samples predicted with the 500 sample calibration model. The objective was to compare the number of QTLs and whether the QTLs detected with FT-NIR co-localize with those detected by pyMBMS. Even though the level of genetic control estimated with FT-NIR data were considerably lower than those estimated with pyMBMS, seven QTLs were identified with the FT-NIR predictions and eight with pyMBMS (Table 5). However, only three of these QTLs mapped to the same intervals in both pyMBMS and FT-NIR (Table 5, Figure 3). These coincidences were detected for  $m/z$  144, lignin and G-lignin, and were all located on the same region of LGXIII. We previously identified this region as having a major effect on wood chemical and growth traits of this family [10]. More specifically, this region explains 56% of the heritable variation for cellulose to lignin ratio, as well as 20%–25% of the heritable variation for biomass.

**Table 3.** Pair-wise comparisons of genotypic correlations between chemical traits for both NIR and pyMBMS.

Trait	Lignin		S-lignin		G-lignin		S/G		C6		C5		C6/C5		<i>m/z</i> 144		C6/L	
	MS	NIR	MS	NIR	MS	NIR												
<b>Lignin</b>	—	<u><b>1.00</b></u>	<b>0.92</b>	<b>0.99</b>	<i>0.51</i>	<b>0.79</b>	<b>0.70</b>	<b>0.90</b>	<b>-0.99</b>	<b>-1.00</b>	<b>-0.81</b>	<b>-1.00</b>	<b>-1.00</b>	<b>-0.98</b>	<b>-1.00</b>	<b>-0.97</b>	<b>-1.00</b>	<b>-1.00</b>
<b>SL</b>	<b>0.92</b>	<b>0.99</b>	—	<u><b>0.99</b></u>	<i>0.14</i>	<b>0.70</b>	<b>0.92</b>	<b>0.96</b>	<b>-0.85</b>	<b>-1.00</b>	<b>-0.62</b>	<b>-0.99</b>	<b>-1.00</b>	<b>-0.97</b>	<b>-0.92</b>	<b>-1.00</b>	<b>-0.93</b>	<b>-1.00</b>
<b>GL</b>	<i>0.51</i>	<i>0.79</i>	<i>0.14</i>	<b>0.70</b>	—	<u><b>0.74</b></u>	<i>-0.26</i>	<i>0.45</i>	<b>-0.64</b>	<b>-0.61</b>	<b>-0.69</b>	<i>-0.58</i>	<i>-0.42</i>	<i>-0.67</i>	<i>-0.57</i>	<i>-0.58</i>	<b>-0.62</b>	<b>-0.71</b>
<b>S/G</b>	<b>0.70</b>	<b>0.90</b>	<b>0.92</b>	<b>0.96</b>	<i>-0.26</i>	<i>0.45</i>	—	<u><b>0.92</b></u>	<i>-0.54</i>	<b>-0.95</b>	<i>-0.29</i>	<b>-0.96</b>	<b>-0.86</b>	<b>-0.93</b>	<b>-0.64</b>	<b>-0.95</b>	<b>-0.64</b>	<b>-0.94</b>
<b>C6</b>	<b>-0.99</b>	<b>-1.00</b>	<b>-0.85</b>	<b>-1.00</b>	<b>-0.64</b>	<b>-0.61</b>	<i>-0.54</i>	<b>-0.95</b>	—	<u><b>0.87</b></u>	<b>0.96</b>	<b>1.00</b>	<b>0.88</b>	<b>1.00</b>	<b>0.98</b>	<b>1.00</b>	<b>0.98</b>	<b>1.00</b>
<b>C5</b>	<b>-0.81</b>	<b>-1.00</b>	<b>-0.62</b>	<b>-0.99</b>	<b>-0.69</b>	<i>-0.58</i>	<i>-0.29</i>	<b>-0.96</b>	<b>0.96</b>	<b>1.00</b>	—	<u><b>0.73</b></u>	<b>0.72</b>	<b>0.99</b>	<b>0.90</b>	<b>1.00</b>	<b>0.88</b>	<b>1.00</b>
<b>C6/C5</b>	<b>-1.00</b>	<b>-0.98</b>	<b>-1.00</b>	<b>-0.97</b>	<i>-0.42</i>	<i>-0.67</i>	<b>-0.86</b>	<b>-0.93</b>	<b>0.88</b>	<b>1.00</b>	<b>0.72</b>	<b>0.99</b>	—	<u><b>1.00</b></u>	<b>0.95</b>	<b>1.00</b>	<b>0.96</b>	<b>1.00</b>
<b><i>m/z</i> 144</b>	<b>-1.00</b>	<b>-0.97</b>	<b>-0.92</b>	<b>-1.00</b>	<i>-0.57</i>	<i>-0.58</i>	<b>-0.64</b>	<b>-0.95</b>	<b>0.98</b>	<b>1.00</b>	<b>0.90</b>	<b>1.00</b>	<b>0.95</b>	<b>1.00</b>	—	<u><b>0.87</b></u>	<b>0.99</b>	<b>0.98</b>
<b>C6/L</b>	<b>-1.00</b>	<b>-1.00</b>	<b>-0.93</b>	<b>-1.00</b>	<b>-0.62</b>	<b>-0.71</b>	<b>-0.64</b>	<b>-0.94</b>	<b>0.98</b>	<b>1.00</b>	<b>0.88</b>	<b>1.00</b>	<b>0.96</b>	<b>1.00</b>	<b>0.99</b>	<b>0.98</b>	—	<u><b>0.87</b></u>

