

ADVANCES IN EUCALYPT GENETICS: FROM GENES TO ECOSYSTEMS

Brad Potts^{1,2}, Joao Costa e Silva³, Corey Hudson^{1,2}, Rebecca Jones^{1,2}, Matthew Hamilton^{1,2}, Greg Dutkowski^{1,4}, Julianne O'Reilly-Wapstra^{1,2}, Jules Freeman^{1,2}, Dorothy Steane^{1,2}, Joseph Bailey¹, Jennifer Schweitzer¹, Thomas Whitham⁵ and René Vaillancourt^{1,2}

¹School of Plant Science¹ and CRC for Forestry², University of Tasmania, Private Bag 55, Hobart, Tasmania, 7001, Australia.

³Centro de Estudos Florestais, Departamento dos Recursos Naturais, Ambiente e Território, Instituto Superior de Agronomia, Universidade Técnica de Lisboa, Tapada da Ajuda, 1349-017 Lisboa, Portugal.

⁴PlantPlan Genetics, PO Box 1811, Mount Gambier, South Australia, 5290, Australia

⁵Department of Biological Sciences, and the Merriam-Powell Center for Environmental Research, Northern Arizona University, Flagstaff, AZ 86011, USA.

The four years since our last meeting have seen enormous advances in eucalypt genetics on multiple fronts - from the genomic resources required to understand genome structure and function, through to the conceptualization and study of the effects of genes which extend beyond the individual to affect its neighbours, dependent biotic communities and even ecosystem processes. The International Eucalypt Genome Network's (EUCAGEN) 2006 proposal to the US Department of Energy's Joint Genome Institute (DOE-JGI) was successful and saw the public release of a *Eucalyptus* reference genome sequence in 2011 based on *E. grandis*, DOE-JGI's subsequent re-sequencing of the *E. globulus* genome and international co-operation in the development and public release of novel genomic and other resources (<http://web.up.ac.za/eucagen/>; Keller *et al.* 2009; Rengel *et al.* 2009; Neves *et al.* 2011; Paiva *et al.* 2011). As expected (Myburg *et al.* 2007), these initiatives have unleashed an international effort to explore the genome, and clearly placed *Eucalyptus* as a model forest tree system for genetic research.

One of the molecular tools recently developed for *Eucalyptus* is the Diversity Array Technology (DArT) (Sansaloni *et al.* 2010). With wide transferability across *Eucalyptus* species, DArT provides an unprecedented level of resolution for linkage mapping, detection of quantitative trait loci (QTL), comparative mapping, population genetic, phylogenetic and evolutionary studies with thousands of genome-wide markers of known position available (Hudson *et al.* 2011; Kullán *et al.* 2011; Steane *et al.* 2011). Automation of genotyping and the reduced cost per marker have facilitated the development of high-density linkage maps (Hudson *et al.* 2011; Kullán *et al.* 2011). The transferability of DArT across species and pedigrees has allowed the transfer of QTL information between studies and, in combination with an increasing resource of microsatellites (Faria *et al.* 2011) and single nucleotide polymorphisms (SNPs; Paiva *et al.* 2011), has provided a high-resolution link between phenotypes in field populations and the *Eucalyptus* reference genome. Eucalypts are diploid and the number of chromosomes appears to be uniform across the genus ($2n = 22$; Oudjehih and Bentouati 2006). DArT markers form the foundation of a high-resolution comparative mapping study that has demonstrated only minor structural differences amongst three species studied (*E. grandis*, *E. urophylla* and *E. globulus*; Hudson *et al.* 2011). This result argues for high overall genome similarity, and high transferability of the *E. grandis*

Potts BM, Costa e Silva J, Hudson C, Jones J, Hamilton M, Dutkowski G, O'Reilly-Wapstra J, Freeman J, Steane D, Bailey J, Schweitzer J, Whitham T, Vaillancourt R (2011) Advances in eucalypt genetics: From genes to ecosystems. In 'Joining silvicultural and genetic strategies to minimize *Eucalyptus* environmental stresses: from research to practice' (Eds Goncalves JL *et al.*) pp. 26-32. Proceedings of IUFRO Conference on Improvement and Culture of Eucalypts, Porto Seguro, Brazil. 16-18th Nov. 2011 (ESALQ, Piracicaba, Brazil). (invited talk)

