Deviant near-infrared spectra identifies *Corymbia* hybrids

Myralyn Abasolo [a], David J. Lee [b], Carolyn Raymond [c], Roger Meder [d], Mervyn Shepherd [a,*]

[a] Cooperative Research Centre for Forestry, Southern Cross Plant Science, Southern Cross University, PO Box 157, Lismore, NSW 2480, Australia
[b] Faculty of Science, Health, Education and Engineering, University of the Sunshine Coast and Department of Agriculture, Fisheries and Forestry, Locked Bag 4 Maroochydore DC, QLD 4558, Australia
[c] Southern Cross Plant Science, Southern Cross University, PO Box 157, Lismore, NSW 2480, Australia
[d] Division of Plant Industry, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Queensland Bioscience Precinct, 306 Carmody Road, St. Lucia, QLD 4067, Australia

ARTICLE INFO

ABSTRACT

The study examined the potential of Near Infrared Reflectance (NIR) spectroscopy for field diagnosis of hybrids between *Corymbia* (formerly *Eucalyptus*) species. NIR profiles were generated by scanning foliage from a total of 383 hybrid and 533 parental seedlings grown in a common garden and partial least squares discriminant analysis was used to test three-way model power to assign individuals to their appropriate taxon; either a parental or F_1 hybrid class. Using the optimised conditions, fresh foliage from eight-month-old seedlings and a handheld NIR instrument (950–1800 nm), the mean assignment rates for the three hybrid groups ranged from 76% to 90%. Hybrid-parent contrast of NIR spectra deviated more so than parent–parent contrast. The F_1 taxon assignment rates were usually higher than those for parents at 100% and 72%, respectively. Hybrid resolution was even greater for 2nd generation backcross hybrids. Similar to studies of morphology, taxon assignments tended to be more accurate for hybrid groups in which the parental taxa were more divergent. The practical application of this technique for hybrid diagnosis of seedlings in the nursery will require careful attention to control environmental factors because seedling age and storage effects influenced the ability of NIR to identify hybrids. The technique may also necessitate the generation of comparable reference populations, although exclusions approaches to analysis may circumvent the need for reference populations. The application of NIR in field diagnosis will be further complicated by the need to generate global models across environments but such models have been obtained for reliable prediction of chemistries in other situations.

1. Introduction

Gene flow from forest plantations may have negative consequences for the genetic diversity, structure, and composition of native tree populations (Arnold et al., 1999; Potts et al. 2003). Despite the observations that even low levels of gene flow from exotic plantations may constitute a risk (Barbour et al., 2008a), there has been little assessment of the scale and impact of gene flow, and little monitoring continues to be undertaken (Laikre et al., 2010). This, in part, may be due to the difficulty and costs associated with the screening of large numbers of individuals needed to detect low levels of gene flow and often a lack of cost-effective reliable hybrid detection tools (Barbour et al., 2008b).

The spotted gums (Genus *Corymbia* Section *Maculatae*) are the major hardwood plantation taxon in subtropical Australia (Lee, 2007; Lee et al., 2009, 2010; Nichols et al., 2010). The four taxa (*Corymbia citriodora* subsp. *citriodora* (CCC), *C c. variegata* (CCV) and *C. henryi* (CH)) occur naturally as a replacement series that extends along the east coast from Cairns (QLD) in the north, to Bega (NSW) in the south (Hill and Johnson, 1995). The hybrid between spotted gum and *Corymbia torelliana* (Section *Torellianae;* CT), a native of tropical north Queensland (QLD), is also of interest to forestry and is being planted experimentally. *Corymbia torelliana* and spotted gum hybridise naturally in north Queensland where they co-occur, and controlled crossing experiments show that they can be easily crossed in both directions and have early viability (Hill and Johnson, 1995; Dickinson et al., 2010, 2012, 2013).