Note: MS: pyMBMS, NIR: FT-NIR. Bold type means pair-wise traits with correlation > |0.60|, Underline correlations are used for ranking, and Italic correlation pairs are correlations differed by more than 0.2. C6/C5: ratio of C6/C5, S/G: ratio of S-lignin/G-lignin, C6/L: ratio of C6/lignin.

**Table 4.** Pair-wise estimates of genotypic correlations between pyMBMS and FT-NIR.

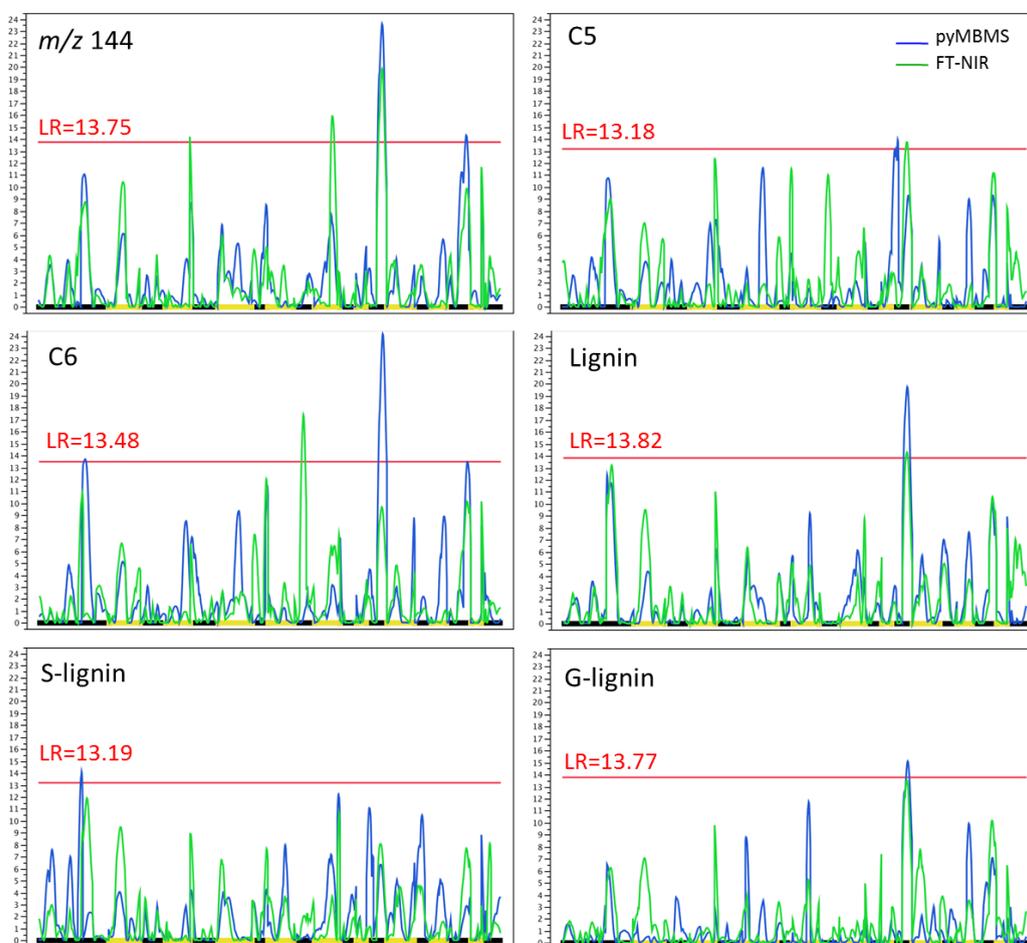
pyMBMS traits	Trait	FT-NIR predicted traits								
		Lignin	SL	GL	S/G	C6	C5	C6/C5	<i>m/z</i> 144	C6/L
<b>Lignin</b>	<b>Lignin</b>	<b>1.00</b>	<b>0.92</b>	<b>0.93</b>	<b>0.80</b>	<b>-0.96</b>	<b>-1.00</b>	<b>-0.89</b>	<b>-1.00</b>	<b>-1.00</b>
<b>SL</b>	<b>SL</b>	<b>1.00</b>	<b>0.99</b>	<b>0.73</b>	<b>0.95</b>	<b>-0.99</b>	<b>-1.00</b>	<b>-0.94</b>	<b>-1.00</b>	<b>-1.00</b>
<b>GL</b>	<b>GL</b>	<i>0.29</i>	<i>0.18</i>	<b>0.74</b>	<i>-0.27</i>	<i>-0.24</i>	<i>-0.28</i>	<i>-0.18</i>	<i>-0.27</i>	<i>-0.31</i>
<b>S/G</b>	<b>S/G</b>	<b>0.85</b>	<b>0.88</b>	<i>0.42</i>	<b>0.92</b>	<b>-0.84</b>	<b>-0.85</b>	<b>-0.82</b>	<b>-0.86</b>	<b>-0.89</b>
<b>C6</b>	<b>C6</b>	<b>-0.76</b>	<b>-0.74</b>	<b>-0.65</b>	<b>-0.63</b>	<b>0.87</b>	<b>0.96</b>	<b>0.74</b>	<b>0.94</b>	<b>0.74</b>
<b>C5</b>	<b>C5</b>	<i>-0.51</i>	<i>-0.49</i>	<i>-0.51</i>	<i>-0.39</i>	<i>0.50</i>	<b>0.73</b>	<i>0.48</i>	<i>0.51</i>	<i>0.48</i>
