

Commentary

Plant scientists celebrate new woody plant genome

'I don't think there is a proper way to celebrate something which makes you happy.'

Matthew Oliphant, Usability Works, 1 April 2006

Our friends often remind us that scientists are strange beasts. As such, it is fitting that we should rejoice at the release of a major new genome by further scholarship and analysis – using the genome to unravel new elements of plant physiology and evolution. In this Feature Issue we celebrate the completion of the genome of *Eucalyptus grandis* (Myburg *et al.*, 2014), the first representative of the plant order Myrtales, an early diverging rosid lineage, and a species-rich genus that evolved in isolation on the Australian continent. Compared to plant genomes sequenced to date, it represents an independent evolutionary experiment on what it means to be a large woody perennial plant evolving in diverse, and often stressful, habitats. *Eucalyptus* contains some of the fastest growing hardwood trees and the tallest flowering plant (*Eucalyptus regnans*) on Earth, and also has many species adapted to extremely dry, hot, and nutrient deficient soils. It also produces a diverse array of plant-specific metabolites (including the well-known eucalyptus oils). Members of the genus are among the most productive manufacturers of lignocellulosic biomass on the planet, which makes the genome a rich source of genetic information to inform the development of new kinds of bioproducts, including chemicals, wood, pulp, and bioenergy (Fig. 1).

We have gathered together 13 scientific papers that look more deeply into the genome, and use it to guide further studies important for plant biology and biotechnology. One major theme of the papers is that the many novelties of the eucalypt genome, both structural and functional, appear to bear the fingerprints of natural selection that – as we probe them more deeply – will provide lessons in adaptive evolution for years to come. The same lessons will inform breeding, biotechnology, and genetic engineering.

The extraordinary diversity of topics covered in this small volume illustrates the wide applicability of the genome resource. For example, several papers provided tools or analyses that advance the state of the genome resource or provide new tools for molecular genetic analyses or breeding. Bartholomé *et al.* (2015b, this issue pp. 1283–1296) used genetic recombination data from a hybrid cross and study of 6000 single nucleotide polymorphism (SNP) markers to create a dense genetic map to more precisely organize and collapse genome segments, guiding a more accurate version 2.0

of the genome assembly in the process. Silva-Junior *et al.* (2015, this issue pp. 1527–1540) created a chip with 60 000 SNP markers based on analysis of 240 representative genomes from 12 species. They show that the chip is effective for analyzing genome diversity in a wide variety of eucalypt species, and in a related study (O. B. Silva-Junior & D. Grattaglia, unpublished), they used it to characterize patterns of recombination, polymorphism, and linkage disequilibrium in depth. Their work showed that this new genotyping resource will be very useful for genetic analysis and genome-assisted breeding.

Several papers were motivated by ecological problems. Plett *et al.* (2015, this issue pp. 1423–1436) used genomic resources and genetic diversity to understand how eucalypts and an important root mutualist, the ectomycorrhizal fungus *Pisolithus*, may respond to climate change. The authors reported high variation in root transcriptomes and susceptibility to infection as a result of exposure to different CO₂ concentrations. Different isolates of *Pisolithus* also colonized *E. grandis* in a complex, differential manner.

Building on their high-density genetic maps, Bartholomé *et al.* (2015a, this issue pp. 1437–1449) studied genetic variation in carbon isotope composition that may help inform about adaptation to drought and breeding for stress tolerance. They report 15 quantitative trait loci (QTLs) for isotope composition ($\delta^{13}\text{C}$), widely used as a surrogate for water use efficiency, and find that most are stable across test environments and unrelated to QTLs for growth rate. This suggests that molecular breeding could be used to improve water use efficiency without impacting productivity; getting this tradeoff right is a long-standing problem in *Eucalyptus* plantation forestry.

Several papers probed macroevolution, comparing *E. grandis* genes or genomic patterns to those of other eucalypt species or more distant taxa. This includes a study by Hudson *et al.* (2015, this issue pp. 1378–1390) that investigated genomic patterns of diversity in six eucalypt species based on 2840 DNA markers. Analysis of range-wide collections of the six species revealed that individual loci varied widely in their diversity and extent of population differentiation, though genomic regions often had similar characteristics in different species. The authors reported the identification of a number of species-differentiating markers that are widely distributed in the genome – indicating the likely occurrence of multiple selective sweeps, or lineage-specific mutations.

