Forest&Wood Products Australia
Knowledge for a sustainable Australia

RESOURCES

PROJECT NUMBER: PNC220-1011

JANUARY 2014

Utility of molecular breeding in forestry

This report can also be viewed on the FWPA website
www.fwpa.com.au
FWPA Level 4, 10-16 Queen Street,
Melbourne VIC 3000, Australia
T +61 (0)3 9927 3200 F +61 (0)3 9927 3288
E info@fwpa.com.au W www.fwpa.com.au

Utility of molecular breeding in forestry

Prepared for

Forest & Wood Products Australia

By

R.J. Kerr, G.W. Dutkowski, T.A. McRae, S. Southerton, B. Thumma and B. Tier

Publication: Utility of molecular breeding in forestry

Project No: PNC220-1011

This work is supported by funding provided to FWPA by the Australian Government Department of Agriculture, Fisheries and Forestry (DAFF).

ISBN: 978-1-921763-90-8

Researcher/s:

Southern Tree Breeding Association Inc: **Thomas A. McRae**
PO Box 1811, Mount Gambier, SA 5290
PlantPlan Genetics Pty Ltd: **Richard J. Kerr and Gregory W. Dutkowski**
Hobart, Tasmania
CSIRO Plant Industry: **Bala Thumma and Simon Southerton**
Canberra, ACT
Animal Genetics and Breeding Unit: **Bruce Tier**
Armidale, NSW

Final report received by FWPA in April 2012

Forest & Wood Products Australia Limited
Level 4, 10-16 Queen St, Melbourne, Victoria, 3000
T +61 3 9927 3200 F +61 3 9927 3288
E info@fwpa.com.au
W www.fwpa.com.au

Executive Summary

This study demonstrates for the first time in an operational tree improvement program that molecular information recorded at the DNA level can make an incremental contribution to genetic and economic gain when coupled with quantitative tree breeding methods. The aim of the project was to investigate the integration of molecular data into routine genetic evaluation of plantation species, in this case *Eucalyptus nitens* commonly known as Shining gum. Shining gum is an important plantation species grown in Tasmania and some areas of Victoria.

Project collaborators (Gunns Limited and Forestry Tasmania) have been running conventional tree improvement programs for Shining gum for more than three decades. The objective of the programs has been to improve tree characteristics which are important commercially, such as growth rate, tree form, wood and fibre quality, disease and pest tolerances. Good progress has been made with current seed orchards and planting stock established with advanced generation genetic material having much improved characteristics. Data from measurements on trees in more than 150 field trials has been used to identify elite genotypes or trees with the best combination of genes for commercially important traits.

The objectives of this project were to: (i) integrate molecular information with phenotypic data for Shining gum in genetic evaluation; (ii) determine the utility or cost benefit of molecular information (changes in DNA code sequence) in improving genetic gain for an economic objective based on multiple traits; and (iii) disseminate the results and findings to the Australian forestry industry.

The genotypic data used in the project was generated by CSIRO Division of Plant Industry as part of the Hottest 100 project funded by FWPA. This project discovered a number of single nucleotide polymorphisms (SNP) that account for a small proportion of the genetic variance observed in wood quality traits. SNP-trait associations have the potential to increase the accuracy of selection in breeding and deployment. Molecular based selection criteria (MBSC) were generated for a set of 92 parent trees in seed orchards which were independent of trees assessed in the initial discovery and validation populations. The MBSC are functions of the SNP allele effects and the frequencies of the SNP alleles in the population.

The molecular data (MBSC) were then incorporated into a typical TREEPLAN analysis that uses phenotypic performance data and pedigree information on 69,029 trees in 27 field trials already residing in the DATAPLAN database for this species. The results are compared to see if the molecular information has led to any further increase in gain for the particular trait affected. Changes in the overall combined economic or multi-trait selection indices are then used to gauge the benefit of pursuing the additional molecular information.

The results are promising with the molecular data, when integrated with phenotypic data, improving the overall quality of the group of deployment clones selected for seed production in an orchard.