<b>C6/C5</b>	<b>C6/C5</b>	<b>-1.00</b>	<b>-1.00</b>	<b>-0.83</b>	<b>-0.93</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>
<b><i>m/z</i> 144</b>	<b><i>m/z</i> 144</b>	<b>-0.84</b>	<b>-0.83</b>	<b>-0.65</b>	<b>-0.74</b>	<b>0.86</b>	<b>0.86</b>	<b>0.86</b>	<b>0.87</b>	<b>0.83</b>
<b>C6/L</b>	<b>C6/L</b>	<b>-0.88</b>	<b>-0.84</b>	<b>-0.78</b>	<b>-0.72</b>	<b>0.86</b>	<b>0.88</b>	<b>0.83</b>	<b>0.86</b>	<b>0.87</b>

Note: Bold type means pair-wise traits with correlation > |0.60|. C6/C5: ratio of C6/C5, S/G: ratio of S-lignin/G-lignin, C6/L: ratio of C6/lignin.

**Table 5.** Number of quantitative trait loci (QTLs) detected with pyMBMS and FT-NIRS for six wood chemical traits. Coincidental QTLs map to the same intervals in the genome.

Trait	Number of QTLs		
	pyMBMS	FT-NIR	Coincident
<i>m/z</i> 144	2	3	1
C5	1	1	0
C6	2	1	0
Lignin	1	1	1
G-lignin	1	1	1
S-lignin	1	0	0
<b>Total</b>	<b>8</b>	<b>7</b>	<b>3</b>

**Figure 3.** QTL profiles for pyMBMS (blue) and FT-NIR (green) predictions for six wood chemistry traits. The FT-NIR predictions were conducted with the 500 samples calibration set. The likelihood ratio (LR) threshold is indicated in each figure (red line), the y-axis is the likelihood ratio and the black and yellow bars on the x-axis delimit each linkage group.