genome sequence information and marker/trait associations amongst the lineages studied. At the population level, our molecular and quantitative genetic studies have focused on the temperate species *E. globulus*, which is part of a complex of four related taxa, variously given species or subspecies status (Potts *et al.* 2004; Potts *et al.* 2007). Using a multi-species consensus map, genome scans of population parameters derived from DArT markers have located genomic regions associated with species differentiation, diversity deserts and introgression (Hudson 2011; R. Jones unpubl.). The natural introgression studied was between *E. globulus* and a rare Tasmanian endemic, *E. cordata* (McKinnon *et al.* 2010; R. Jones unpubl.). Despite evidence of introgression of chloroplast DNA affecting a large proportion of the *E. globulus* gene-pool and porosity of the *E. globulus* nuclear genome, such introgression appears to have only a minor effect on the nuclear gene pools of the two species as they maintain their morphological and overall molecular distinctiveness when the two species co-occur. Insights into the post-zygotic barriers that would maintain species integrity in the face of hybridisation have come from studies of artificial first (Volker *et al.* 2008) and advanced (Costa e Silva *et al.* 2011b) generation hybrids between *E. globulus* and *E. nitens*. These two species are well delineated genetically and differentiating DArT markers have been located on 8 of the 11 linkage groups (Hudson 2011). Significant outbreeding depression occurs in their hybrids and line-cross analyses argue that the development of favorable epistasis within species is a key mechanism underlying their differentiation (Costa e Silva *et al.* 2011b). Given the relatively high genome homology observed to date, such epistasis is likely to be a key mechanism of speciation in *Eucalyptus*.

At lower taxonomic levels the distinction between recognized eucalypt taxa is often less clear and outbreeding depression less obvious. This is seen no better than in the relationship of the four taxa comprising the *E. globulus* complex. While the cores of these taxa are geographically, morphologically and, in most cases, genetically distinct, the phenotypic and molecular genetic variation between them appears continuous (Jones 2009). Intermediate populations occur over large geographical areas, and morphological and molecular affinities often do not correspond. Within *E. globulus* and its intergrades, there is increasing evidence for genetic differentiation in a wide variety of traits of ecological and economic significance, at both broad- (Stackpole *et al.* 2010) and local- (Foster *et al.* 2007) spatial scales. There is molecular and other evidence for contemporary and historic barriers to gene flow (Foster *et al.* 2007; Jones 2009; Jones *et al.* 2011), which could contribute to population divergence through drift. However, at this micro-evolutionary level there is increasing evidence that natural selection has played a significant role in shaping the geographic patterns of quantitative genetic variation observed in wood property (Hamilton *et al.* 2010; Stackpole *et al.* 2011), developmental (Hamilton *et al.* 2011), and stress-related (Costa e Silva *et al.* 2006; Tibbits *et al.* 2006; Dutkowski and Potts 2011) traits. A similar conclusion is reached from recent studies of other eucalypt species (Shepherd *et al.* 2010).

From a quantitative genetics perspective, understanding the potential for eucalypt populations to genetically adapt to diverse abiotic and biotic stresses (Myburg *et al.* 2007; Teulières *et al.* 2007; Guimarães *et al.* 2010), requires knowledge of the genetic architecture of the relevant traits, particularly their levels of additive genetic variance and co-variation. Of the wide variety of traits studied in *E. globulus*, the additive genetic control (and thus narrow-sense heritability) of growth and survival traits themselves is low and there is evidence for significant non-additive genetic control of growth under outcrossing which appears to be comparable to the additive genetic variation in some (Li *et al.* 2007; Volker *et al.* 2008; Callister *et al.* 2011), but not all populations (Costa e Silva *et al.* 2004, 2009). However, growth and survival are also subject to severe inbreeding depression, which may vary with

level of inbreeding, age, site, and population. At the population level, this inbreeding depression does not appear to be due to epistasis (Costa e Silva *et al.* 2011a), but rather to dominance variation associated with a genetic load of rare and recessive deleterious alleles (Costa e Silva *et al.* 2010a, b). For diameter growth, the dominance variation expressed under inbreeding was estimated to be nearly 10-fold greater than the dominance variance associated with random mating and the additive variance (Costa e Silva *et al.* 2010b). With self-incompatibility, outcrossing rates and inbreeding depression varying amongst female parents (Mimura *et al.* 2009; Costa e Silva *et al.* 2010b; McGowen *et al.* 2010), these adverse dominance effects expressed under inbreeding appear to be a major driver of differences in growth amongst open-pollinated families (Costa e Silva *et al.* 2010b; see also Bush *et al.* 2011). Not until later ages, when substantial mortality of inbred progeny has occurred, do the additive genetic differences between parents appear to be manifest in the standing genetic variation in growth of their open-pollinated families (Costa e Silva *et al.* 2010b).