Selecting the best 30 genotypes (and roguing the others) based on phenotypic data alone among a group of 92 available in orchards for use as seed parents resulted in a predicted marginal return of $554 net present value per hectare compared with deploying unselected baseline seed. By integrating the molecular based selection criteria in the TREEPLAN evaluation, an additional 10 percent or incremental gain of $55 NPV per hectare can be achieved. This gain is the marginal gain or extra profit due to the deployment of improved genetics. Although the group of orchard parents used for the comparisons are not the best candidates identified by TREEPLAN in the population it demonstrates the utility of the approach.

Table of Contents

Executive Summary i

Introduction 3

Aim 3

Objectives 3

Background 3

Discussion 5

The extent to which the Research Project Objectives were achieved 8

Description of outputs achieved 9

Details of all Intellectual Property Rights and Confidential Information created in the course of carrying out the Research Project 10

Commercial implementation of the results 10

Difficulties encountered 10

Researcher's recommendations for any further research or actions needed to further assist commercial exploitation of the results of the Research Project 11

References 13

Acknowledgements 15

Researcher's Disclaimer 15

Introduction

Aim

The aim of the project is to investigate the integration of molecular data into routine genetic evaluation of plantation species such as *Eucalyptus nitens*.

Objectives

- Integrate molecular based selection criteria with phenotypic data for *E. nitens* in a species wide TREEPLAN analysis.
- Determine the utility of molecular information in improving genetic gain for an economic objective based on multiple traits.
- Disseminate results and findings to the Australian forestry industry.

Background

Genetic improvement of planting stock can lift productivity and increase economic returns in plantation forestry. Past gains made in tree improvement programs have been based exclusively on traditional methods of quantitative genetics. That is, the predicted genetic merit of a tree is based on its own performance as an individual, or parent, and that of its relatives, for traits measured in field trials grown under commercial conditions. Use of restricted maximum likelihood (REML) to estimate variance components, which are then used in best linear unbiased prediction (BLUP) to predict genetic values has become the paradigm in genetic evaluation, much like what had occurred in livestock improvement. Kerr et al. (2012 in preparation) describe a BLUP-based genetic evaluation system that was specifically designed to integrate forest tree data from hundreds of historical and current progeny trials. The system is now being used to routinely update genetic values on a program-wide basis in several temperate species (*Pinus radiata*, *Eucalyptus globulus* and *E. nitens*) grown commercially in Australia.

The aim of the present study was to investigate the integration of molecular data into routine genetic evaluation of one of these species, *Eucalyptus nitens*, commonly known as Shining gum. Shining gum is an important plantation species grown in Tasmania and some areas of Victoria. Study collaborators (Gunns Limited and Forestry Tasmania) have been running conventional tree improvement programs for Shining gum for more than three decades. Data from measurements on trees in more than 150 field trials have been integrated into a single, multivariate analysis, which predicts genetic values for approximately 35 traits. Values for these 35 traits are summarised into values for a reduced set of breeding objective traits that have a direct economic impact. An economic index is then used to combine information on breeding objective traits into a single value that is indicative of the net present value of the genotype for deployment. Selection decisions are usually based on the index value.

The molecular data arose from an association genetics study to examine allelic variation in genes influencing kraft pulp yield (KPY) and cellulose content (CC). Approximately 100 cell wall genes had been selected based on previous research (Thumma et al. 2005, 2009) and literature searches. Full-length gene sequences of these candidate genes were obtained by gene walking. Single nucleotide polymorphisms (SNP) were identified by high throughput sequencing in unrelated individuals. Preliminary analyses have determined that at least six SNP are consistently associated with variation in KPY and CC in at least two validation samples, and may explain in excess of 10% of the total additive genetic variation.

The proposed mechanism for integrating the SNP data is to form molecular-based observations, which are a function of the estimated allelic effects and allele frequencies. The molecular-based trait is included in the multivariate analysis by treating it as if it were a conventional (field-based) trait, with a genetic and error variance, and genetic co-variances with other traits. Predicted genetic values for molecular- and field-based traits reduce to a prediction of genetic values for breeding objective traits, which in turn are multiplied by their economic weights to form a prediction of the index value. This mechanism for integration of molecular data guarantees that selection pressure is not disproportionately used up on increasing the frequency of the favourable allele at individual loci.