In addition to its use in forestry, CT has been widely planted as an amenity tree, but in recent years this has declined as CT was recognised as a weed in some shires of northern New South Wales (Forests NSW) and southern QLD (Hill and Johnson, 1995; NCWAC, 2003; Kingston et al., 2004). Plantings of CT stock give rise to two situations where gene flow monitoring from plantings may be warranted. Firstly, where CT is planted as an amenity tree near native spotted gum because there is potential for introgression of CT into native forest gene pools. And second, where the *Corymbia* F_1 hybrid (CT × spotted gums, F_1, sensu Lee (2007)) is planted near native spotted gum, and hence there is potential for backcrossing and further introgression of CT into native gene pools. This introgression

* Corresponding author. Tel.: +61 2 6620 3412; fax: +61 2 6620 3398/3442.
E-mail address: mervyn.shepherd@scu.edu.au (M. Shepherd).

may be undesirable because the same attribution of adaptability and vigour, that make the hybrids attractive for plantations, may also contribute to the ability of the hybrids to invade novel and disturbed environments (Lee, 2007; Lee et al., 2009; Abasolo et al., 2012; Shepherd et al., unpublished).

Near infrared (NIR) spectroscopy has been widely used to predict chemical compositions in agricultural crops and to predict the wood chemistry of trees (e.g. Jones et al., 2006; Poke and Raymond, 2006; Maranan and Laborie, 2008; Meder and Meglen, 2012). NIR has also been shown to have the potential to evaluate the nutritional value of forests in ecological studies (Foley et al., 1998; Stolter et al., 2006; Meder et al., 2007) and for selection of wood quality in tree breeding programmes (Schimleck, 2007; Meder et al., 2011; Meder and Schimleck, 2011). To a lesser extent, NIR has also been used for hybrid identification in plants but its potential in this regard, as far as these authors are aware, has been restricted to *Eucalyptus* and *Betula* (Atkinson et al., 1997; Humphreys et al., 2008). The main value of near infrared spectroscopy in hybrid screening is where it is difficult to identify hybrids by morphology (Tibbits, 1988; Abasolo et al., 2012) but it may offer other advantages such as reduced laboratory time and non-destructive sampling relative to other methods (Foley et al., 1998; Ebbers et al., 2002; Humphreys et al., 2008).

The principle behind NIR is similar to that of our human perception of colour in the visible light region (400–700 nm) of the electromagnetic spectrum. We see an object as being green, because all the other wavelengths of light, except green (wavelength ca. 510 nm), are absorbed by the object and only green is reflected back to our eyes. In the same manner, infrared radiation (700 nm–1 mm) incident on an object is either absorbed or reflected at various wavelengths. Near infrared spectrometers operate in the near and short wavelength range of the infrared region, typically between ca. 800 and 2500 nm dependent on the detector technology. Absorption of incident energy occurs due to interaction of the incident radiation with the variety of molecular bonds present in the object matrix, such as the stretching and bending modes of bonds between carbon, hydrogen, oxygen, nitrogen and other elements present in organic materials (Givens et al., 1997). It is the variety in these bond arrangements (e.g. the C–H bond of a methylene differs slightly from the C–H bond of a methyl and is completely different to a C–O bond) that in turn gives rise to differences in infrared absorption so that different chemical constituents will have different absorbance in the NIR spectra (Foley et al., 1998; Tsuchikawa, 2007; Schwanninger et al., 2011).

One of the limitations that must be addressed in the application of NIR to predict chemical composition (or in taxonomic classification), however, is the requirement for calibration and validation of an empirical chemical model (Foley et al., 1998; Richardson et al., 2004). A global model, based on a large sample size is ideal because such a model should be broadly applicable (Wallis and Foley, 2003). For some applications such as varietal identification, the NIR spectral differences among samples are often tested against a previously developed global model, for example in coffee (Bertrand et al., 2006; Posada et al., 2009) and oranges (Cen et al., 2007). It can be difficult to identify a particular compound that explains the difference between samples as the NIR profile corresponds to a functional group (Foley et al., 1998), in which the compounds comprising the functional group have overlapping NIR spectra (Curran, 1989). Global models exist for limited wood properties, specifically the Kraft pulp yield of hardwoods (Downes et al., 2009, 2010, 2011) and softwoods (Hodge and Woodbridge, 2010). Addressing these points is necessary for a wider application of NIR (i.e. application in different environments and different ages) for hybrid identification without the need of a reference population growing in the same environment and the same age as the taxon to be classified.