Kersting *et al.* (2015, this issue pp. 1328–1336) studied the evolutionary changes of protein domains encoded in the *Eucalyptus* genome across key angiosperm evolutionary nodes. By reference to gene ontology and transcriptome databases, they suggested that natural selection played a major role in domain gains, losses, and specialization – and that these were overrepresented in *Eucalyptus*'s many tandemly duplicated genes (versus those from whole genome duplication). Many cases of enrichment were associated with reproduction and stress related genes, and were particularly prominent in genes related to pollen development.

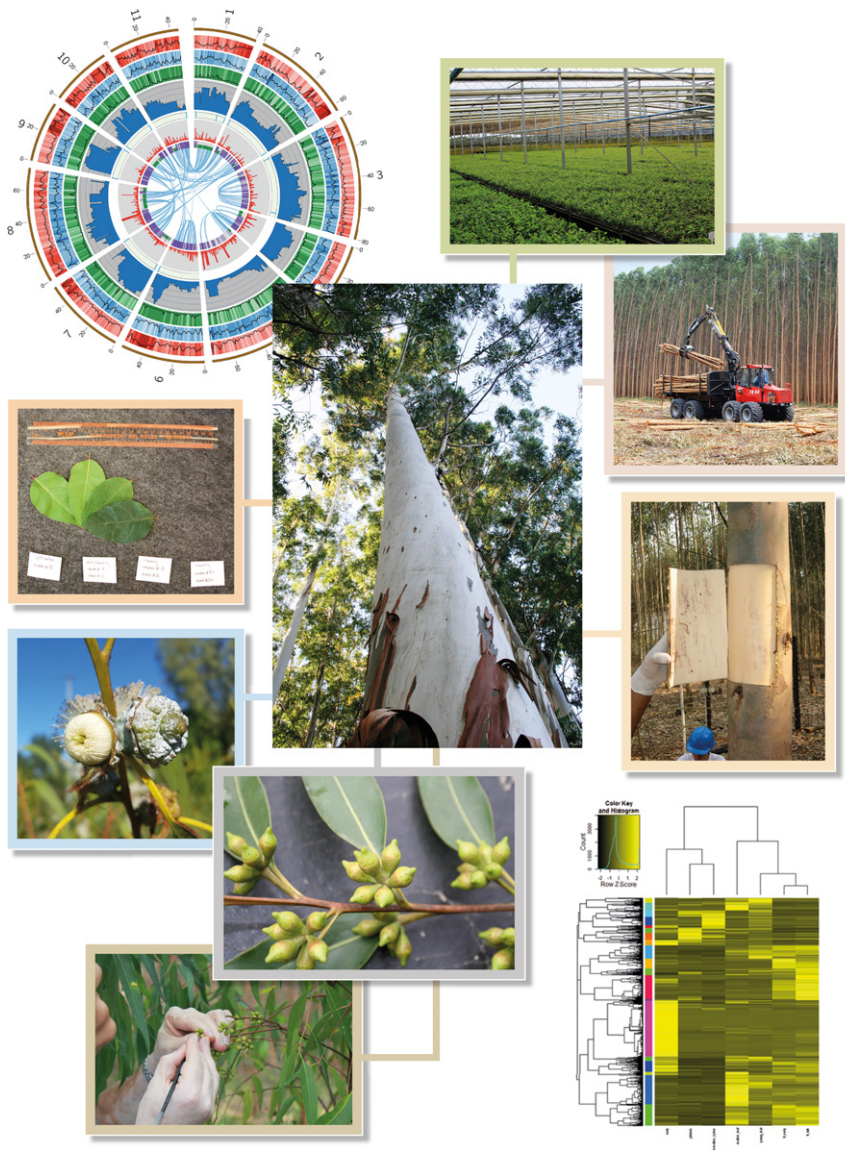


Fig. 1 This eucalypt genomics Feature Issue covers a wide swath of biology and translational tools. The subject matter includes comparative genome mapping and gene expression profiles, studies of genes that control wood properties, growth, and flowering, and tools that will speed applications in breeding, biotechnology, and ecological management (photographs courtesy of Steven H. Strauss and Alexander A. Myburg).

Two papers focused on large, multifunctional transcription factor gene families and their roles in development, wood biology and stress responses. Hussey *et al.* (2015, this issue pp. 1337–1350) studied the NAC protein family, comprising 189 proteins in *E. grandis*. They found that many NAC genes had differential tissue expression in accordance with their roles in development, and that a number of genes exhibited modified expression in response to cold treatment, suggesting a role in coordinating gene expression in response to stress. Similarly, Soler *et al.* (2015, this issue pp. 1364–1377) characterized the 141 members of the R2R3 subfamily of MYB proteins encoded in *E. grandis* with respect to structure, phylogeny, and differential tissue expression. They revealed several groups of MYB genes that are expanded in woody plants and may be involved in woody plant-specific processes such as cambial development.