A current research frontier lies in linking phenotypic variation to the genome through association genetics (Thumma *et al.* 2009; Sexton *et al.* 2010; Southerton *et al.* 2010; Thavamanikumar *et al.* 2011; Kulheim *et al.* 2011) or genome wide (Grattapaglia and Resende 2011) approaches. There is also a new frontier in linking genetics and ecology, through community and ecosystem genetics (Whitham *et al.* 2006; Whitham *et al.* 2008). This research looks beyond the direct genetic effects on the tree phenotype to study the flow-on effects on neighbours, the associated biota and even ecosystem processes. Such extended indirect genetic effects and feedbacks are especially important in forest trees, which dominate many natural and artificial terrestrial ecosystems, provide habitat for numerous dependent organisms, and supply many important ecosystem services (Whitham *et al.* 2006). We have been studying these extended genetic effects in *Eucalyptus globulus*, following natural colonization of base-population trials by local fungi, insect and marsupial species. These Australian trials have provided a robust experimental system in which to study individual species and community level responses to genetic variation in a foundation tree species. In one field trial, *E. globulus* trees from different races supported different arthropod and fungal communities in their canopy (Barbour *et al.* 2009c), trunk (Barbour *et al.* 2009b) and litter (Barbour *et al.* 2009a); and there was even differences in soil nitrogen levels (Bailey *et al.* 2012). There is the potential for complex feedbacks between tree genetics and dependent biotic communities, with reciprocal and cascading effects likely (Whitham *et al.* 2006). We have been studying the drivers of these biotic responses (Miller *et al.* 2011; O'Reilly-Wapstra *et al.* 2012), the genetic co-variance amongst dependent species, feedbacks to tree fitness (O'Reilly-Wapstra *et al.* 2012), potential correlated responses to selection and, more recently, the importance of indirect genetic effects in these artificial forests. Placing forest tree genetics into this broader community and ecosystem framework offers a more holistic approach to meet the challenges posed by human-induced global change and to sustainably manage the expanding area of artificial forests across the globe (Whitham *et al.* 2006; Whitham *et al.* 2010).

References

- Bailey J, *et al.* (2012) From genes to ecosystems: Emerging concepts bridging ecological and evolutionary dynamics. In 'The Ecology of Plant Secondary Metabolites: from Genes to Landscapes'. (Eds G Iason, *et al.*) pp. in press. (Cambridge University Press, Cambridge, UK)
- Barbour RC, *et al.* (2009a) A footprint of tree-genetics on the biota of the forest floor. *OIKOS* 118, 1917-1923.
- Barbour RC, *et al.* (2009b) Biodiversity consequences of genetic variation in bark characteristics within a foundation tree species. *Conservation Biology* 23, 1146-1155.
- Barbour RC, *et al.* (2009c) A geographic mosaic of genetic variation within a foundation tree species and its community-level consequences. *Ecology* 90, 1762-1772.