As expected, the QTL profiles of pyMBMS and FT-NIR tend to be more similar for traits that have stronger genotypic correlation between both estimates. For example, for lignin ( $r = 1.0$ ), C6 ( $r = 0.87$ ) and *m/z* 144 ( $r = 0.87$ ) the QTL profiles with pyMBMS and FT-NIR tend to co-vary (Figure 3). Generally for these traits, when a QTL is detected with one of the two techniques there is a peak on the QTL profile

of the other that may or may not be above significance threshold. Conversely, for C5 sugars, which have weak correlation between estimates obtained with pyMBMS and FT-NIR ( $r = 0.73$ ), the QTL profiles are quite different, even though the one significant QTL was detected with both methods.

#### 4. Discussion

The low cost and high throughput of FT-NIR offers the opportunity for efficiently phenotyping the thousands of wood samples needed for dissecting genetic and environmental control of wood lignocellulose composition. However, because FT-NIR is an indirect method, robust calibration models with high prediction precision need to be developed and broadly validated. In this study, pyMBMS chemical composition data and FT-NIR spectra were used to calibrate and predict the composition of 1505 poplar samples from a single family, grown under high and low nitrogen [10].

##### 4.1. Calibration, Prediction, and Sample Size

FT-NIR calibration models to estimate wood chemical composition have been reported for a number of species; however, most of these models were calibrated with wood chemistry data obtained from wet chemical methods using modest sample sizes and moderate calibration and prediction results were obtained [27,41,43–46]. Recently, global NIR models were developed with multiple pine species from multiple sites with high calibration and validation precision for lignin ( $R^2$  of calibration and validation 0.97 and 0.95 respectively) and cellulose ( $R^2$  of calibration and validation 0.84 and 0.72 respectively) content [47,48]. For *Eucalyptus*, more than 40 species across Australia (720 samples) were used to calibrate NIR models to predict Kraft pulp yield with  $R^2 = 0.91$  and low standard error of cross validation (1.36%) [49]. However, the utility of these global calibrations for genetic analyses has not been reported.

For indirect methods requiring calibration, an important question is how many samples are needed in the calibration set to develop strong predictive models that can be used for environmental and genetic analyses of wood chemistry. Although the  $R^2$  of calibration models were slightly weaker with 150 and 250 compared with the 500 and 750 sample sets (Table 1), the  $R^2$  of prediction models increased for all components with increasing sample size, with the strongest coefficients being obtained with 500 or 750 sample sets. Random sets of 500 samples yielded calibration models with  $R^2$  that ranged from 0.56 to 0.87 (Table 1) and were sufficient for developing good predictive models for all components. Overall, stronger calibration and prediction statistics were obtained with lignin than carbohydrates (Table 1), likely reflecting the quality of the estimation of lignin compared with the carbohydrate data from pyMBMS [11,50,51]. Reported calibration and prediction statistics using wet chemical methods were stronger than predictions obtained with pyMBMS [47]. One reason could be that the sum of peak intensities, arising from the breakdown of molecules during pyrolysis might be a mixture of several chemical components.

#### 4.2. Genetic Parameters and Environmental Control

A few reports have demonstrated that NIR predictions can be used to estimate genetic parameters. For *Pinus pinaster Ait* and *Pinus taeda*, Isik *et al.* [52], Perez *et al.* [34] and Gaspar *et al.* [35] showed NIR as a potential tool to estimate heritabilities and genetic correlations for tree selection in physical and chemical properties. For *Eucalyptus globulus* and *Eucalyptus nitens*, the application of NIR focused on predicting the genetic parameters for cellulose [28,33], pulp yield [12,30,31], lignin [29]. Schimleck *et al.* [15,32] investigated the cellulose content predicted with a NIR calibration model as an alternative to estimate Kraft pulp yield in *Eucalyptus nitens*. NIR analysis provided estimates of genetic parameters that were as good as direct cellulose assessment, demonstrating that the heritability and genetic gain could be estimated by NIR. However, these studies suffer from the lack of validation of the accuracy and reliability of the estimates from NIR predictions.