It has been suggested that markers should be used by industry to screen existing clonal seed orchards to identify inferior genotypes for culling and to screen putative genotypes entering the orchards. This is a valid proposal but should be achieved within the framework of a genetic evaluation system that combines all information. Many quantitative traits affect profitability and quantitative trait variation is most likely caused by a very large number of loci, each with small effect. Unless SNP can be discovered that can account for all the variation, for all traits with an economic impact, culling or screening only on the basis of genotype statuses for SNP marker is likely to reduce profitability.

In summary the objectives of this study are to: describe how SNP marker data can be combined to form a molecular-based phenotype; describe how the molecular-based phenotypes are incorporated into routine genetic evaluation; and to determine the utility or cost benefit of molecular-based phenotypes in improving genetic gain for an economic index based on multiple traits. As part of the last objective we will be including a scenario where genotypes are screened for entering a seed orchard on the basis of genotype status for SNP loci only.

Discussion

Neale and Savolainen (2004) proposed candidate-gene-based association studies as a method of choice for identifying important alleles affecting quantitative traits in forest tree species. They argue association studies will lead to more efficient methods of marker-assisted breeding. However to date there have been no reports of the practical application of the results of association studies in forest tree breeding. Resende et al. (2012) suggest that association studies have yet to reveal loci that explain a sufficient proportion of total heritable variance, before results can have practical utility. Wegrzyn et al. (2010) identified 13 single-marker associations in nine candidate genes that were significantly associated with lignin composition in poplars. Each marker explained a small proportion of the phenotypic variance ranging from 1.2% to 3.8%. Gonzales-Martinez et al. (2010) identified multiple SNP-trait associations in loblolly pine affecting earlywood specific gravity. They temper the claim of 40% of the additive variance explained with the need to undertake further validation across different populations and field-testing environments.

The results of the present study suggest that a panel of SNP that explains as little as 8% of the total additive variance for a selection criteria trait has practical benefit within the context of screening genotypes for use in deployment. This statement, though encouraging, must also be considered with some caution. The population samples used in this study do not adequately represent the complete *E. nitens* population. Thus the markers identified in this study should only be used for screening and selecting genotypes belonging to the sub-races represented. More validation is required if screening is to be attempted in other sub-races.

A significant aspect of this work has been to specify procedures for the incorporation of molecular data into routine genetic evaluation. The Southern Tree Breeding Association (STBA) has opted to develop a comprehensive genetic evaluation system (TREEPLAN) that is capable of analysing in excess of a hundred traits, measured across potentially hundreds of progeny trials using a multi-trait and multi-site BLUP model customised for forest tree data. The approach taken is to construct molecular-based phenotypes, which assign to molecular-based traits that are included in the BLUP analysis. The approach requires no substantial changes to current BLUP software and is in a sense removed from the QTL discovery process.

Will this approach remain relevant as advances are made in the areas of genomics and statistical methodologies? To address this question it is convenient to summarise the process into a series of four steps:

a) Discovery, estimation and validation of molecular effects
b) Prediction of a molecular phenotype
c) Validation of the molecular phenotype
d) Incorporation of the molecular phenotype into regular or routine evaluation

We note here that use of "genomic" or "molecular" is interchangeable and both are used to refer to any information gathered at the DNA level. Step a) is generally the domain of dedicated institutions or research groups with expertise in molecular research. There is expanding array of options for estimating effects. In the present study simple linear regression sufficed because there were a small number of SNP located in targeted candidate genes and the discovery and validation population samples were free of any significant structure. If the whole genome has been covered by a dense SNP set and/or population stratification existed then more complex procedures and models are required. On the one hand there are methods that enforce model sparsity in keeping with the assumption that most loci have a weak or no effect on the phenotype. There are variable selection methods such as Bayes B (Meuwissen et al. 2001) or Bayesian shrinkage methods (e.g. Xu, 2003; Mutshinda and Sillanpaa, 2011). Alternatively, there are models that do not enforce sparsity and effectively divide the whole genome into a series of independent segments. Each segment acts as a proxy for the sum of the effects of QTL it carries. The additive genetic variance can be divided evenly among all segments and genomic effects are predicted using BLUP. This procedure is often referred to as genomic BLUP or GBLUP. Step b) is needed to convert marker locus effects into parameters from standard quantitative genetics theory. When locus effects are considered fixed as was the case in the present study the additive variance stems from random sampling of genotypes (Gianola et al. 2009) and assuming complete LD between the marker and the QTL. If Bayesian methods are used Gianola et al. (2009) highlight some problems relating marker effects to additive genetic variance and provide expressions for their conversion. Under GBLUP Resende et al. (2011) give an expression to predict the genomic breeding value (GBV), which is essentially what we have termed a molecular-based phenotype. Step c) is required for determining the recovered additive variance captured by the markers and for checking the accuracy of the molecular-based phenotype. We have suggested a bi-variate model (6) that analyses the molecular-based phenotype together with its corresponding field-based phenotype under certain constraints. Resende et al. (2011), when using GBLUP have suggested cross-validation schemes where individuals within the validation and training sets are interchanged repeatedly.

