the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

The Plant Journal (2009) **58**, 706–714 doi: 10.1111/j.1365-313X.2009.03808.x

TECHNICAL ADVANCE

Rapid analysis of poplar lignin monomer composition by a streamlined thioacidolysis procedure and near-infrared reflectance-based prediction modeling

Andrew R. Robinson and Shawn D. Mansfield*

Department of Wood Science, University of British Columbia, 4030-2424 Main Mall, Vancouver, BC, V6T 1Z4, Canada

Received 7 November 2008; revised 6 January 2009; accepted 14 January 2009; published online 26 February 2009.
*For correspondence (fax 1604 822 9104; e-mail shawnman@interchange.ubc.ca).

SUMMARY

Determination of the physico-chemical attributes of plant cell walls, such as lignin content and composition, is of paramount importance in germplasm screening and for evaluating the results of plant breeding and genetic engineering. There are escalating needs for analyses to be robust, reproducible, accurate, and efficient. We have recently modified an established protocol for discrimination of lignin monomers, thioacidolysis, with the goal of increasing sample throughput while maintaining accuracy and reducing equipment load and consumption of reagents. Numerous methodological changes related to volume scaling, selection of the processing vessel, and sample handling were addressed. The revised protocol permitted rapid processing of some 50 or more samples per person per day. A direct comparison between methods using hybrid poplar (*Populus alba* × *tremula*) wood samples, resulted in quantities of *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) lignin monomers that were equivalent to those derived from the original protocol. The revised methodology was then applied to quickly generate phenotypic trait data from 267 hybrid poplar trees (including wild type and eight *C4H::F5H* transgenic lines), for the development of a near-infrared-based model for predicting the proportion of lignin monomers across a broad phenotypic range of S:G. The resulting partial least squares regression model performed well under full cross-validation, giving strong, linear relationships between actual and predicted monomer proportions, and very high predictive accuracy for the predominant G and S monomers. This research brings considerable refinement to the thioacidolysis procedure, and establishes a method for rapidly and accurately quantifying cell-wall lignin composition that could effectively be employed in routine phenotypic screening platforms.

Keywords: thioacidolysis, lignin, syringyl, guaiacyl, *p*-hydroxyphenyl, near-infrared reflectance.

INTRODUCTION

The lignin heteropolymer is an integral cell-wall constituent that significantly influences the physical properties of plants via its involvement in architectural support, water transport, and defense. Lignins comprise the second most abundant polymer class in the biosphere, and their biosynthesis renders these polymers among the more complex biomacromolecules synthesized by plants. This intricate macromolecule is assembled via the combinatorial free radical coupling of monolignol precursors derived from three *p*-hydroxycinnamyl alcohols varying in their degree of methoxylation (Ralph *et al.*, 2004), resulting in varying proportions of *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) lignin subunits according to plant species, tissue type, and response to environmental stress (Campbell and Sederoff, 1996). The effective use of plants for a range of natural and industrial purposes is largely dependent on the extent to which the plant cell wall is lignified. This is particularly true when considering biomass conversion for bioenergy, forage digestibility in ruminants, or processes such as pulping and papermaking from woody materials. The effective removal

of lignin is expensive and often the limiting factor in many applications. Lignin is therefore a key plant breeding or genetic engineering target to provide improvements in cell-wall conversion. Consequently, analytical techniques that permit the precise determination of the abundance and chemical attributes of lignin have become important tools in the analysis of plant cell walls. Advances in lignin analysis, in terms of both the efficiency and accuracy of determination, are now necessary due to the demand for medium- to high-throughput phenotypic screens to identify valuable plant germplasm for specific applications (e.g. bioenergy crops), especially in the case when the individuals originate from large breeding studies, mutant populations, or mapping families.

One important aspect of lignin composition that can affect the utility of the plant cell wall is the proportional content resulting from the different monomeric units. For example, in chemical pulping, the lignin monomer ratio has been shown to have a significant impact on the efficiency of delignification (both pulp yield and residual lignin content) and the bleachability of pulp (Huntley *et al.*, 2003; Stewart *et al.*, 2006). An established technique for determining monomeric composition is the thioacidolysis reaction. In this approach, the uncondensed arylglycerol-β-aryl ether-linked *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) monomers of lignin may be cleaved from the polymer and the abundances determined (Rolando *et al.*, 1992). Although proven and valuable, the established protocol for this analysis is time-consuming and requires large amounts of reagents and equipment, especially where the processing of large sample batches is concerned; it is therefore not suitable for high-throughput screens.