No comparison of direct and indirect estimates of wood chemical composition and of genetic parameter estimates has been reported previously. Heritability estimates with FT-NIRS were a little lower than the estimates with pyMBMS except for C6 (Table 2). Figure 2 demonstrates that estimates of genetic control were more similar between FT-NIR predictions and pyMBMS data when larger 500 and 750 sample sizes were used for NIR calibration. This suggests that for genetic analyses, using more samples in the calibration is important, even though the statistics of calibration and prediction were similar between the 250 and 500 sample sets. Consequently, with pyMBMS data, sample sizes of 500 or more for NIR calibration should be used when the goal is to partition variances and estimate heritability.

Pair-wise genetic correlations obtained with FT-NIR predicted chemical traits were close to those obtained with the direct pyMBMS method. For example, predicted lignin was highly correlated with predicted S-lignin (0.99), G-lignin (0.79), S/G (0.90), C6 (−1.00), C5 (−1.00), C6/C5 (−0.98), *m/z* 144 (−0.97) and C6/lignin (−1.00) (Table 3), where pyMBMS lignin performed in a similar way. We also investigated the genetic correlation for chemical components between pyMBMS data and FT-NIR predictions, and they had the similar correlation estimates for most of the chemical traits, except for 17 pairs where correlations differed by more than 0.2. Consequently, the results proved FT-NIR can provide a robust model for genetic correlations [13].

Previously, Novaes *et al.* [10] observed that nitrogen application during early growth of *Populus* significantly increased the content of C5 (hemicellulose) and C6 (cellulose), and decreased lignin content. With FT-NIR predictions, nitrogen fertilization was also highly significant for all chemical traits, with the C6 increase being larger than the C5 increase, and total lignin decrease was much more than S-lignin and G-lignin (Table 2). This shows that the nitrogen effect could be of major importance for enhancing cellulose content in wood.

#### 4.3. QTL Mapping

In total 7 QTLs identified with FT-NIR and 8 with py-MBMS, but only three were colocalized to the same interval. With the pyMBMS data, one QTL was identified for S-lignin but none were identified with FT-NIR. This relatively low coincidence of QTLs between the two methods requires additional investigation to understand whether the seven QTLs detected only with FT-NIR are real, even though a the stringent ( $\alpha = 0.05$ , with 1000 genome-wide permutation tests) QTL threshold used

gives statistical strength to the hypothesis that the novel FT-NIR QTLs are real. However, it is important to note the low prediction  $R^2$  of C6 (49%) and C5 (52%), for which two QTLs were identified only with FT-NIR predictions. Because coincident QTLs for lignin and  $m/z$  144 were detected with FT-NIR predictions and pyMBMS data, this suggests that FT-NIR can be used when calibration models yield good predictions ( $R^2 > 0.80$ ). Once a good calibration model is obtained, FT-NIR is less costly and faster than direct methods, such as wet chemistry. FT-NIR may be especially interesting for genetic analyses of wood chemistry, given the fact that these studies, such as QTL and association studies, require the phenotyping of thousands of individuals.

## 5. Conclusions

These results show that FT-NIR, coupled with pyMBMS for model calibration, is an appropriate high-throughput method for wood chemical calibration. Our results show good and moderate calibration and prediction results for lignin and carbohydrates, respectively. Our study demonstrates the similarity of the direct and indirect estimates of wood chemical composition and genetic parameter estimates. Strong genetic correlations are obtained between the FT-NIR and pyMBMS. However, QTLs detected by FT-NIR and pyMBMS are quite different.

## Acknowledgments

We thank Dudley A. Huber and Patricio R. Munoz for advice on the analysis of the data with ASReml. This research is supported by the Florida Energy Systems Consortium.

## Conflicts of Interest

The authors declare no conflict of interest.

## References

1. McCarl, B.A.; Schneider, U.A. Climate change-Greenhouse gas mitigation in us agriculture and forestry. *Science* **2001**, *294*, 2481–2482.
2. McKeand, S.; Mullin, T.; Byram, T.; White, T. Deployment of genetically improved loblolly and slash pines in the south. *J. For.* **2003**, *101*, 32–37.
3. Fox, T.R.; Jokela, E.J.; Allen, H.L. The development of pine plantation silviculture in the southern united states. *J. For.* **2007**, *105*, 337–347.
4. Li, X.B.; Huber, D.A.; Powell, G.L.; White, T.L.; Peter, G.F. Breeding for improved growth and juvenile corewood stiffness in slash pine. *Can. J. For. Res.* **2007**, *37*, 1886–1893.
5. Greaves, B.L.; Borralho, N.M.; Raymond, C.A. Breeding objective for plantation eucalypts grown for production of kraft pulp. *For. Sci.* **1997**, *43*, 465–472.
6. Peter, G.F.; White, D.E.; Torre, R.D.L.; Singh, R. The value of forest biotechnology: A cost modelling study with loblolly pine and kraft linerboard in the southeastern USA. *Int. J. Biotechnol.* **2007**, *9*, 415–435.