We have purposely made step d) our primary focus, since it has appeared not to have been given much attention by forest geneticists up to this point. Our interest in step d) stems from a conviction that the breeding goal consists of numerous traits, and given that markers will unlikely capture all the genetic variance for all traits in the breeding goal, conventional pedigree and phenotypic data is still required. A multi-trait individual tree model is our choice because it can account for changes in trait variances due to previous selection on correlated traits (Kennedy and Sorenson, 1988). In dairy cattle breeding genetic evaluation that incorporates whole genome information is currently implemented as a multistep procedure. Routine evaluation without genomic information is completed first. Genomic effects are then estimated using adjusted phenotypes arising from the first step, usually for a small, elite group of bulls. Finally genomic breeding values are estimated using a selection index. Misztal et al. (2009) have suggested a unified procedure based on modifying the numerator relationship matrix, **A**, to account for genomic information. The resultant matrix

expresses the actual realised proportion of the genome that two individuals share, when those individuals have genomic information, but expresses the expectation of the proportion otherwise. We are currently investigating the use of this approach in our multi-trait BLUP model, in anticipation that whole genome SNP sets will become available in the near future. The approach, in the context of our system, will negate the need to split selection criteria traits in field-based and molecular-based.

In the meantime the current system for incorporating molecular data will be used to apply the results of numerous association studies that are currently being undertaken in all species that are of interest to the STBA. Improved genetic gain, hence profitability is likely to increase as a result.

The extent to which the Research Project Objectives were achieved

The project was successful in achieving the objectives.

The current study is a milestone achievement because it is the first time we have been able to successfully demonstrate added value from the use of molecular (SNP) data in an operational tree improvement program. That is, the molecular data can influence the choice of genotypes (parents) used in a seed orchard. It shows that for the first time after decades of research that molecular data can be used to complement conventional tree breeding methods based on quantitative genetics and phenotypic data. It also highlights the utility of the TREEPLAN genetic evaluation system, considering all data (molecular and phenotypic) and full pedigree in an industry wide analysis.

Integrate molecular based selection criteria with phenotypic data for E. nitens in a species wide TREEPLAN analysis.

The project was successful in integrating molecular data into routine genetic evaluation of *Eucalyptus nitens*, which is an important plantation species grown in temperate Australia. Molecular based selection criteria were generated using data from the Hottest 100 project. Phenotypic data associated with genotypes (trees) tested in 151 field trials (or orchards) currently resides in the DATAPLAN system. The trials have been established and trees assessed over decades of tree improvement by nine different companies and/or agencies. A restricted subset of data from 27 of these trials was used in the comparative TREEPLAN analysis done to integrate molecular phenotypes.

Determine the utility (cost benefit) of molecular information in improving genetic gain for an economic objective based on multiple traits.

Proprietary bio-economic models of Gunns Limited and Forestry Tasmania are used for TREEPLAN evaluations in *E. nitens*. Economic indices based on these models were used to value the differences in genetic merit of different genotypes in the populations. This approach allowed us to make an objective appraisal (in $ net present value per ha) of the merit of incorporating additional molecular information into the evaluation. The project has successfully achieved this objective, but more testing and comprehensive evaluations are required to confirm the findings.

Disseminate results and findings to the Australian forestry industry.