In contrast, several spectroscopic techniques have been employed to analyze/screen plant cell walls, including near-infrared (NIR) reflectance (Blakeney and Flinn, 2005; Huang *et al.*, 2007; Via *et al.*, 2007) and Fourier transform infrared (FT-IR) spectroscopies (Chen *et al.*, 1998; Carpita and McCann, 2002). Near-infrared reflectance spectroscopy uses infrared overtones and combination vibrations, while FT-IR employs mid-infrared regions based primarily on functional and fundamental vibrations. Near-infrared reflectance spectroscopy is a promising technique for the rapid determination of the physical and chemical properties of wood based on calibration and estimation rather than direct measurement. Although, the NIR signatures are often overlapping and make direct structural assignment difficult, there are several good examples of NIR reflectance-based modeling in the determination of lignin composition, in terms of both monomer ratio and total lignin content (Takayama *et al.*, 1997; Bailleres *et al.*, 2002; Yeh *et al.*, 2004; Sykes *et al.*, 2005; Alves *et al.*, 2006; Jones *et al.*, 2006; Poke and Raymond, 2006; Yamada *et al.*, 2006; Huang *et al.*, 2007; Li *et al.*, 2007; Maranan and Laborie, 2008). However, this trait prediction process still first requires the measurement of NIR spectra and compositional trait(s) (e.g. the determination of H, G, and S monomer composition of lignin by wet chemical techniques such as thioacidolysis or nitrobenzene oxidation) in a 'calibration' population that covers the phenotypic range of interest, in order for a predictive model to be generated. Furthermore, distinct prediction models are typically required for each plant species, or at least for each genus of interest. However, once an accurate model has been established, the easily obtained NIR data are all that will be required to allow the estimation of compositional traits in samples with undetermined characteristics.

The research herein describes a modified thioacidolysis protocol that enables the medium- to high-throughput screening of plant cell-wall lignin monomer composition. We describe the details of a series of simple revisions to the original thioacidolysis protocol that reduce its resource demands while permitting a marked increase in sample throughput. Furthermore, we utilize the described wet-chemical procedure to rapidly develop an accurate NIR reflectance-based prediction model for lignin monomer composition in hybrid poplar.

RESULTS

Thioacidolysis reaction

An initial experiment comparing the performance of the original method (Lapierre *et al.*, 1985; Rolando *et al.*, 1992) and our revised thioacidolysis method was conducted to determine the relative efficacy of quantification of *p*-hydroxyphenyl, guaiacyl, and syringyl lignin monomer subunits in woody plant tissue. In order to estimate the monomer proportions and total monomer yield (as determined by both methods) ground, extract-free wood samples from single trees of three lines of hybrid poplar P717 (*Populus alba* × *tremula*) were analysed, with five technical replicates conducted on each sample. These trees represent two transgenic lines, both of which exhibit markedly altered wood lignin chemistry, and the wild-type control. The transgenic lines included an over-expressing *C4H::F5H* line, which exhibited a considerable increase in the S:G ratio of lignin, without significant effect on *p*-hydroxyphenyl monomers or the relative lignin content of the wood, as previously determined by thioacidolysis (Meyer *et al.*, 1996, 1998; Franke *et al.*, 2000; Huntley *et al.*, 2003), and an RNAi-suppressed line targeting *p-coumaroyl-CoA 3-hydroxylase* (*C3′H*), with substantially reduced levels of cell-wall lignification and elevated levels of *p*-hydroxyphenyl units (Coleman *et al.*, 2008).

The intention was to detect, measure, and combine the contributions from all of the three major lignin monomer subunits liberated from the cell-wall lignin, where H monomers represent 4-hydroxyphenyl, G monomers represent 4-hydroxy-3-methoxyphenyl, and S monomers represent

4-hydroxy-3,5-dimethoxyphenyl subunits. The monomers were detected as both the *erythro*- and *threo*-isomers of (H, G, or S)-CHR-CHR-CH_2-R and (H, G, or S)-CHR-CHR-CHR_2, as per Rolando *et al.* (1992). In practice, clean, discrete measurements of each form were only possible for the G- and S-derived monomers, as the H-derived monomers were at trace levels in most samples, and the peaks frequently did not resolve well from one another in gas chromatography. Approximate elution times under the conditions described herein were: 22.2–23.5 min for H-derived monomers, 23.8–24.3 min for G-derived monomers, and 25.8–26.4 min for S-derived monomers (Figure 1).

Lignin subunit ratios and total monomer yield differed between the lines, with the wild type exhibiting an S:G ratio of ~70:30 as is typical of many angiosperm deciduous species, the over-expressing *C4H::F5H* transgenic line exhibiting an elevated S:G ratio (~84:15), and the *C3′H:RNAi* transgenic line exhibiting a decrease in total monomer yield of ~40% (compared to wild-type), with a concurrent increase (~10%) in H monomers largely at the expense of G monomers (Table 1).

It is apparent from Figure 2 and Table 1 that the averaged composition of lignin monomers, liberated as a result of thioacidolysis, was very similar when comparing the