7. Kien, N.D.; Quang, T.H.; Jansson, G.; Harwood, C.; Clapham, D.; von Arnold, S. Cellulose content as a selection trait in breeding for kraft pulp yield in *Eucalyptus urophylla*. *Ann. For. Sci.* **2009**, *66*, 1–8.
8. Quang, T.H.; Kien, N.D.; von Arnold, S.; Jansson, G.; Thinh, H.H.; Clapham, D. Relationship of wood composition to growth traits of selected open-pollinated families of *Eucalyptus urophylla* from a progeny trial in Vietnam. *New For.* **2010**, *39*, 301–312.
9. Zhang, D.; Zhang, Z.; Yang, K. 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. For. Res.* **2006**, *36*, 2015–2023.
10. Novaes, E.; Osorio, L.; Drost, D.R.; Miles, B.L.; Boaventura-Novaes, C.R.; Benedict, C.; Dervinis, C.; Yu, Q.; Sykes, R.; Davis, M. Quantitative genetic analysis of biomass and wood chemistry of populus under different nitrogen levels. *New Phytol.* **2009**, *182*, 878–890.
11. Li, X. Breeding for Improved Growth, Wood Quality, and Chemistry for Southern Pines by Combining Quantitative Genetics and Association Mapping. Ph.D. Thesis, University of Florida, Gainesville, FL, USA, 15 December 2009.
12. Raymond, C.A.; Schimleck, L.R.; Muneri, A.; Michell, A.J. Genetic parameters and genotype-by-environment interactions for pulp-yield predicted using near infrared reflectance analysis and pulp productivity in *Eucalyptus globulus*. *Int. J. For. Genet.* **2001**, *8*, 213–224.
13. Schimleck, L.R. Near infrared spectroscopy: A rapid, non-destructive method for measuring wood properties and its application to tree breeding. *N. Z. J. For. Sci.* **2008**, *38*, 14–35.
14. Sykes, R.; Li, B.; Isik, F.; Kadla, J.; Chang, H.M. Genetic variation and genotype by environment interactions of juvenile wood chemical properties in *Pinus taeda* L. *Ann. For. Sci.* **2006**, *63*, 897–904.
15. Kube, P.D.; Raymond, C.A.; Banham, P.W. Genetic parameters for diameter, basic density, cellulose content and fibre properties for *Eucalyptus nitens*. *For. Genet.* **2001**, *8*, 285–294.
16. Evans, R.J.; Milne, T.A. Molecular characterization of the pyrolysis of biomass. 1. Fundamentals. *Energy Fuels* **1987**, *1*, 123–137.
17. Evans, R.J.; Milne, T.A. Molecular characterization of the pyrolysis of biomass. 2. Applications. *Energy Fuels* **1987**, *1*, 311–319.
18. Agblevor, F.A.; Evans, R.J.; Johnson, K.D. Molecular-beam mass-spectrometric analysis of lignocellulosic materials: I. Herbaceous biomass. *J. Anal. Appl. Pyrolysis* **1994**, *30*, 125–144.
19. Voelker, S.L.; Lachenbruch, B.; Meinzer, F.C.; Jourdes, M.; Ki, C.; Patten, A.M.; Davin, L.B.; Lewis, N.G.; Tuskan, G.A.; Gunter, L. Antisense down-regulation of 4CL expression alters lignification, tree growth, and saccharification potential of field-grown poplar. *Plant Physiol.* **2010**, *154*, 874–886.
20. Sewell, M.M.; Davis, M.F.; Tuskan, G.A.; Wheeler, N.C.; Elam, C.C.; Bassoni, D.L.; Neale, D.B. Identification of qtls influencing wood property traits in loblolly pine (*Pinus taeda* L.). II. Chemical wood properties. *Theor. Appl. Genet.* **2002**, *104*, 214–222.
21. Tuskan, G.; West, D.; Bradshaw, H.D.; Neale, D.; Sewell, M.; Wheeler, N.; Megraw, B.; Jech, K.; Wiselogle, A.; Evans, R. Two high-throughput techniques for determining wood properties as part of a molecular genetics analysis of hybrid poplar and loblolly pine. *Appl. Biochem. Biotechnol.* **1999**, *77*, 55–65.

