Near Infrared Spectroscopy for Cost Effective Screening of Foliar Oil Characteristics in a *Melaleuca cajuputi* Breeding Population

LAURENCE R. SCHIMLECK*

CSIRO Forestry and Forest Products, Private Bag 10, Clayton South MDC, Victoria, 3169, Australia

JOHN C. DORAN

CSIRO Forestry and Forest Products, P.O. Box E4008, Kingston, ACT, 2604, Australia

ANTO RIMBAWANTO

Centre of Forest Biotechnology and Tree Improvement, Jalan Palagan Tentara Pelajar Km 15, Purwobinangan, Pakem, Sleman, Yogyakarta 55582, Indonesia

The identification of *Melaleuca cajuputi* leaf samples (trees) that demonstrate enhanced oil characteristics using near infrared (NIR) spectroscopy is described. Leaf samples from an unthinned *M. cajuputi* seedling seed orchard in Indonesia were collected and air-dried, and their 1,8-cineole content and oil concentrations were determined. NIR spectra of the leaves were obtained, and calibrations for 1,8-cineole content and oil concentration were developed using spectra that had been selected using spectral features; that is, no knowledge of 1,8-cineole content or oil concentration was used to select the calibration samples. The calibrations were used to predict the 1,8-cineole content and oil concentration of the remaining samples. It was demonstrated that NIR spectroscopy could be used to identify leaf samples that had high 1,8-cineole contents and oil concentrations. The technique has the potential to greatly reduce the time involved in ranking large numbers of samples for these attributes, as is a requirement in tree breeding programs to enhance oil production.

KEYWORDS: Near infrared spectroscopy; *Melaleuca cajuputi*; oil concentration; 1,8-cineole content; leaves

INTRODUCTION

Leaves of *Melaleuca cajuputi* subspecies *cajuputi* Powell (Myrtaceae), hereafter referred to as *M. cajuputi*, are the principal source of medicinal cajuput oil widely used as a folk medicine throughout Southeast Asia (*1*). Indonesia is the major supplier of the oil (>300 tons/annum), which is extracted from the leaf by steam distillation. A selection and breeding program to enhance oil yield (a combination of leaf biomass/ha and oil concentration in those leaves) and quality (% 1,8-cineole of total oil) from *M. cajuputi* plantations on Java commenced in 1998 with the planting of the first seedling seed orchard at Paliyan near Yogyakarta (*2*).

Improvement in heritable traits such as oil concentration and percent 1,8-cineole in seedlots from orchards of this type requires screening of orchard trees for the traits of interest.

Heavy culling of orchard trees follows, based ideally on a combined selection index incorporating leaf biomass, oil concentration, and oil quality data, to leave only the very best genotypes to cross-pollinate and produce improved seed. In the case of the Paliyan orchard, a total of about 500 trees were screened for their foliar oil characteristics, but in larger orchards, the sample numbers are often in the thousands. Conventional methods of determining foliar oil traits require extraction of oil by steam distillation or solvent methods followed by chemical analysis of the oil using gas liquid chromatography (*3*, *4*). They are laboratory-based, time-consuming, and thus expensive especially for large sample sets. Clearly, a rapid method of ranking the value of a tree as an oil producer and particularly one that could be employed in the field would be highly advantageous in accelerating the selection process and reducing costs.

A rapid alternative to existing methods of analysis is provided by near infrared (NIR) spectroscopy. It has been demonstrated that NIR spectroscopy can be used to predict 1,8-cineole content

* To whom correspondence should be addressed. Present address: Warnell School of Forest Resources, Room 4-330, The University of Georgia, Athens, GA 30602-2152. Tel.: (706)583 0464. Fax: (706)542 8356. E-mail: lschimleck@smokey.forestry.uga.edu.

in leaves (*5, 6*) and eucalyptus oil (*7*). The study by Ebbers et al. (*6*) is particularly relevant as it was conducted on the leaves of *Eucalyptus*, a close relative of *Melaleuca*.