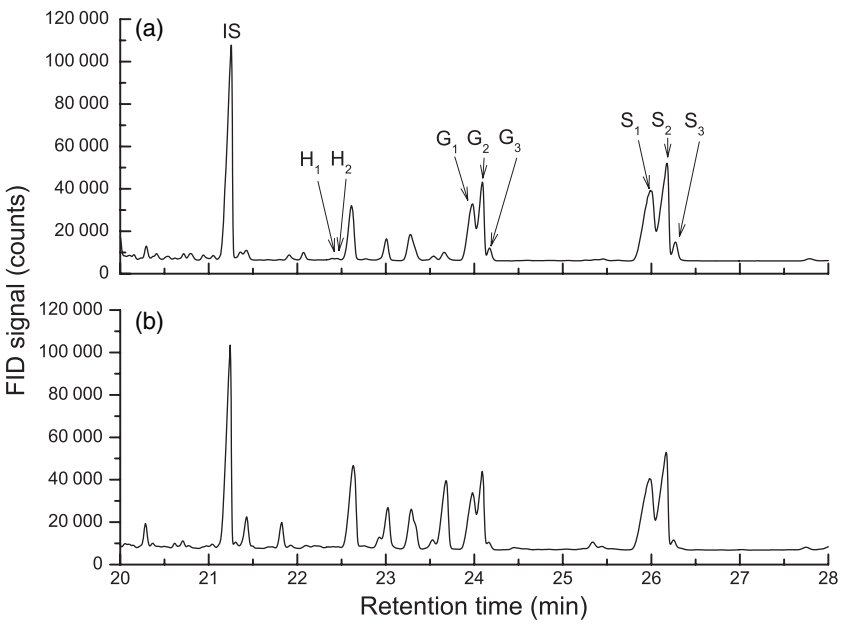

Figure 1. Typical gas chromatography flame ionization detector (FID) traces for wild-type P717 poplar thioacidolysis products.
Products prepared by (a) original, large-scale method, and (b) revised, small-scale method.
H_1, *erythro-p*-hydroxyphenyl-CHR-CHR-CH_2R; H_2, *threo-p*-hydroxyphenyl-CHR-CHR-CH_2R; G_1, *erythro*-Guaiacyl-CHR-CHR-CH_2R; G_2, *threo*-guaiacyl-CHR-CHR-CH_2R; G_3, guaiacyl-CH_2-CHR-CHR_2; S_1, *erythro*-syringly-CHR-CHR-CH_2R; S_2, *threo*-syringyl-CHR-CHR-CH_2R; S_3, syringyl-CH_2-CHR-CHR_2.

Table 1 Lignin yield and proportion measured by original and revised thioacidolysis reaction procedures

Line/ treatment[a]	Thiolignin yield (μmol g^{-1} Klason lignin)		Thiolignin composition (proportion)					
			Mean			SD		
	Mean	SD	H	G	S	H	G	S
P7-WT-R	1798.30	19.32	0.0037	0.3128	0.6835	0.0002	0.0034	0.0033
P7-WT-O	1941.93	89.14	0.0031	0.3161	0.6808	0.0002	0.0023	0.0024
P7-CF-R	2013.52	108.50	0.0039	0.1490	0.8471	0.0006	0.0015	0.0013
P7-CF-O	2129.42	326.68	0.0043	0.1527	0.8430	0.0009	0.0007	0.0004
P7-CR-R	1574.62	54.67	0.1212	0.2355	0.6433	0.0038	0.0049	0.0023
P7-CR-O	1649.18	254.84	0.1053	0.2413	0.6534	0.0035	0.0018	0.0033

[a]Line/treatment descriptors: P7 = P717 hybrid poplar (*Populus alba* × *tremula*), WT = wild type, CF = *C4H::F5H*, CR = *C3′H:RNAi*, and suffixes 'O' and 'R' indicate the use of 'original' or 'revised' protocol.
H, G, and S = *p*-hydroxyphenyl, guaiacyl and syringyl lignin monomers, respectively, liberated by thioacidolysis.
Mean and standard deviation (SD) values are based on five technical reaction replicates conducted with homogenized, ground, and extracted wood samples.