22. Jannink, J.L.; Lorenz, A.J.; Iwata, H. Genomic selection in plant breeding: From theory to practice. *Brief. Funct. Gen.* **2010**, *9*, 166–177.
23. Nicolai, B.M.; Beullens, K.; Bobelyn, E.; Peirs, A.; Saeys, W.; Theron, K.I.; Lammertyn, J. Nondestructive measurement of fruit and vegetable quality by means of nir spectroscopy: A review. *Postharvest Biol. Technol.* **2007**, *46*, 99–118.
24. Alishahi, A.; Farahmand, H.; Prieto, N.; Cozzolino, D. Identification of transgenic foods using NIR spectroscopy: A review. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2010**, *75*, 1–7.
25. Tsuchikawa, S. A review of recent near infrared research for wood and paper. *Appl. Spectrosc. Rev.* **2007**, *42*, 43–71.
26. So, C.L.; Via, B.K.; Groom, L.H.; Schimleck, L.R.; Shupe, T.F.; Kelley, S.S.; Rials, T.G. Near infrared spectroscopy in the forest products industry. *For. Prod. J.* **2004**, *54*, 6–16.
27. Jones, P.D.; Schimleck, L.R.; Peter, G.F.; Daniels, R.F.; Clark, A. Nondestructive estimation of wood chemical composition of sections of radial wood strips by diffuse reflectance near infrared spectroscopy. *Wood Sci. Technol.* **2006**, *40*, 709–720.
28. Raymond, C.A.; Schimleck, L.R. Development of near infrared reflectance analysis calibrations for estimating genetic parameters for cellulose content in *Eucalyptus globulus*. *Can. J. For. Res.* **2002**, *32*, 170–176.
29. Poke, F.S.; Potts, B.M.; Vaillancourt, R.E.; Raymond, C.A. Genetic parameters for lignin, extractives and decay in *Eucalyptus globulus*. *Ann. For. Sci.* **2006**, *63*, 813–821.
30. Silva, J.C.; Borralho, N.M.G.; Araujo, J.A.; Vaillancourt, R.E.; Potts, B.M. Genetic parameters for growth, wood density and pulp yield in *Eucalyptus globulus*. *Tree Genet. Gen.* **2009**, *5*, 291–305.
31. Stackpole, D.J.; Vaillancourt, R.E.; Downes, G.M.; Harwood, C.E.; Brad, P.M. Genetic control of kraft pulp yield in *Eucalyptus globulus*. *Can. J. For. Res.* **2010**, *40*, 917–927.
32. Schimleck, L.R.; Kube, P.D.; Raymond, C.A. Genetic improvement of kraft pulp yield in *Eucalyptus nitens* using cellulose content determined by near infrared spectroscopy. *Can. J. For. Res.* **2004**, *34*, 2363–2370.
33. Hamilton, M.G.; Raymond, C.A.; Harwood, C.E.; Potts, B.M. Genetic variation in *Eucalyptus nitens* pulpwood and wood shrinkage traits. *Tree Genet. Gen.* **2009**, *5*, 307–316.
34. Perez, D.d.S.; Guillemain, A.; Alazard, P.; Plomion, C.; Rozenberg, P.; Carlos Rodrigues, J.; Alves, A.; Chantre, G. Improvement of *Pinus pinaster* ait elite trees selection by combining near infrared spectroscopy and genetic tools. *Holzforschung* **2007**, *61*, 611–622.
35. Gaspar, M.J.; Alves, A.; Louzada, J.L.; Morais, J.; Santos, A.; Fernandes, C.; Almeida, M.H.; Rodrigues, J.C. Genetic variation of chemical and mechanical traits of maritime pine (*Pinus pinaster* aiton). Correlations with wood density components. *Ann. For. Sci.* **2011**, *68*, 255–265.
36. Drost, D.R.; Novaes, E.; Boaventura-Novaes, C.; Benedict, C.I.; Brown, R.S.; Yin, T.; Tuskan, G.A.; Kirst, M. A microarray-based genotyping and genetic mapping approach for highly heterozygous outcrossing species enables localization of a large fraction of the unassembled *Populus trichocarpa* genome sequence. *Plant J.* **2009**, *58*, 1054–1067.
37. Lander, E.S.; Green, P.; Abrahamson, J.; Barlow, A.; Daly, M.J.; Lincoln, S.E.; Newburg, L. Mapmaker: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* **1987**, *1*, 174–181.