The results and findings of the study were provided to Gunns Limited and Forestry Tasmania on line (subject to proprietary restrictions) for use in deployment programs. The molecular based selection criteria are available in DATAPLAN for further TREEPLAN evaluations as new phenotypic data is generated. Dr Kerr presented results and findings to industry personnel at several Technical Advisory Committee meetings in 2011. Representatives from

industry, including Gunns and Forestry Tasmania, participated. The importance and commercial relevance of the work was highlighted at an industry genome roundtable (sponsored by FWPA) held 26 July 2011 in Melbourne in association with the Botanical Congress. Dr McRae reported general findings at the recent AusTimber 2012 ForestWorks conference in Mount Gambier promoting the importance of genetics in modern forestry practices to industry. We plan to publish the innovative results in a leading journal, as well as in various industry media outlets in the coming months as part of a broader promotion of the benefits of genetics research.

The main message for senior executives of plantation grower and timber processing companies is that we have been able to successfully demonstrate added value with molecular (SNPs) data when incorporated with phenotypic data into routine genetic evaluation using TREEPLAN. Adding molecular information into the *E. nitens* analysis resulted in a $55 NPV per ha incremental improvement above the predicted marginal return of $554 NPV above baseline for a set of improved seed orchard parents identified using conventional tree breeding methods based on quantitative genetics and phenotypic data. That is, molecular information has the potential to increase marginal gain due to improved genetics (not total productivity) by a further 10 percent, based on the results from the study. Although the group of orchard parents used for the comparisons were not the best candidates (marginal gains of up to $1500 NPV for a kraft pulp objective) identified by TREEPLAN in the population, an incremental gain was achieved.

Description of outputs achieved

Molecular-based phenotypes were generated and validated using data from the Hottest 100 project.

Integrated TREEPLAN analysis was successfully completed using molecular- and field-based phenotypes associated with the genetic resources of *E. nitens* and the results stored in DATAPLAN.

Economic indices to compare the value of genetic material described using molecular and/or phenotypic information were generated and provided online to industry partners (Gunns Limited and Forestry Tasmania).

The current study is a milestone achievement because it is the first time we have been able to successfully demonstrate added value from the use of molecular (SNP) data in an operational tree improvement program. Marginal gain from using molecular based selection criteria was demonstrated through its impact on overall gain for the economic objective.

Details of all Intellectual Property Rights and Confidential Information created in the course of carrying out the Research Project

Background Intellectual Property

Background Intellectual Property made available for this project remains the property of the contributing party.

Genetic material and associated information (pedigree and tree measurement data) provided to the project remain the property (and confidential information) of the contributing party.

Molecular data and associated technologies made available for use in this project remain the property of the contributing party.

DATAPLAN®, TREEPLAN®, associated software and methods remain the property of the

Southern Tree Breeding Association Inc., PlantPlan Genetics Pty Ltd and partners.

Project Intellectual Property

Genetic values and economic indices generated using Background Intellectual Property will be provided on a confidential basis to the party contributing the Background Intellectual Property for their internal use.

CSIRO Plant Industry will own Intellectual Property connected with SNP-phenotypic trait associations. This includes knowledge of primer sequences and other relevant art pertaining to establishment of the associations.

Enhancements made to DATAPLAN®, TREEPLAN® and associated software as part of the project is owned by the Southern Tree Breeding Association Inc.

CSIRO Plant Industry and PlantPlan Genetics Pty Ltd each retain the Background Intellectual Property relating to the methods they have already developed prior to the project commencing.

Commercial implementation of the results

The molecular based selection criteria can now be routinely included in current and future TREEPLAN analyses.

Difficulties encountered

None.

Researcher's recommendations for any further research or actions needed to further assist commercial exploitation of the results of the Research Project

The amount of useful molecular data is currently very limited. It would be ideal to increase the number of associations such that the markers explain a greater proportion of the additive genetic variance observed for the trait. The Hottest 1000 project will hopefully address this issue.

Additionally, the SNP-trait associations were discovered using population samples that do not adequately represent the whole *E. nitens* resource. In reality, only two sub-races have been sampled. For greater applicability in international/national genetic evaluation analyses, samples representing the whole resource should be used in any future studies.