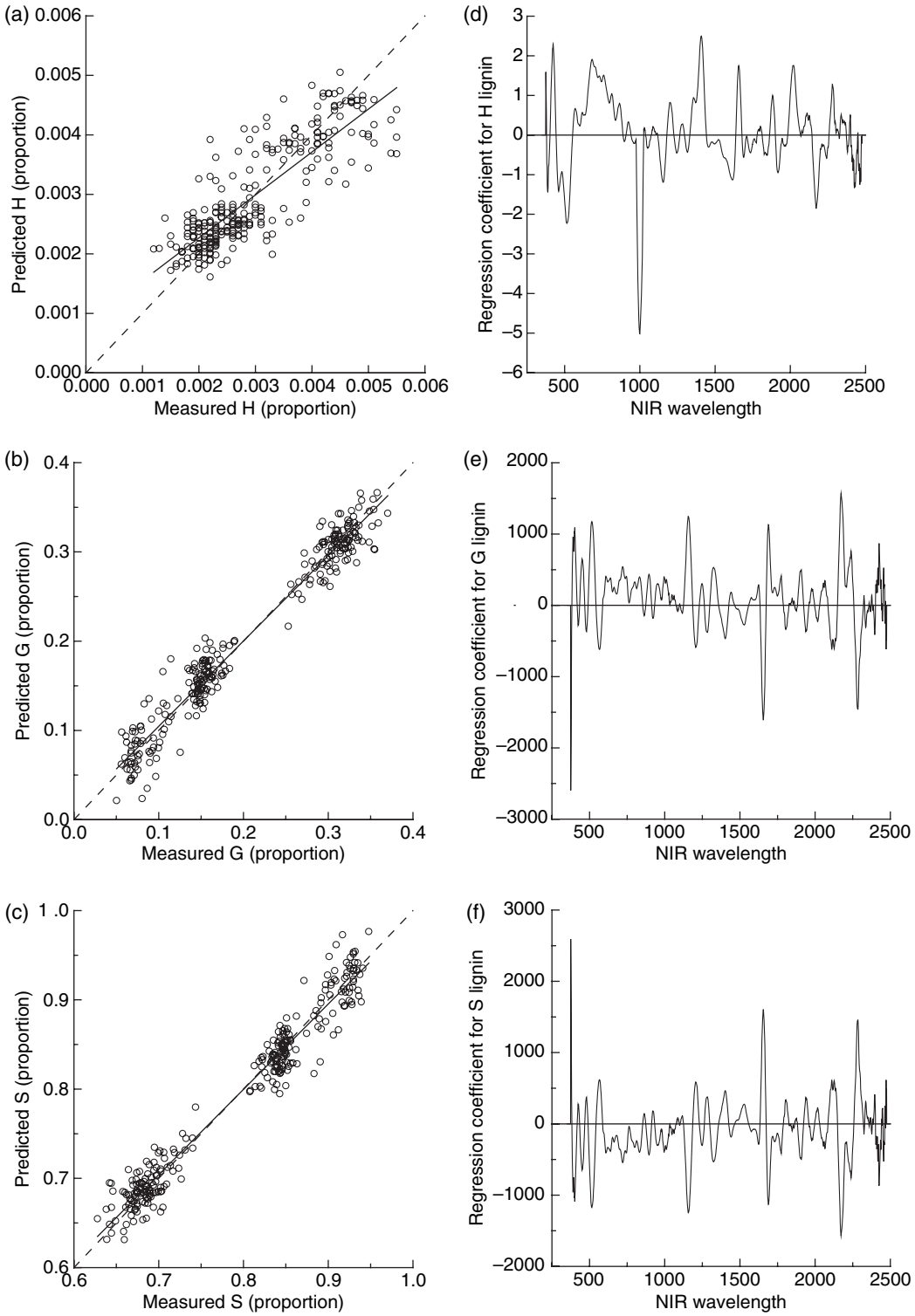

Figure 2. Near-infrared reflectance-based partial least squares regression (PLSR) prediction modeling of lignin monomer proportion in P717 poplars exhibiting variation in syringyl:guaiacyl (S:G) ratio due to transgene activity.

Results presented are derived from a model using the first eight principal components, under cross-validation.

(a)–(c) Plots of measured versus predicted values for lignin monomer proportion for *p*-hydroxyphenyl (H), G, and S monomers, respectively. Solid lines represent regression lines of best fit, while dashed lines represent optimal 1:1 relationships between measured and predicted values.

(d)–(f) Corresponding plots of regression coefficients for the first derivative NIR wavelengths with the PLSR-derived prediction model, for H, G, and S monomers, respectively (also based on the first eight principal components).

original and revised methods, with differences generally falling within a single standard deviation. Thioacidolysis yield (as measured by the sum of the *threo*- and *erythro*-isomers of H, G, and S monomers) was also very similar. Furthermore, the level of error does not appear to favor any monomer over another.

Traditional thioacidolysis techniques permit, on average, the preparation of about six to ten samples per day prior to monomer quantification by gas chromatography (GC) analysis. In contrast, the modified method permits upwards of 50 individual samples to be prepared in a single day, which can then be run on the GC for analysis the subsequent day. This substantially increases preparation number, and offers a medium- to high-throughput analytical protocol for accurately determining plant cell-wall lignin monomer subunit composition. Furthermore, the reduced volumes of reaction solvents (5- to 10-fold reduction in dioxane, BF_3, and ethane thiol) and purifying steps (up to 20-fold reduction in methylene chloride rinse volume etc.) offer added savings in preparation costs.

Prediction of proportion of lignin monomers via NIR reflectance spectroscopy

Following confirmation that the modified thioacidolysis protocol was suitable for quantifying the lignin monomer composition of plant cell-wall moieties, 267 independent thioacidolysis reactions were run to determine the monomer composition of wild-type and transgenic wood samples originating from individual trees exhibiting a board spectrum of lignin monomer compositions. These same samples were then NIR-scanned and the relationship between proportions of H, S, and G lignin and the NIR data (both raw and derivative transformed) modeled by partial least squares regression (PLSR) analysis, with complete cross-validation. Substantial variation in monomer composition was achieved by including trees from eight *C4H::F5H* transformed lines originating from independent transformation events and that displayed a change in S:G ratio when compared with the wild-type control trees. These were the same lines previously analyzed by Huntley *et al.* (2003). The sample and phenotypic structures of the dataset are described in Table 2.

A predictive model including the first 8 (of 15 calculated) principal components was deemed most suitable, as this model accounted for 95.896% of variance in the response variables, while the addition of subsequent components resulted in only marginal improvement (97.669% with all 15 components). The resultant NIR reflectance prediction model proved to be highly accurate under cross-validation, as indicated in the plots of measured versus predicted monomer proportion (Figure 2) and in the corresponding table of regression descriptors (Table 3). As anticipated (due to the low level of and consequent difficultly in quantifying H monomers), the accuracy of prediction was notably less for the H proportion than for G and S.