38. Zeng, Z.-B. Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. *Proc. Natl. Acad. Sci.* **1993**, *90*, 10972–10976.
39. Wang, S.; Basten, C.; Zeng, Z. *Windows QTL Cartographer 2.5*; Department of Statistics, North Carolina State University: Raleigh, NC, USA, 2007.
40. Churchill, G.A.; Doerge, R.W. Empirical threshold values for quantitative trait mapping. *Genetics* **1994**, *138*, 963–971.
41. Poke, F.S.; Raymond, C.A. Predicting extractives, lignin, and cellulose contents using near infrared spectroscopy on solid wood in eucalyptus globulus. *J. Wood Chem. Technol.* **2006**, *26*, 187–199.
42. Sykes, R.; Yung, M.; Novaes, E.; Kirst, M.; Peter, G.; Davis, M. High-throughput screening of plant cell-wall composition using pyrolysis molecular beam mass spectroscopy. In *Biofuels: Methods and Protocols*; Mielenz, J.R., Ed.; Humana Press: New York, NY, USA, 2009; Volume 581, pp. 169–183.
43. Hou, S.; Li, L. Rapid characterization of woody biomass digestibility and chemical composition using near-infrared spectroscopy. *J. Integr. Plant Biol.* **2011**, *53*, 166–175.
44. Kelley, S.S.; Rials, T.G.; Snell, R.; Groom, L.H.; Sluiter, A. Use of near infrared spectroscopy to measure the chemical and mechanical properties of solid wood. *Wood Sci. Technol.* **2004**, *38*, 257–276.
45. Nkansah, K.; Dawson-Andoh, B.; Slahor, J. Rapid characterization of biomass using near infrared spectroscopy coupled with multivariate data analysis: Part 1 yellow-poplar (*Liriodendron tulipifera* L.). *Bioresour. Technol.* **2010**, *101*, 4570–4576.
46. Wright, J.A.; Birkett, M.D.; Gambino, M.J.T. Prediction of pulp yield and cellulose content from wood samples using near-infrared reflectance spectroscopy. *Tappi J.* **1990**, *73*, 164–166.
47. Hodge, G.R.; Woodbridge, W.C. Global near infrared models to predict lignin and cellulose content of pine wood. *J. Near Infrared Spectrosc.* **2010**, *18*, 367–380.
48. Hodge, G.R.; Woodbridge, W.C. Use of near infrared spectroscopy to predict lignin content in tropical and sub-tropical pines. *J. Near Infrared Spectrosc.* **2004**, *12*, 381–390.
49. Downes, G.M.; Meder, R.; Hicks, C.; Ebdon, N. Developing and evaluating a multisite and multispecies NIR calibration for the prediction of kraft pulp yield in eucalypts. *South. For.* **2009**, *71*, 155–164.
50. Sykes, R.; Kodrzycki, B.; Tuskan, G.; Foutz, K.; Davis, M. Within tree variability of lignin composition in populus. *Wood Sci. Technol.* **2008**, *42*, 649–661.
51. Davis, M.; Elam, C.; Wiselogel, A.; Evans, R.; Tuskan, G.; West, D.; Megraw, R.; Wheeler, N.; Jech, K.; Sewell, M. Application of Pyrolysis Molecular Beam Mass Spectrometry for the Determination of Loblolly Pine and Hybrid Poplar Cell Wall Composition; TAPPI Pulping: Atlanta, Georgia, 1999; pp. 1077–1082.
52. Isik, F.; Mora, C.R.; Schimleck, L.R. Genetic variation in pinus taeda wood properties predicted using non-destructive techniques. *Ann. For. Sci.* **2011**, *68*, 283–293.