The STBA has been encouraged that the current project has helped to bridge the disconnection between molecular genetics research and operational breeding; and is optimistic about future collaboration. We have built species wide databases for the plantation industry which houses all phenotypic data, pedigree and associated information for *E. nitens* and other plantation species in DATAPLAN. The TREEPLAN system has the utility to routinely incorporate phenotypic and molecular based selection criteria into species wide genetic evaluations. It is essential molecular geneticists engage with these delivery platforms to ensure research results are adopted in routine tree improvement programs. Information in isolation is not very useful. In summary we recommend that personnel associated with these delivery platforms are represented in future molecular genetics projects that are targeting species bred by the STBA or by STBA affiliated companies.

Alternative approaches to the generation and integration of molecular data need to be explored. Firstly, there is the SNP discovery phase, followed by the detection of significant associations that can be validated across multiple population samples. The paradigm followed was essentially one of choosing the appropriate candidate genes and zeroing in on putative QTL in those genes. Given that linkage disequilibrium (LD) is so weak in forest tree populations the gamble is being able to hit on a SNP that is very close to the causative mutation, if not it actually being the causative mutation. The integration of the generated molecular data was then a process of summing up the effects of individual QTL into a molecular-based phenotype. The assumption being that each QTL (or small chunk of chromosome) is independent to the other QTL.

What if the paradigm was changed such that the whole genome could be effectively covered by SNP marker, rather than zeroing in on certain loci? This implies effectively dividing up the genome into a series of independent segments. The molecular geneticists would argue that given that LD is so weak, the segments would have to be very narrow and an enormous

number of SNP markers are needed. Given the rate at which sequencing technologies are advancing, we anticipate in the near future routinely genotyping trees for a vast number of SNP may not be such an insurmountable problem. Under such a paradigm, alternative strategies for integrating the molecular data, such as use of a genomic relationship matrix, would need to be explored. Given the current project had a limited budget we could not explore such strategies. Feasibility studies should be commenced that investigate the use of statistical techniques now adopted by the animal breeders, who currently routinely genotype elite animals for vast numbers of SNP markers.

There is a perceived lack of preparation for how marker genotypes are going to be stored on an industry wide basis. We anticipate DATAPLAN will be adequate for storing molecular based phenotypes, but not adequate as a repository for raw genotypes. There needs to be a lot more discussion regarding how large amounts of raw genotype information is going to be stored, converted to molecular-based phenotypes, and linked to DATAPLAN. This discussion should involve all groups with an interest in *E. nitens* breeding and deployment. That is, the plan must be formulated at a national or even an international level. Formal linkages between DATAPLAN and databases/repositories for raw genotype data need to be established.

References

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007). TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23 (19): 2633-2635

Churchill GA, Doerge RW (1994). Empirical Threshold Values for Quantitative Trait Mapping. Genetics 138 (3): 963-971

Falconer DS, Mackay TFC (1996). Introduction to Quantitative Genetics. Longman, England

Gianola D, de los Campos G, Hill WG, Manfredi E, Fernando R (2009). Additive genetic variability and the Bayesian alphabet. Genetics 183 (1): 347-363.

Gilmour AR, Gogel BJ, Cullis BR, Thompson R (2009). ASReml User Guide. Release 3.0. Hemel Hempstead, HP1 1ES, UK

Gonzalez-Martinez SC, Wheeler NC, Ersoz E, Nelson CD, Neale DB (2007). Association genetics in *Pinus taeda* L. I. Wood property traits. Genetics 175 (1): 399-409.

Hamilton MG, Raymond CA, Potts BM (2008). Short Note: The Genetic Correlation Between Air-dried Density and Basic Density in *Eucalyptus Nitens* Wood Cores. Silvae Genet 57 (4-5): 210-212

Kennedy BW, Sorenson DA(1988). Properties of mixed model methods for prediction of genetic merit. In: B.S. Weir EJE, M.M. Goodman, G. Namkong (ed) Second International Conference on Quantitative Genetics, Raleigh, NC, Sinauer Assoc. Sunderland, MA, pp 91-103

Kerr, RJ, Thumma, B, Southerton, S, Tier, B, Dutkowski, GW, McRae, TA and Harwood, C (2010). Operational testing of molecular breeding to improve pulpwood value in *Eucalyptus nitens*. (Confidential Client Report PPG10-04).