DISCUSSION

Revised thioacidolysis procedure

In this post-genomic era there is a pressing need to develop high-throughput phenotyping tools to work in parallel with, and complement, the rapidly advancing functional genomics toolbox (i.e. trancriptomics, proteomics, and metabolomics). Furthermore, advanced breeding programs, mapping populations and mutant populations, as well as our

Table 2 Lignin properties and sample counts of tree lines employed in near-infrared reflectance-based modeling of lignin monomer proportions by partial least squares regression

		Thiolignin yield (μmol g^{-1} Klason lignin)		Proportion of lignin monomers by class[b]					
				Hydroxyphenyl		Guaiacyl		Syringyl	
Tree line[a]	Tree count	Mean	SD	Mean	SD	Mean	SD	Mean	SD
P7-WT	104	2321.59	280.12	0.0028	0.0010	0.3158	0.0209	0.6814	0.0206
P7-CF-A	9	2519.57	100.95	0.0028	0.0006	0.2753	0.0197	0.7219	0.0196
P7-CF-B	82	2504.30	275.10	0.0033	0.0012	0.1556	0.0116	0.8411	0.0110
P7-CF-C	7	2750.77	152.40	0.0025	0.0006	0.1473	0.0143	0.8502	0.0146
P7-CF-D	5	2634.48	76.71	0.0023	0.0006	0.1360	0.0381	0.8617	0.0379
P7-CF-E	5	2771.64	147.65	0.0026	0.0003	0.1103	0.0116	0.8872	0.0117
P7-CF-F	12	2790.52	150.8	0.0028	0.0005	0.0883	0.0125	0.9089	0.0126
P7-CF-G	6	2773.49	116.35	0.0032	0.0006	0.0854	0.0069	0.9114	0.0069
P7-CF-H	37	2707.53	242.47	0.0028	0.0009	0.0803	0.0390	0.9169	0.0393
Total	267								

[a]Tree line descriptors: P7 = P717 hybrid poplar (*Populus alba* × *tremula*), WT, wild type; CF = *C4H::F5H*, and an alphabetic suffix denotes identity of transgenic lines derived from single transformation events. Lines have been ordered according to increasing guaiacyl:syringyl lignin monomer ratio.

[b]Lignin monomer proportions were determined by the revised thioacidolysis procedure, described herein. SD, standard deviation.

Table 3 Performance indicators of a near-infrared reflectance-based partial least squares regression model for prediction of lignin monomer proportion, as determined by a comparison between actual measured and model-predicted values. Indicators are provided for both model calibration and model cross-validation scenarios. Corresponding graphical representations of the cross-validation scenario are provided in Figure 1a–c

	Model calibration			Model validation		
Descriptor[a]	H	G	S	H	G	S
Slope	0.731	0.956	0.968	0.721	0.956	0.956
Offset	0.008	0.090	0.255	0.008	0.090	0.346
Correlation	0.855	0.979	0.984	0.843	0.979	0.979
R-square	0.731	0.957	0.968	0.710	0.957	0.958
RMSEC/RMSEP	0.005	0.202	0.176	0.005	0.202	0.201

[a]Descriptor explanations are as follows: Slope = slope of the regression line between measured and predicted values (ideally = ±1); Offset = *Y*-intercept of the regression line; Correlations = correlation coefficient between measured and predicted values; *R*-square = coefficient of determination (a measure of the degree of fit of the regression); RMSEC/RMSEP = root mean square error of calibration/prediction (the average modeling error).
H, G, and S = *p*-hydroxyphenyl, guaiacyl and syringyl lignin monomers, respectively.

traditional plant resources, require techniques that permit the accurate rapid characterization of plant cell walls, with minimal inputs (time, reagents, and equipment). Thus, the principal aim of this research was to modify an accepted protocol (i.e. thioacidolysis) to improve throughput while concurrently maintaining accuracy. The resultant protocol could then serve either as a standalone platform for analysis of plant cell-wall monomer composition, or as the wet chemical procedure required to establish predictive models based on other analytical tools, such as spectroscopy. Because independent models are typically required on a species by species basis, due to the substantial variation in the chemical and ultrasturctural architecture of plants, extensive calibration sets continue to be needed and the greater the number of samples, the higher the statistical accuracy.

Revisions to the procedure were attempted to increase the efficiency of the traditional thioacidolysis reaction employed to liberate lignin monomers for quantification by GC. Several iterations and modifications were attempted (data not shown) but, attesting to the development by the originators of the procedures, in the end the established chemical reaction was retained and most revisions were methodological rather than empirical. Essentially, the reaction volume was scaled down with as few changes to chemical ratios as possible. Volume scaling was central to achieving higher efficiency, as this allowed simultaneous processing of many samples via the use of alternative equipment. Simply reducing the reaction volume to small-scale 1-ml levels facilitated large-scale incubation of several reactions together in a temperature-controlled heating dry-block. Furthermore, the reduction in reaction volumes permitted the aqueous quenching and the subsequent organic solvent extraction procedures to be conducted in a single 5-ml glass vial without the need for a separatory funnel. Additionally, the final volume of extract in organic solvent was less than 2 ml, which allowed for the rapid evaporation of a large number of samples together in a Vacufuge (Eppendorf, http://www.eppendorf.com). Ultimately, these simple modifications now permit 50 or more samples to be processed per person, per day, without the need for specialized equipment and with substantially lower chemical input.

Equivalence of results from original and revised thioacidolysis protocols

Although the procedural revisions mainly involved scaling and equipment choice, there were some changes that potentially could have affected the analysis results. Comparison between the performance of the original and revised protocols, as applied to samples with highly deviating lignin monomer composition, was intended to assess any possible effects.