The objective of this study was to explore the applicability of NIR spectroscopy as an initial screening tool for ranking trees of *M. cajuputi* based on commercially desirable foliar oil traits. Specific aims in achieving this goal were (i) to develop calibrations for 1,8-cineole content and oil concentration using NIR spectra obtained from selected *M. cajuputi* leaf samples; (ii) to predict the 1,8-cineole content and oil concentration of a large number of *M. cajuputi* leaf samples in a separate set; and (iii) to rank the leaf samples based on NIR-predicted 1,8-cineole contents and oil concentrations and to compare these rankings with laboratory determined data.

EXPERIMENTAL PROCEDURES

Samples. A total of 315 air-dried leaf samples from 2 year old trees of *M. cajuputi* growing in the Paliyan seedling seed orchard were available for analysis. The orchard had already been thinned to half original stocking based on growth characteristics. One hundred grams of fresh leaves was harvested from the lower crown of each remaining tree, labeled to origin, and bagged. After they were collected, the leaf samples were air-dried at room temperature in Yogyakarta (ca. 28 °C) for 1 month. They were then air-freighted to Canberra, where after fumigation in methyl bromide and irradiation, they were released to CSIRO Forestry and Forest Products for characterization of their oils. These treatments are not known to affect oil properties.

Conventional Methods of Estimating Oil Concentration and 1,8-Cineole Content. *Determination of Oil Concentration.* The foliar essential oils of each selected tree were obtained by steam distillation of a subsample (usually 25 g) of each air-dried sample of leaves for 2 h with cohobation in a modified Dean and Stark apparatus (*3*). Following distillation, pentane (1 mL) was added to maximize oil recovery, which was then dried over sodium sulfate and pentane evaporated at room temperature. A separate subsample (usually 3 g) of each collection was oven-dried at 70 °C for 12 h for moisture content determination, so that the oil concentration could be expressed as a percentage of leaf dry weight (W/W% DW).

Determination of 1,8-Cineole Percentage (of Total Oils). 1,8-Cineole content was determined by gas chromatography (GC) using a Hewlett-Packard 6890A gas chromatograph, with a 3390A integrator, an Alltech AT 35 column (60 m × 0.25 mm i.d.), and a flame ionization detector operating at 300 °C. A 1 μL sample of oil in ethanol was injected at 200 °C. Hydrogen was used as the carrier gas (40−50 cm³/min) with an oven temperature program of 10 °C/min from 50 (1 min) to 250 °C (4 min). Major components were identified by comparing retention times against laboratory-supplied Aldrich/Sigma pure components.

NIR Analysis. *Samples and Their Preparation.* The 315 samples (trees) for this study covered the full range of values for oil concentration and percent 1,8-cineole found in the orchard population, as determined by steam distillation GC. Because of the large variation in the physical size of the samples, the air-dried leaf samples were prepared for NIR analysis by cutting them into large pieces using scissors.

Measurement of NIR Spectra and Spectral Manipulation. The cut leaves of each sample were placed in a FOSS NIRSystems large sample cup (NR-7070). The NIR spectra were measured in diffuse reflectance mode from samples held in a spinning sample holder in a FOSS NIRSystems Inc. model 5000 scanning spectrophotometer. The spectra were collected at 2 nm intervals over the wavelength range of 1100−2500 nm. The instrument reference was a ceramic standard. Fifty scans were accumulated for each sample, and the results were averaged. After the spectrum had been obtained, the sample cup was emptied and repacked and a duplicate spectrum was obtained.

The duplicate spectra were averaged and then imported into the WinISI II (version 1.50) software package (Infrasoft International, Port Matilda, PA). A standard normal variate (SNV) plus detrend math

Table 1. Summary Statistics for 1,8-Cineole Contents and Oil Concentrations for the Calibration and Prediction Sets[a]

leaf parameter	calibration set (81 samples)				prediction set (234 samples)			
	min	max	avg	SD	min	max	avg	SD
1,8-cineole content (%)	2.3	63.3	45.2	16.3	3.8	73.9	50.6	11.5
oil concentration (W/W% DW)	0.34	4.78	2.11	0.75	0.41	4.76	2.13	0.71

[a] Note that samples excluded from the calibration development were not used in the determination of the summary statistics.

treatment was used to normalize the spectra. The spectra were then converted to the second derivative using a gap width of 4 nm.