Kerr, RJ, Dutkowski, GW, Tier, B, Li, L and McRae, TA (2012). Implementing extensive multivariate BLUP analysis in forestry. In preparation.

Meuwissen TH, Hayes BJ, Goddard ME (2001). Prediction of total genetic value using genome-wide dense marker maps. Genetics 157 (4):1819-1829

Misztal I, Legarra A, Aguilar I (2009). Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. J Dairy Sci 92 (9): 4648-4655.

Mutshinda CM, Sillanpaa MJ (2011). Bayesian shrinkage analysis of QTLs under shape-adaptive shrinkage priors, and accurate re-estimation of genetic effects. Heredity 107 (5):405-412. **Neale DB, Savolainen O** (2004) Association genetics of complex traits in conifers. Trends Plant Sci 9 (7): 325-330.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81 (3): 559-575.

Resende MDV, Resende MFR, Sansaloni CP, Petroli CD, Missiaggia AA, Aguiar AM, Abad JM, Takahashi EK, Rosado AM, Faria DA, Pappas GJ, Kilian A, Grattapaglia D (2012). Genomic selection for growth and wood quality in Eucalyptus: capturing the missing heritability and accelerating breeding for complex traits in forest trees. New Phytol 194 (1): 116-128.

Schimleck L, Kube P, Raymond C, Michell A, French J (2005). Estimation of whole-tree kraft pulp yield of *Eucalyptus nitens* using near-infrared spectra collected from increment cores. Can J Forest Res 35 (12): 2797-2805. **Schneeberger M, Barwick SA, Crow GH, Hammond K** (1992). Economic indices using breeding values predicted by BLUP. J Anim Breed Genet 109 (1-6): 180-187.

Southerton SG, MacMillan CP, Bell JC, Bhuiyan N, Downes G, Ravenwood IC, Joyce KR, Williams D, Thumma BR (2010). Association of allelic variation in xylem genes with wood properties in *Eucalyptus nitens*. Aust Forestry 73 (4): 259-264

Team RDC (2005). R: A language and environment for statistical computing, reference index version 2.2.1. R Foundation for Statistical Computing, Vienna, Austria

Thumma BR, Matheson BA, Zhang D, Meeske C, Meder R, Downes GM, Southerton SG (2009). Identification of a Cis-acting regulatory polymorphism in a Eucalypt COBRA-like gene affecting cellulose content. Genetics 183 (3): 1153-1164.

Thumma BR, Nolan MF, Evans R, Moran GF (2005). Polymorphisms in cinnamoyl CoA reductase (CCR) are associated with variation in microfibril angle in *Eucalyptus spp*. Genetics 171 (3):1257-1265.

Thumma B, MacMillan C, Southerton S, Williams D, Joyce K and Ravenwood I (2010). Accelerated breeding for high pulp yield in *E. nitens* using DNA markers identified in 100 cell wall genes: The Hottest 100. FWPA report PNC052-0708

Wegrzyn JL, Eckert AJ, Choi M, Lee JM, Stanton BJ, Sykes R, Davis MF, Tsai CJ, Neale DB (2010). Association genetics of traits controlling lignin and cellulose biosynthesis in black cottonwood (*Populus trichocarpa*, *Salicaceae*) secondary xylem. The New phytologist 188 (2): 515-532.

Xu S (2003). Estimating polygenic effects using markers of the entire genome. Genetics 163 (2): 789-801

Acknowledgements

The authors gratefully acknowledge the financial assistance of Forest and Wood Products Australia Ltd, PlantPlan Genetics, Southern Tree Breeding Association, CSIRO, Forestry Tasmania, Gunns Ltd, and the Cooperative Research Centre for Forestry. The authors especially are grateful to Dr Chris Harwood, Kelsey Joyce and Dr Dean Williams for their interest and motivation.

Researcher's Disclaimer

Field performance of trees is influenced by management techniques, environmental conditions and the genetic composition of each tree. The authors and their employers do not guarantee the actual performance of trees in plantations based upon TREEPLAN genetic values due to other factors directly influencing the performance of each tree. TREEPLAN genetic values are also affected by the amount and quality of pedigree, measurement and genetic parameter information, and the models used. While due diligence will be made to use all this information appropriately, the derived genetic values will reflect any deficiencies.