One major concern during scaling was the decision to supply the same amount of sample (as in the traditional method) to the modified reaction in order to maintain weighing accuracy and good sample representation and homogeneity. This meant that, in the revised protocol, the ratio between reaction volume and sample weight was decreased 10-fold compared with the original method. In their discussion of the analysis of lignocellulose residues (i.e. extractive-free wood flour) using thioacidolysis, Rolando *et al.* (1992) suggested that because the reagent is quickly consumed through reaction with glycosidic units in the sample, it is advisable to provide surplus reagent in order to ensure that this does not become a limiting factor. They suggest a minimum of 0.5 ml of reagent per mg of tissue – a ratio that was reduced considerably in the revised protocol. Another concern was the decision to reduce both the relative volume and the number of organic solvent extraction cycles to isolate the lignin breakdown products from the aqueous mixture, post-reaction. This was done primarily to reduce sample handling time, and to keep the final extract volume within the 2-ml capacity needed for ease of sample handling (a single glass vial) and evaporation by SpeedVac. However, this carried with it the risk of decreasing product yield or skewing monomer representation.

The results indicate that the revisions did marginally reduce yield (by ~5%), but had a limited influence on

monomer quantification (Table 1). Yield in thioacidolysis is ascribed to the β–O–4-ether content and the extent of its cleavage; only β-ether units that are either free-phenolic or are further 4–O–β-etherified will release the thioacidolysis monomers. However, it has been shown that thioacidolysis does not completely cleave all existing β-ethers (Ralph and Grabber, 1996), and as such the minor reductions in observed yield in the revised protocol may be a function of using lower reagent loadings. Furthermore, in poplar, some lignin units are gamma-*p*-hydroxybenzoylated; these units do not completely release the (unacylated) monomer (Grabber *et al.*, 1996), and therefore as a consequence of reagent scaling may account for the under-quantification. The yield is therefore always an underestimate of β-ether content, but if linkage analysis is the objective, then the modified protocol will deliver a further slight underestimation compared with the traditional method. However, monomer proportion, rather than yield, is the primary concern of the thioacidolysis method. As long as the yield is high enough to facilitate accurate measurement of relative abundance, this analysis has achieved its goal. The shifts observed in monomer proportion were small enough (and frequently statistically insignificant) that the measurements generated by the original and revised protocols could be considered 'substantially equivalent'. This study therefore indicates that for wood of varied lignin composition, from 12-week-old hybrid poplar, as little as 0.1 ml of reaction mixture per mg of tissue is sufficient to maintain a good yield, and accurate quantification. From a purely pragmatic point of view, especially for use as a large population-screening tool, the slight decrease in yield and minimal shifts in monomer proportions observed under the revised protocol are suggested to be acceptable, given the considerable resource- and time-related benefits it provides.

Predictive modeling of proportion of lignin monomers

According to visual and numerical indicators, the performance of the NIR reflectance-based predictive model for the proportion of lignin monomers was excellent, and comparable to the results achieved in other species and applications (Bailleres *et al.*, 2002; Alves *et al.*, 2006; Yamada *et al.*, 2006; Maranan and Laborie, 2008). Notably, prediction of G and S monomers in cross-validation exhibited minimal error across the range of phenotypic extremes. Because the correlation coefficients and slopes of the regression lines of the predicted value of these monomers against their measured value are approaching 1, a high level of confidence could be placed in further predictions made by this model for additional, similar wood specimens. The comparatively low accuracy of prediction for H monomers is probably related to the ordinarily low abundance of this monomer in poplar in combination with the error possibly introduced when measuring near the baseline of GC traces, and the limited variation observed in H monomer abundance across the samples modeled (the *C4H::F5H* construct had little effect on relative H monomer content).

The cross-validation regression plots (Figure 2) support the relationship existing between G and S monomer proportion in angiosperm wood of *C4H::F5H* modified trees, as has previously been shown (Huntley *et al.*, 2003) by thioacidolysis (Table 2). This is most apparent in the plots of regression coefficients for NIR wavelengths against the prediction model, in which the strength and importance of individual NIR wavelengths for G and S monomer prediction are complementary (Figure 2e, f). It is probable that the strength of this relationship contributed substantially to the accuracy of the predictive model, as PLSR involving multiple response variables takes implicit advantage of such patterns.

CONCLUDING REMARKS

It is apparent from these findings that the revised thioacidolysis protocol described here is a useful alternative to the classic procedure. This new protocol promotes a substantial increase in sample processing efficiency while markedly reducing equipment load and reagent consumption, and achieves this with little effect on the determination of lignin composition. The increased performance under this protocol makes thioacidolysis an effective tool for rapidly generating the large sets of phenotypic data required for NIR reflectance-based predictive modeling of lignin compositional traits. In applying such modeling to the wood of hybrid poplar, it was confirmed that data generated via the revised thioacidolysis protocol were of sufficient quality to allow accurate modeling of the relationships existing between NIR reflectance spectra and the proportion of the primary monomeric constituents in lignin (as measured by thioacidolsysis), across a broad phenotypic range.

EXPERIMENTAL PROCEDURES

Plant material and sample preparation

All trees were propagated from apical explants by sterile tissue culture techniques (3 weeks' growth cycle) and then repotted into a soil medium. Trees were then grown in a temperature-controlled greenhouse under a fixed 16-h photoperiod with supplemental lighting (radiant flux density of 300 W m^{-2}). Daily watering with fertigated water was achieved by flood-table irrigation. The trees were harvested 12 weeks after transfer to the greenhouse, and individual wood samples collected from each stem 10 cm above the base of the root collar. The samples were then stripped of bark and pith and oven-dried (50°C) for 2 days. Once dry, the wood was ground to a flour in a Wiley mill (40 mesh), and extracted for 12 h with hot acetone in a Soxhlet apparatus to remove extractives.