- Bush D, *et al.* (2011) Marker-based adjustment of the additive relationship matrix for estimation of genetic parameters - an example using *Eucalyptus cladocalyx*. *Tree Genetics & Genomes* 7, 23-35.
- Callister AN, *et al.* (2011) Genetic analysis of *Eucalyptus globulus* diameter, straightness, branch size, and forking in Western Australia. *Canadian Journal of Forest Research* 41, 1333-1343.
- Costa e Silva J, *et al.* (2004) Additive and non-additive genetic parameters from clonally replicated and seedling progenies of *Eucalyptus globulus*. *Theoretical and Applied Genetics* 108, 1113-1119.
- Costa e Silva J, *et al.* (2006) Genotype by environment interaction for growth of *Eucalyptus globulus* in Australia. *Tree Genetics and Genomes* 2, 61-75.
- Costa e Silva J, *et al.* (2009) Genetic parameters for growth, wood density and pulp yield in *Eucalyptus globulus*. *Tree Genetics & Genomes* 5, 291-305.
- Costa e Silva J, *et al.* (2010a) Effects of inbreeding on population mean performance and observational variances in *Eucalyptus globulus*. *Annals of Forest Science* 67, 605p1-605p9.
- Costa e Silva J, *et al.* (2010b) Genetic variation and parental performance under inbreeding for growth in *Eucalyptus globulus*. *Annals of Forest Science* 67, 606p1-606p8.
- Costa e Silva J, *et al.* (2011a) The effects of age and environment on the expression of inbreeding depression in *Eucalyptus globulus*. *Heredity* 107, 50-60.
- Costa e Silva J, *et al.* (2011b) Epistasis causes outbreeding depression in eucalypt hybrids. *Tree Genetics and Genomes* DOI: 10.1007/s11295-011-0437-8.
- Dutkowski G, *et al.* (2011) Drought tolerant *Eucalyptus globulus* provenances have thick bark, high basic density and grow poorly on wet sites. this volume
- Faria DA, *et al.* (2011) Genotyping systems for *Eucalyptus* based on tetra-, penta-, and hexanucleotide repeat EST microsatellites and their use for individual fingerprinting and assignment tests. *Tree Genetics & Genomes* 7, 63-77.
- Foster SA, *et al.* (2007) Parallel evolution of dwarf ecotypes in the forest tree *Eucalyptus globulus*. *New Phytologist* 175, 370-380.
- Grattapaglia D, *et al.* (2011) Genomic selection in forest tree breeding. *Tree Genetics & Genomes* 7, 241-255.
- Guimarães LMS, *et al.* (2010) Genetic control of *Eucalyptus urophylla* and *E. grandis* resistance to canker caused by *Chrysosporthe cubensis*. *Genetics and Molecular Biology* 33, 525-531.
- Hamilton M, *et al.* (2010) Genetic correlations between pulpwood and solid-wood selection and objective traits in *Eucalyptus globulus*. *Annals of Forest Science* Published online: 18 May 2010
- Hamilton MG, *et al.* (2011) The genetic variation in the timing of heteroblastic transition in *Eucalyptus globulus* is stable across environments. *Australian Journal of Botany* 59, 170-175.
- Hudson C, *et al.* (2011) High synteny and colinearity among *Eucalyptus* genomes revealed by high-density comparative genetic mapping. *Tree Genetics and Genomes* in press.
- Hudson CJ (2011) Eucalypt genomics: Linkage mapping, QTL analysis and population genomic studies using Diversity Arrays Technology. PhD thesis (submitted), University of Tasmania.
- Jones R (2009) Molecular evolution and genetic control of flowering in the *Eucalyptus globulus* species complex. PhD thesis, University of Tasmania.
- Jones R, *et al.* (2011) Genetic and environmental control of flowering time in *Eucalyptus globulus* ssp. *globulus*. *Tree Genetics and Genomes* in press.
- Keller G, *et al.* (2009) Development and functional annotation of an 11,303-EST collection from *Eucalyptus* for studies of cold tolerance. *Tree Genetics & Genomes* 5, 317-327.
- Kulheim C, *et al.* (2011) The molecular basis of quantitative variation in foliar secondary metabolites in *Eucalyptus globulus*. *New Phytologist*, doi: 10.1111/j.1469-8137.2011.03769.x.
- Kullan ARK, *et al.* (2011) High-density genetic linkage maps with over 2,400 sequence-anchored DArT markers for genetic dissection in an F2 pseudo-backcross of *Eucalyptus grandis* × *E. urophylla*. *Tree Genetics and Genomes* in press.
- Li Y, *et al.* (2007) The genetic architecture of a *Eucalyptus globulus* full-sib breeding population in Australia. *Forest Genetics* 12, 167-179.
- McGowen M, *et al.* (2010) Sources of variation in self-incompatibility in the Australian forest tree, *Eucalyptus globulus*. *Annals of Botany* 105, 737-745.
- McKinnon GE, *et al.* (2010) Recurrent nuclear DNA introgression accompanies chloroplast DNA exchange between two eucalypt species. *Molecular Ecology* 19, 1367-1380.
- Miller AM, *et al.* (2011) Field screening for genetic-based susceptibility to mammalian browsing. *Forest Ecology and Management* 262, 1500-1506.
- Mimura M, *et al.* (2009) Comparison of contemporary mating patterns in continuous and fragmented *Eucalyptus globulus* native forests. *Molecular Ecology* 18, 4180-4192.
- Myburg AA, *et al.* (2007) *Eucalyptus*. In 'Genome mapping and molecular breeding in plants.' (Ed. C Kole) pp. 115-160. (Springer: New York, NY, USA)
- Neves LG, *et al.* (2011) A high-density transcript linkage map with 1,845 expressed genes positioned by microarray-based Single Feature Polymorphisms (SFP) in *Eucalyptus*. *BMC Genomics* 12.
- O'Reilly-Wapstra J, *et al.* (2012) Selection for anti-herbivore plant secondary metabolites: a *Eucalyptus* system. In 'The Ecology of Plant Secondary Metabolites: from Genes to Landscapes'. (Eds G Iason, *et al.*) pp. in press. (Cambridge University Press, Cambridge, UK)
- Oudjehih B, *et al.* (2006) Chromosome numbers of the 59 species of *Eucalyptus* L'Herit. (Myrtaceae). *Caryologia* 59, 207-212.
- Paiva JAP, *et al.* (2011) Advancing *Eucalyptus* genomics: identification and sequencing of lignin biosynthesis genes from deep-coverage BAC libraries. *BMC Genomics* 12.