Calibration Development. WinISI II software was used to select samples for calibration development. At this point, it was assumed that no information relating to the 1,8-cineole content and oil concentration of the samples was available and sample selection was based solely on the software to identify spectrally unique samples.

WinISI II uses a neighborhood concept to identify samples for calibration development. A description of how samples are selected using this method is described in the ISI Windows Near Infrared Software manual (*8*). A neighborhood H of 1.2 was used, 81 samples were identified for the development of calibrations for 1,8-cineole content and oil concentration, and the remaining samples (234) formed the prediction set. A statistical summary of the calibration and prediction sets is given in **Table 1**.

The calibrations were developed using modified partial least squares (MPLS) regression (*9*) and the math-treated spectra. Calibrations were developed with four cross-validation segments. The standard error of cross-validation (SECV) (determined from the residuals of each cross-validation phase), the standard error of calibration (SEC) (determined from the residuals of the final calibration), and the coefficient of determination (R^2) were used to assess calibration performance.

The standard error of prediction (SEP) gives a measure of how well a calibration predicts the parameter of interest for a set of unknown samples that are different from the calibration test set. The ratio of performance to deviation (RPD), calculated as the ratio of the standard deviation (SD) of the reference data to the SEP, was used to determine the predictive ability of the calibrations (*10*). An RPD of greater than 2.5 is considered satisfactory for screening.

RESULTS AND DISCUSSION

1,8-Cineole content and oil concentration in leaves of individual trees are two important criteria that must be evaluated when selecting trees for retention (or culling) in breeding populations of tree improvement programs aimed at increasing the yield and quality of cineole-rich oil. Presently, the identification of trees is based on wet chemical methods, such as steam distillation or solvent extraction followed by GC. When several hundred leaf samples are screened, the analysis is time-consuming and costly. The objective of this study was to explore the applicability of NIR spectroscopy as an initial screening tool for ranking trees of *M. cajuputi* based on commercially desirable foliar oil traits.

Variation of NIR Spectra. The NIR diffuse reflectance spectra of the *M. cajuputi* leaf samples were examined to investigate if spectral variation could be related to variation in leaf chemical composition. Considerable variation was observed between 1650−1800 and 2200−2400 nm (**Figure 1**). Similar regions (1700−1800 and 2200−2300 nm) were observed for the transmission NIR spectrum of pure 1,8-cineole (*6*). The strongest correlation with 1,8-cineole was observed with the
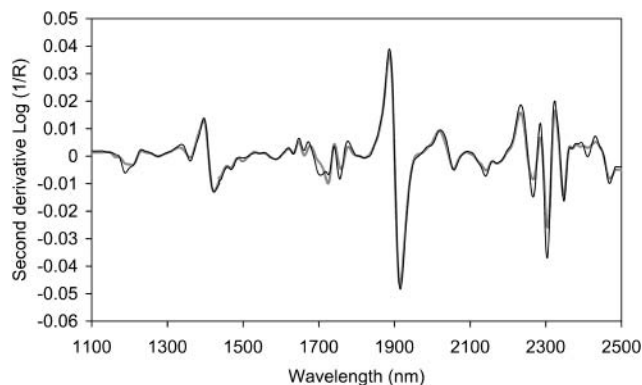
Figure 1. Variation in diffuse reflectance NIR spectra of two *M. cajuputi* leaf samples. The thick, light line represents a sample that had a low 1,8-cineole content (3.2%) and oil concentration (0.55%, W/W% DW) while the thin, dark line represents a sample that had a high 1,8-cineole content (61.9%) and oil concentration (4.6%, W/W% DW).

band at 1768 nm ($R^2 = 0.53$), while the band at 1696 nm had the strongest relationship with oil concentration ($R^2 = 0.21$).

PLS Calibrations Based on Selected Samples. A total of 81 samples (as selected by the WinISI II software) were used for the development of 1,8-cineole content and oil concentration calibrations. The 1,8-cineole content calibration had strong calibration statistics. Four samples were omitted from the calibration because of their large residuals (where the residual equals the measured value minus the NIR fitted value) during cross-validation. The relationship between measured 1,8-cineole content and NIR-fitted 1,8-cineole content was good over the range of the calibration (**Figure 2a**). The distribution of 1,8-cineole contents in the calibration set was biased toward samples with high 1,8-cineole contents because the majority of samples (62%) used to form the calibration and prediction sets had 1,8-cineole contents greater than 50%.