Thioacidolysis procedure

Wood samples involved in the comparison between thioacidolysis procedures were initially processed as per Rolando *et al.* (1992),

using 10 mg of ground, extract-free oven-dried wood flour as the substrate, and a ratio of 1 ml of reaction mixture to 1 mg of sample. The same samples, as well as those used in NIR reflectance-based predictive modeling, were also processed according to the following revised protocol. For each sample, 10 mg of ground, extract-free oven-dried wood flour was weighed into a 5-ml glass Wheaton vial with Teflon-lined screw-cap. One milliliter of freshly made reaction mixture [2.5% boron trifluoride etherate and 10% ethanethiol, in recently distilled dioxane (v/v)] was added to each vial and blanked with nitrogen gas prior to sealing. Vials were then placed together in a dry heating block (100°C) for 4 h, with periodic (hourly) manual agitation. The reaction was halted by placing the reactions at –20°C for 5 min. Internal standard (5 mg ml^{-1} tetracosane in methylene chloride, 0.2 ml) was then added to each vial with enough 0.4 M sodium bicarbonate to bring the reaction pH to between 3 and 4 (~0.3 ml, as determined by pH indicator paper). To extract the reaction products from the aqueous mixture, 2 ml of water and 1 ml of methylene chloride were added to each vial, which was then recapped, vortexed, and allowed to settle, phase-separating the upper (aqueous) and lower (organic, and containing lignin breakdown products) phases. An aliquot (1.5 ml) of the organic phase was removed by autopipette, and simultaneously cleared of residual water and filtered by passing through a Pasteur pipette packed with a small tissue-paper plug and an inch (~50 mg) of granular anhydrous sodium sulfate, and transferred directly into a 2-ml polypropylene microfuge tube. Samples were then collectively evaporated to dryness in an Eppendorf Vacufuge (approximately 1.5 h at 45°C), and resuspended in 1 ml of methylene chloride. Samples were derivatized by combining 20 μl of resuspended sample with 20 μl of pyridine and 100 μl of *N,O*-bis(trimethylsilyl) acetamide (Sigma, http://www.sigmaaldrich.com/). After incubation for at least 2 h at 25°C, 1 μl of this reaction product was analyzed by GC.

Gas chromatography

Gas chromatography (http://www.chem.agilent.com) was conducted on a Hewlett Packard 5890 series II instrument, fitted with an autosampler, splitless injector, flame ionization detector (FID), and a 30-m RTX5ms 0.25-mm internal diameter capillary column. One-microliter injections were separated using helium as a carrier gas at 1 ml min^{-1}. Inlet and detector temperatures were set to 250°C, while the oven profile consisted of: initial temperature 130°C, hold 3 min, ramp temperature 3°C min^{-1} for 40 min to give a final temperature of 250°C, hold 5 min, cool. Peak identification was consistent with that in Rolando *et al.* (1992).

NIR reflectance spectroscopy

The light reflectance of wood samples across the NIR spectrum was measured with an Analytical Spectral Devices Inc., QualitySpec Pro NIR spectrophotometer (http://www.asdi.com/), equipped with a round 1.5-cm diameter sample window (Muglite, Analytical Spectral Devices). The scanning range was from 350–2500 nm, with a 2-nm interval, interpolated to 1 nm.

Statistical analysis

Prediction modeling was conducted using the PLSR package provided in The UNSCRAMBLER 9.1 software (Camo Technologies, http://www.camotechnologies.com/), employing full cross-validation as a modeling option. Prior to PLSR, NIR reflectance data were transformed into the Savitzky–Golay first derivative, with the averaging/smoothing process spanning 25 wavelengths either side of each data point, and an order of 2 for the polynomial approximation process.

ACKNOWLEDGEMENTS

Funding for this project was provided by the Natural Sciences and Engineering Research Council of Canada Grants held by SDM. AR would also like to recognize the Top Achiever Doctoral Scholarship (Bright Future Scholarships, Tertiary Education Commission, Wellington, New Zealand) held between 2002 and 2005.

REFERENCES

Alves, A., Schwanninger, M., Pereira, H. and Rodrigues, J. (2006) Calibration of NIR to assess lignin composition (H/G ratio) in maritime pine wood using analytical pyrolysis as the reference method. *Holzforschung*, **60**, 29–31.

Bailleres, H., Davrieus, F. and Pichavant, F.H. (2002) Near infrared analysis as a tool for rapid screening of some major wood characteristics in a eucalyptus breeding program. *Ann. For. Sci.* **59**, 479–490.

Blakeney, A.B. and Flinn, P.C. (2005) Determination of non-starch polysaccharides in cereal grains with near-infared reflectance spectroscopy. *Mol. Nutr. Food Res.* **49**, 546–550.

Campbell, M.M. and Sederoff, R.R. (1996) Variation in lignin content and composition - Mechanism of control and implications for the genetic improvement of plants. *Plant Physiol.* **110**, 3–13.

Carpita, N.C. and McCann, M.C. (2002) The functions of cell wall polysaccharides in composition and architecture revealed through mutations. *Plant Soil*, **247**, 71–80.

Chen, L.M., Carpita, N.C., Reiter, W.D., Wilson, R.H., Jeffries, C. and McCann, M.C. (1998) A rapid method to screen for cell-wall mutants using discriminant analysis of Fourier transform infrared spectra. *Plant J.* **16**, 385–392.