In this study, we examined the potential of NIR analysis to resolve among genealogical classes by comparing the NIR spectra of standardised tissue material of *Corymbia* hybrids and their parental taxa. Specifically, three factors that may have effect on NIR profile were compared; age of seedlings, tissue preparation and the influence of NIR instrument wavelength range. The effects of these three factors on the ability of NIR to discriminate between taxa were determined.

2. Methods

2.1. Materials and experimental design

The materials used in this study were from two section of the genus *Corymbia*. The first was from section *Maculatae* (i.e. the spotted gums) represented by three of the four recognised taxa (CCV, CCC and CH; *Corymbia maculata* was not studied) and the second was from section *Torellianae* (CT) (Hill and Johnson, 1995; Parra-O. et al., 2009).

The seedlings used were sown and grown at the Department of Agriculture, Fisheries and Forestry glasshouse facility at Gympie, QLD as detailed in Abasolo et al. (2012). The parental taxa, CCC, CCV, CH, and CT were represented by 10 open-pollinated families of: CCC (ex Kirrima, QLD [17°39′S, 146°5′E] and Yeppoon, QLD [23°07′S, 150°44′E]); CCV (ex Richmond Range, NSW [28°40′S, 152°42′E] and Woondum, QLD [26°15′S, 152°49′E]); CH (ex Lockyer, QLD [27°30′S, 152°04′E], Nerang, QLD [27°59′S, 153°20′E] and Myrtle, NSW [29°08′S, 152°05′E]); CT (ex Helenvale, North Queensland [15°43′S, 145°14′E] and a landrace from the Gympie region [no latitude and longitude available but molecular studies indicate that this material was from Kuranda, QLD (16°5′S, 145°36′E) (McVey, 2004). The parental taxa were open-pollinated thus, it is possible that hybrids were inadvertently included in the experiment, especially for species that occur in sympatry (Abasolo et al., 2012).

The hybrid taxa studied were from controlled pollinations and consisted of outbred F_1 interspecific hybrids (F_{1H}, F_{1V}, F_{1C}), and outbred backcross to CT (BCT). Where possible, parental control populations were composed of open-pollinated seedlots of trees from the same provenance (or nearby provenances) as the parents used for hybrid crosses but they were not the same individuals (data not shown; see Abasolo et al. (2012)). All analyses were based on hybrid groups as follows: (1) CT, F_{1C}, CCC for the CCC group; (2) CT, F_{1V}, BCT and CCV for the CCV group; and (3) CT, F_{1H}, CH for the CH group. The number of families in the hybrid taxa was lower and more variable than the parental taxa as fewer hybrid families were available for study (Table 1).

The nursery experiment was designed and analysed as Incomplete Block Design, with a layout generated with CycDesigN Version 1.2 software (Whitaker et al., 2006). It consisted of five replicates of nine blocks, each block comprising 24 individuals, arrayed in the block as four rows of two plots of three individuals per family (Total layout nominally accommodated = 1080 individuals).

2.2. Collection of NIR spectra

Whole, intact, fresh leaves of each individual seedling were scanned using a handheld near infrared spectrometer (Polychromix Phazir Model 1018, now Thermo Scientific, Tewksbury, MA, USA, www.thermoscientific.com) operating in the nominal wavelength range (950–1800 nm) on four- and eight-month-old seedlings.

Thawed leaf samples of eight-month-old seedlings were scanned using the original and a second handheld NIR (microPhazir RX, Model 1624, Thermo Scientific – Portable Optical Analysis