The oil concentration calibration had calibration statistics inferior to those of the 1,8-cineole content calibration. Despite the moderate R^2, the relationship between NIR-fitted oil concentration and measured oil concentration was good over the range of the calibration set (**Figure 2b**). Three samples were omitted from the calibration because of their large residuals. Unlike 1,8-cineole content, the oil concentration distribution was relatively normal.

Prediction of 1,8-Cineole Content and Oil Concentration. The 1,8-cineole content and oil concentration of the remaining 234 samples were predicted by the respective calibrations. The relationship between measured 1,8-cineole content and NIR-predicted 1,8-cineole content ($R^2 = 0.63$) was quite good considering the variation in the sample set (**Figure 3a**). The SEP was considerably higher than the SEC but similar to the SECV. The RPD (1.44) was low indicating that if the 1,8-cineole content calibration was used it would only be suitable for initial screening.

Two samples had residuals greater than ±20, one sample had a low 1,8-cineole content (11.8%) that was overestimated by the calibration (residual = 21.4), and the other was underestimated. Considering the distribution of samples in the calibration set, the overestimation of samples with low 1,8-cineole content would be expected.

The relationship between measured oil concentration and NIR-predicted oil concentration ($R^2 = 0.54$) was reasonable (**Figure 3b**). The SEP was similar to the SEC reported for the oil concentration calibration, but the RPD was again low (1.45).
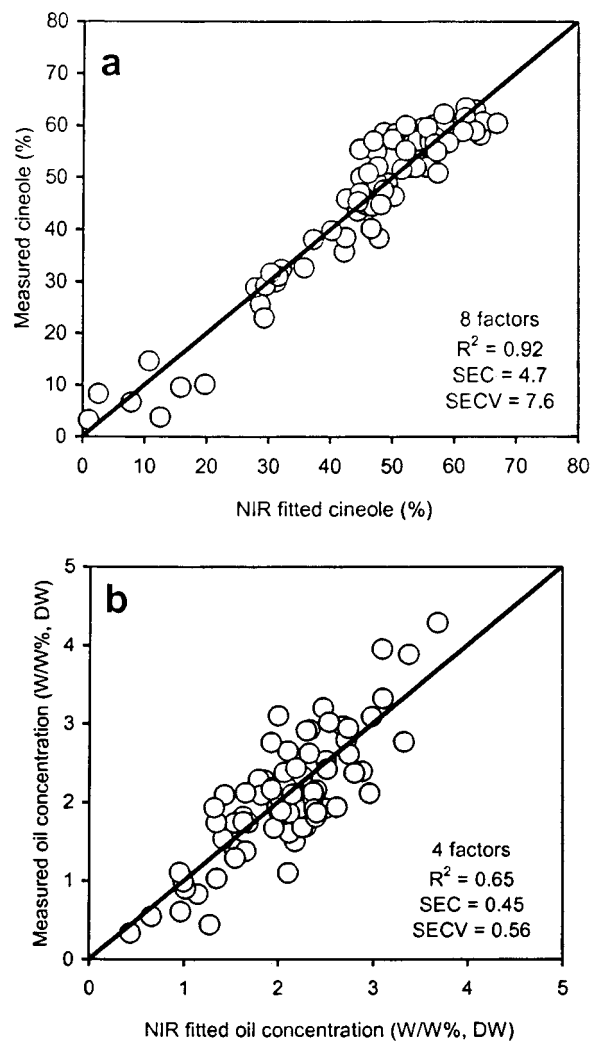

Figure 2. Relationship between measured values and NIR estimated values for (**a**) 1,8-cineole content and (**b**) oil concentration.