Coleman, H.D., Park, J.Y., Nair, R., Chapple, C. and Mansfield, S.D. (2008) RNAi-mediated suppression of *p*-coumaroyl-CoA 3′-hydroxylase in hybrid poplar impacts lignin deposition and soluble secondary metabolism. *Proc. Natl Acad. Sci. USA*, **105**, 4501–4506.

Franke, R., McMichael, C.M., Meyer, K., Shirley, A.M., Cusumano, J.C. and Chapple, C. (2000) Modified lignin in tobacco and poplar plants over-expressing the Arabidopsis gene encoding ferulate 5-hydroxylase. *Plant J.* **22**, 223–234.

Grabber, J.H., Quideau, S. and Ralph, J. (1996) *p*-Coumaroylated syringyl units in maize lignin; implications for β-ether cleavage by thioacidolysis. *Phytochem.* **43**, 1189–1194.

Huang, A.M., Jiang, Z.H. and Li, G.Y. (2007) Determination of holocellulose and lignin content in Chinese fir by near infrared spectroscopy. *Spectrosc. Spectr. Anal.* **27**, 1328–1331.

Huntley, S.K., Ellis, D., Gilbert, M., Chapple, C. and Mansfield, S.D. (2003) Significant increases in pulping efficiency in C4H-F5H-transformed poplars: Improved chemical savings and reduced environmental toxins. *J. Agric. Food. Chem.* **51**, 6178–6183.

Jones, P.D., Schimleck, L.R., Peter, G.F., Daniels, R.F. and Clark, A. (2006) Nondestructive estimation of wood chemical composition of sections of radial wood strips by diffuse reflectance near infrared spectroscopy. *Wood Sci. Technol.* **40**, 709–720.

Lapierre, C., Monties, B. and Rolando, C. (1985) Thioacidolysis of lignin: comparison with acidolysis. *J. Wood Chem. Technol.* **5**, 277–292.

Li, G.Y., Huang, A.M., Wang, G., Qin, D.C. and Jiang, Z.H. (2007) Rapid determination of mason lignin content in bamboo by NIR. *Spectrosc. Spectr. Anal.* **27**, 1977–1980.

Maranan, M.C. and Laborie, M.P.G. (2008) Rapid prediction of the chemical traits of hybrid poplar with near infrared spectroscopy. *J. Biobased Mater. Bioenergy*, **2**, 57–63.

Meyer, K., Cusumano, J.C., Somerville, C. and Chapple, C.C.S. (1996) Ferulate-5-hydroxylase from Arabidopsis thaliana defines a new family of cytochrome P450-dependent monooxygenases. *Proc. Natl Acad. Sci. USA*, **93**, 6869–6874.

Meyer, K., Shirley, A.M., Cusumano, J.C., Bell-Lelong, D.A. and Chapple, C. (1998) Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in Arabidopsis. *Proc. Natl Acad. Sci. USA*, **95**, 6619–6623.

Poke, F.S. and Raymond, C.A. (2006) Predicting extractives, lignin, and cellulose contents using near infrared spectroscopy on solid wood in *Eucalyptus globulus*. *J. Wood Chem. Technol.* **26**, 187–199.

Ralph, J. and Grabber, J.H. (1996) Dimeric β-ether thioacidolysis products resulting from incomplete ether cleavage. *Holzforschung*, **50**, 425–428.

Ralph, J., Lundquist, K., Brunow, G. *et al.* (2004) Lignins: natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids. *Phytochem. Rev.* **3**, 29–60.

Rolando, C., Monties, B. and LaPierre, C. (1992) Thioacidolysis. In *Methods in Lignin Chemistry* (Lin, S. and Dence, C., eds). Berlin: Springer-Verlag, pp. 334–349.

Stewart, J.J., Kadla, J.F. and Mansfield, S.D. (2006) The influence of lignin chemistry and ultrastructure on the pulping efficiency of clonal aspen (Populus tremuloides Michx.). *Holzforschung*, **60**, 111–122.

Sykes, R., Li, B.L., Hodge, G., Goldfarb, B., Kadla, J. and Chang, H.M. (2005) Prediction of loblolly pine wood properties using transmittance near-infrared spectroscopy. *Can. J. For. Res.* **35**, 2423–2431.

Takayama, M., Johjima, T., Yamanaka, T., Wariishi, H. and Tanaka, H. (1997) Fourier transform Raman assignment of guaiacyl and syringyl marker bands for lignin determination. *Spectroc. Acta Pt. A-Molec. Biomolec. Spectr.* **53**, 1621–1628.

Via, B.K., So, C.L., Groom, L.H., Shupe, T.F., Stine, M. and Wikaira, J. (2007) Within tree variation of lignin, extractives, and microfibril angle coupled with the theoretical and near infrared modeling of microfibril angle. *IAWA J.* **28**, 189–209.

Yamada, T., Yeh, T.F., Chang, H.M., Li, L.G., Kadla, J.F. and Chiang, V.L. (2006) Rapid analysis of transgenic trees using transmittance near-infrared spectroscopy (NIR). *Holzforschung*, **60**, 24–28.

Yeh, T.F., Chang, H.M. and Kadla, J.F. (2004) Rapid prediction of solid wood lignin content using transmittance near-infrared spectroscopy. *J. Agric. Food. Chem.* **52**, 1435–1439.

