GENETICS

Supporting Information
http://www.genetics.org/cgi/content/full/genetics.109.106591/DC1

Identification of a *Cis*-Acting Regulatory Polymorphism in a Eucalypt *COBRA*-Like Gene Affecting Cellulose Content

Bala R. Thumma, Bronwyn A. Matheson, Deqiang Zhang, Christian Meeske, Roger Meder, Geoff M. Downes and Simon G. Southerton

DOI: 10.1534/genetics.109.106591

FILE S1

CELLULOSE AND KRAFT PULP YIELD (KPY) PREDICTIONS BASED ON NIR SPECTROSCOPY

Cellulose content was predicted using near infrared reflectance (NIR) spectroscopy described in (SCHIMLECK *et al.* 2004; HAMILTON *et al.* 2008)

Determination of cellulose content

Near infrared reflectance (NIR) spectra were measured on ground wood samples from the cores. The NIR spectra were measured on the wood meal from each core in diffuse reflectance mode in a scanning spectrometer (NIR Systems Inc., Model 5000). A ceramic standard was used as the instrument reference. Spectra were collected at 2-nm intervals over the 1100–2500 nm wavelength range. Fifty scans were acquired per sample and the results were averaged. The Vision® software was used to convert the data to the second-derivative mode using a segment width of 10 nm and a gap width of 0 nm.

Cellulose calibrations

Cellulose content for each wood-meal sample was predicted from the NIR spectra using a previously developed NIR calibration. Twenty nine samples (representing a wide range of NIR predicted cellulose contents) from Meunna were selected for analysis of crude cellulose content. Chemical assays were done on these samples according to the diglyme method of WALLIS *et al.* (1997): NIR calibrations were done according to the procedures outlined by SCHIMLECK *et al.* (2004): Calibrations were developed within the Vision® software (version 2.51) using partial least-squares (PLS) regression with four cross validation segments and a maximum of 10 factors (vectors) as described in SCHIMLECK et al (2004).

The degree of fit of the NIR calibration to the chemical assay data was measured by the standard error of calibration (SEC) (WORKMAN 1992). The calibration resulted in a coefficient of determination (R^2) of 0.89 and a SEC of 0.65 for the calibration set. A further 10 samples (not included in the calibration sets) representing a wide range of NIR predicted cellulose contents were selected for chemical assay to test the predictability of NIR spectra analysis compared to chemically assayed cellulose content. High coefficient of determination (R^2, 0.91) and relatively small standard error of prediction (SEP, 0.97) indicated that the NIR calibrations could be used to accurately predict cellulose content. The calibration developed for Meunna samples was applied to the Ridgely population to estimate cellulose content.

Kraft pulp yield (KPY) predictions in both populations were based on calibration (R^2, 0.89; SEC, 0.59) obtained for the Gog site (about 41°S) in northern Tasmania (SCHIMLECK et al. 2005).

There are 420 families within the trial at Meunna. Initially we collected two wood cores from 300 trees for NIR and SilviScan analysis in 2002. Cellulose and pulp yield values were predicted from ground wood core samples as discussed above. The trait values from 300 trees were used in the initial analysis of SNP-trait associations (Table 3).

From the rest of 120 families, we collected one wood core for NIR analysis in 2006. Intact cores were used for NIR spectra analysis. Spectra were collected from increment cores each 5mm radially. All spectra were collected using a Bruker MPA FT-NIR instrument and calibrations generated using the OPUS Quant software. Cellulose content in increment core samples was determined using a calibration developed within the CRC Forestry, containing 722

samples of predominantly plantation *E. globulus* and *E. nitens* from a range of sites and age classes across southern Australia. The coefficient of determination for the 8 factor calibration was 0.89 (SEC = 0.76) and for cross validation, 0.87 (SECV = 0.77). KPY in cores was predicted using a calibration (Downes *et al* 2007) based on 728 samples, representing over 40 different species of eucalypt. The coefficient of determination for the 9 factor calibration was 0.89 (SEC = 1.52) and for cross validation, 0.88 (SECV = 1.56). To make the data from two samplings comparable, we analysed the data by removing the last four years from extra 120 samples. As cellulose and pulp yield data from the two samplings were predicted using different calibration curves, we normalised the data from both samplings by adjusting means to zero and variance to one, so that data from both samplings can be analysed together. This data was used in the final analysis of SNP7 associations (Table4).

References:

DOWNES G.M., CATELA F. and MEDER R. (2007) Developing and evaluating a Global NIR Calibration for the prediction of Kraft Pulp Yield in Eucalypts. IUFRO Group 2.08.03 "Improvement and Culture of Eucalypts", Keynote speaker, Durban, South Africa 22-26 October 2007

HAMILTON, M.G., RAYMOND, C.A., HARWOOD, C.E. and POTTS, B.M. (2008) Genetic variation in *Eucalyptus nitens* pulpwood and wood shrinkage traits. *Tree Genetics and Genomes*, 10.1007/s11295-008-0179-4

Schimleck, L.R., Kube, P.S. and Raymond, C.A. (2004) Genetic improvement of kraft pulp yield in *Eucalyptus nitens* using cellulose content determined by near infrared spectroscopy. *Canadian Journal of Forest Research*, **34**, 2363-2370.

SCHIMLECK, L., KUBE, P., RAYMOND, C., MICHELL, A. and FRENCH, J. (2005) Estimation of whole-tree kraft pulp yield of Eucalyptus nitens using near-infrared spectra collected from increment cores. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere*, **35**, 2797-2805.

Wallis, A.F.A, Wearne, R.H. and Wright, P.J. (1997) New approaches to the rapid analysis of cellulose in wood. In: Proceedings of the 9th International symposium on wood and pulping chemistry. Canadian Pulp and Paper Association, Montreal, Quebec, pp C3-1–C3-4.