Rankings Based on Predicted 1,8-Cineole Contents. On the basis of NIR-predicted 1,8-cineole content, samples that had 1,8-cineole contents greater than 50% were identified (**Table 2**). On the basis of the NIR predictions, 68% of samples with 1,8-cineole contents greater than 50% were correctly identified. Of the samples with measured 1,8-cineole contents greater than 50%, 30 had 1,8-cineole contents greater than 60%. If the 50% cutoff was used, NIR spectroscopy correctly identified 27 (90%) of these.

Rankings Based on Predicted Oil Concentration. A total of 41 samples were identified by NIR spectroscopy as having oil concentrations greater than 2.6% (W/W% DW). Of these samples, 31 actually had oil concentrations greater than 2.6%. Many samples (16) with oil concentrations greater than 2.6% were not identified because of the underestimation of oil concentration. In terms of ranking, the 41 samples actually included 23 of the top 41 samples and all of the top 10.

Of the 41 samples identified by NIR spectroscopy as having oil concentrations greater than 2.6%, all but two had a measured 1,8-cineole content greater than 41%, 31 had 1,8-cineole contents greater than 50%, and eight of these had 1,8-cineole contents greater than 60% demonstrating that NIR spectroscopy could be used to identify leaf samples with both high 1,8-cineole content and oil concentration.
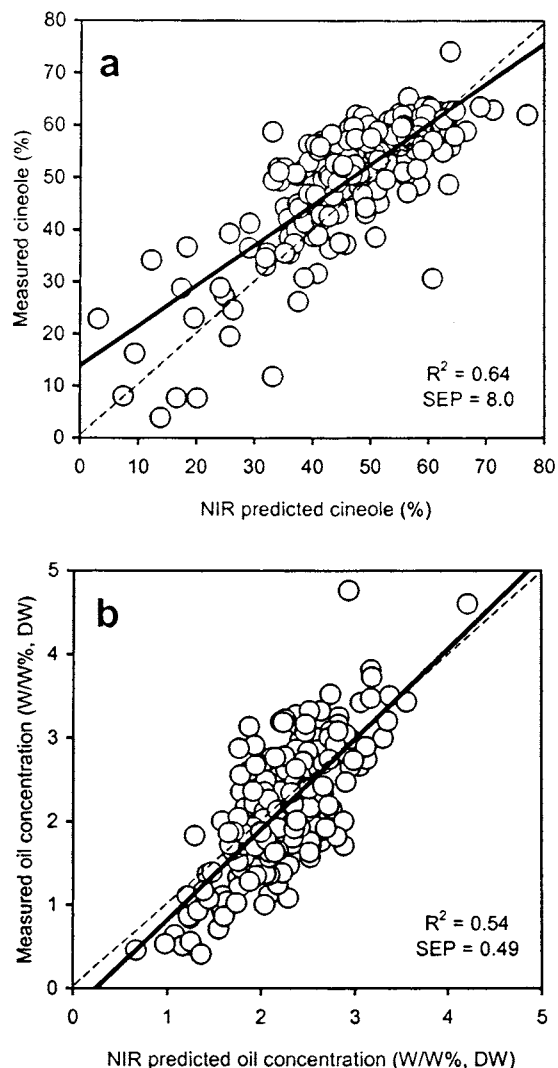
Figure 3. Relationship between measured values and NIR predicted values for (**a**) 1,8-cineole content and (**b**) oil concentration.

Table 2. Number of Samples with 1,8-Cineole Content Greater Than 50% as Measured by Steam Distillation and Predicted by NIR Spectroscopy

		1,8-cineole content (% of total oil) measured by steam distillation		
		>50%	<50%	total
1,8-cineole content (% of total oil) predicted by NIR	>50%	104	8	112
	<50%	50	72	122
	total	154	80	234

Cost Per Sample: NIR Spectroscopy vs Wet Chemistry. It is proposed that NIR spectroscopy could be used to identify samples for further analysis to determine their actual 1,8-cineole contents and oil concentrations. The analysis of only selected samples would greatly reduce the time and cost of initial screening work. The cost of evaluating each sample by NIR is estimated by the authors to be in the range of AUD15 to AUD20 (AUD = Australian dollars) per sample as compared to commercial rates for steam distillation GLC of about AUD66 per sample and solvent extraction GC of AUD55 (I. Southwell, personal communication, 2002).