- Potts B, *et al.* (2007) Advances in reproductive biology and seed production systems of *Eucalyptus*: The case of *Eucalyptus globulus*. Eucalypts and Diversity: Balancing Productivity and Sustainability, IUFRO conference, Durban, South Africa, 22-26 October 2007 p. in press
- Potts BM, *et al.* (2004) Exploration of the *Eucalyptus globulus* gene pool. *Eucalyptus* in a changing world. Aveiro, Portugal. (Eds NMG Borralho, *et al.*) pp. 46-61. (RAIZ, Instituto Investigação de Floresta e Papel)
- Rengel D, *et al.* (2009) A new genomic resource dedicated to wood formation in *Eucalyptus*. *BMC Plant Biology* 9, (27 March 2009).
- Sansaloni CP, *et al.* (2010) A high-density Diversity Arrays Technology (DArT) microarray for genome-wide genotyping in *Eucalyptus*. *Plant Methods* 6.
- Sexton TR, *et al.* (2010) Genetic association studies in *Eucalyptus pilularis* I Smith (blackbutt). *Australian Forestry* 73, 254-258.
- Shepherd M, *et al.* (2010) Geographical and historical determinants of microsatellite variation in *Eucalyptus pilularis*. *Canadian Journal of Forest Research* 40, 1051-1063.
- Southerton SG, *et al.* (2010) Association of allelic variation in xylem genes with wood properties in *Eucalyptus nitens*. *Australian Forestry* 73, 259-264.
- Stackpole D, *et al.* (2011) Genetic variation in the chemical components of *Eucalyptus globulus* wood. *G3: Genes, Genomes, Genetics* 1, 151-159
- Stackpole DJ, *et al.* (2010) Genetic control of kraft pulp yield in *Eucalyptus globulus*. *Canadian Journal of Forest Research* 40, 917-927.
- Steane DA, *et al.* (2011) Population genetic analysis and phylogeny reconstruction in *Eucalyptus* (Myrtaceae) using high-throughput, genome-wide genotyping. *Molecular Phylogenetics and Evolution* 59, 206-224.
- Teulières C, *et al.* (2007) Stress studies in *Eucalyptus*. *Plant Stress* 1, 197-215.
- Thavamanikumar S, *et al.* (2011) The significance of single nucleotide polymorphisms (SNPs) in *Eucalyptus globulus* breeding programs. *Australian Forestry* 74, 23-29.
- Thumma BR, *et al.* (2009) Identification of a cis-acting regulatory polymorphism in a eucalypt COBRA-like gene affecting cellulose content. *Genetics* 183, 1153-1164.
- Tibbits WN, *et al.* (2006) Genetic variation in frost resistance of *Eucalyptus globulus* ssp *globulus* assessed by artificial freezing in winter. *Australian Journal of Botany* 54, 521-529.
- Volker PW, *et al.* (2008) Genetic parameters of intra- and inter-specific hybrids of *Eucalyptus globulus* and *E. nitens*. *Tree Genetics & Genomes* 4, 445-460.
- Whitham TG, *et al.* (2006) A framework for community and ecosystem genetics: from genes to ecosystems. *Nature Reviews Genetics* 7, 510-523.
- Whitham TG, *et al.* (2008) Perspective - Extending genomics to natural communities and ecosystems. *Science* 320, 492-495.
- Whitham TG, *et al.* (2010) A community and ecosystem genetics approach to conservation biology and management. In 'Molecular insights into natural resource conservation and management'. (Eds A DeWoody, *et al.*) pp. 50-73. (Cambridge University Press)