Application of NIR Spectroscopy to Fresh Leaf Samples. If the methods described in this study were to be used in practice, the most time-consuming aspect would be sample drying. In a recent study of two eucalyptus species (*Eucalyptus melliodora* and *Eucalyptus globulus*), Ebbers et al. (6) showed that NIR spectra collected from freeze-dried whole leaf samples gave superior calibration statistics for 1,8-cineole content as compared to calibrations developed using freeze-dried ground leaf samples. These findings show that whole leaf samples can be tested, assuming that a sample holder large enough to accommodate whole leaf samples is available. As compared to the leaves of *E. globulus*, the leaves of *M. cajuputi* are relatively small and could easily be held by the NIRSystems Transport Module (with a large rectangular sample cell).

Ebbers et al. (6) also showed that spectra collected from fresh leaf samples could be used to develop calibrations for 1,8-cineole content. The 1,8-cineole calibration reported for *E. melliodora* using fresh leaf samples had a higher R^2 (0.97) than the calibrations developed using freeze-dried whole leaf (0.94) and freeze-dried ground leaf samples (0.88) (6). The 1,8-cineole calibration developed for *E. globulus* using fresh leaf samples was not as good ($R^2 = 0.74$), and it was suggested that interference from waxes present on the surface of *E. globulus* leaves was responsible. These findings indicate that collection of NIR spectra from fresh leaves and the subsequent development of calibrations are possible, negating the need to dry samples. It is therefore possible that NIR spectroscopy could be used in field situations to make decisions about which trees should be sampled for steam distillation and GC.

LITERATURE CITED

(1) Doran, J. C. Cajuput oil. In *Tea Tree. The Genus Melaleuca*; Southwell, I. A., Lowe, R. F., Eds.; In *Medicinal and Aromatic Plants—Industrial Profiles*; Hardman, R., Ed.; Harwood Academic Press: Amsterdam, 1999; Vol. 9, pp 221−233.

(2) Susanto, M.; Doran, J.; Arnold, R.; Rimbawanto, A. Genetic variation in growth and oil characteristics of *Melaleuca cajuputi* subsp. *cajuputi* and potential for genetic improvement. *J. Trop. For. Sci.* **2003**, in press.

(3) Brophy, J. J.; House, A. P. N.; Boland, D. J.; Lassak, E. V. Digests of the essential oils of 111 species from northern and eastern Australia. In *Eucalyptus Leaf Oils*; Boland, D. J., Brophy, J. J., House, A. P. N., Eds.; Inkata Press: Melbourne, 1991; pp 29−155.

(4) Baker, G. R.; Lowe, R. F.; Southwell, I. A. Comparison of oil recovered from tea tree leaf by ethanol extraction and steam distillation. *J. Agric. Food Chem.* **2000**, *48*, 4041−4043.

(5) Schulz, H.; Drews, H. H.; Kruger, H. Rapid NIRS determination of quality parameters in leaves and isolated essential oils of *Mentha* species. *J. Essent. Oil Res.* **1999**, *11* (2), 185−190.

(6) Ebbers, M. J. H.; Wallis, I. R.; Dury, S.; Floyd, R.; Foley, W. J. Spectrometric predictions of secondary metabolites and nitrogen in fresh *Eucalyptus* foliage: towards remote sensing of the nutritional quality of foliage for leaf eating marsupials. *Aust. J. Bot.* **2002**, *50*, 761−768.

(7) Wilson, N. D.; Watt, R. A.; Moffat, A. C. A near-infrared method for the assay of cineole in *Eucalyptus* oil as an alternative to

the official BP method. *J. Pharm. Pharmacol.* **2001**, *53*, 95−102.

(8) Infrasoft International. *ISI Windows Near Infrared Software*, *WinISI II*, version 1.50, 2000.

(9) Shenk, J. S.; Westerhaus, M. O. New standardisation and calibration procedures for near infrared reflectance spectroscopy. *Crop Sci.* **1991**, *31*, 469−474.

(10) Williams, P. C.; Sobering, D. C. Comparison of commercial near infrared transmittance and reflectance instruments for the analysis of whole grains and seeds. *J. Near Infrared Spectrosc.* **1993**, *1*, 25−33.