A number of papers have biotechnology, most often wood biotechnology, as their primary motivation. Carocha *et al.* (2015, this issue pp. 1297–1313) dissected the genomic structure and expression of 11 protein families in the phenylpropanoid pathway for which there is biochemical evidence for a role in lignification.

Quantitative polymerase chain reaction (qPCR), RNA-seq, and phylogenetic analysis with other published genomes was used to characterize a number of gene family members, including several not previously linked to lignification. MacMillan *et al.* (2015, this issue pp. 1314–1327) characterized the 18 member fascilin-like arabinogalactan (FLA) protein family in *E. grandis*. They found that overexpression of some xylem-expressed members can modify microfibril angle and stem mechanical properties. Such changes in cell wall properties are often observed in a specialized xylem tissue called tension wood that is formed on the upper side of leaning or bent hardwood stems. Mizrahi *et al.* (2015, this issue pp. 1351–1363) characterized changes in physicochemical properties and performed transcriptome-wide expression profiling during tension wood formation. They showed that the observed changes in cell wall chemistry in tension wood (relative increase in cellulose content and decrease in lignin) were associated with down-regulation of many lignin biosynthetic genes and differential expression of hemicellulose modifying genes; these observations could have applications in genetic modification of woody biomass properties.

Many wood-specific genes identified in *Eucalyptus* can only be functionally characterized in woody model plants such as *Populus*, an effort that relies on detailed knowledge of orthology and redundancy in the model system. Towards this aim, Hefer *et al.* (2015, this issue pp. 1391–1405) compared the developing xylem and leaf transcriptomes of *Populus trichocarpa* and *E. grandis* in depth. They identified a set of several hundred single gene pairs with strongly conserved xylem expression patterns that likely share functional roles, but also found a set of genes with discordant expression that might reflect the morphological and functional differences between the wood of the two species.

Finally, in a change from the wood-focused norm, Vining *et al.* (2015, this issue pp. 1406–1422) compared the transcriptomes of early and mature stages of floral development in *E. grandis* to that of vegetative tissues, including leaves, roots, and stems. The goal was to identify gene targets and pathways important to regulation of flowering time, floral organ structure, and fertility. They report extensive floral-dominant expression, with nearly 5000 genes differentially expressed between pooled floral and vegetative tissue types. Nearly 500 genes were differentially expressed between the early and nearly mature stages of floral bud development, and 40 floral-dominant genes appeared to be novel or highly diversified from those of other species. The many differentially expressed genes provide new tools for analysis of development of the distinctive eucalypt flowers and fruits, as well as new means for the modification of flowering, including to induce reproductive sterility and to accelerate the onset of flowering – both of which have applications in accelerated domestication and genetic improvement.

The eucalypt genome has, almost instantly, lived up to its promise. In addition to providing a safer-than-average way to party, the papers in this issue show that the genome sequencing investment by the US Department of Energy – and the associated inputs and collaborative efforts from many scientific and industrial communities around the globe – is truly an accomplishment to celebrate.

Steven H. Strauss^{1*} and Alexander A. Myburg^{2*}

¹Department of Forest Ecosystems and Society, Oregon State University, 321 Richardson Hall, Corvallis, OR 97331-5752, USA;

²Department of Genetics, Forestry and Agricultural Biotechnology Institute (FABI), Genomics Research Institute (GRI), University of Pretoria, Private Bag X20, Pretoria 0028, South Africa
(*Authors for correspondence: tel +1 541 737 6578, email steve.strauss@oregonstate.edu (S.H.S.); tel +27 12 420 4945, email zander.myburg@up.ac.za (A.A.M.))

References

- Bartholomé J, Mabila A, Savelli B, Bert D, Brendel O, Plomion C, Gion J-M. 2015a. Genetic architecture of carbon isotope composition and growth in *Eucalyptus* across multiple environments. *New Phytologist* **206**: 1437–1449.
- Bartholomé J, Mandrou E, Mabila A, Jenkins J, Nabihoudine I, Klopp C, Schmutz J, Plomion C, Gion J-M. 2015b. High-resolution genetic maps of *Eucalyptus* improve *Eucalyptus grandis* genome assembly. *New Phytologist* **206**: 1283–1296.
- Carocha V, Soler M, Hefer C, Cassan-Wang H, Ferevereiro P, Myburg AA, Paiva JAP, Grima-Pettenati J. 2015. Genome-wide analysis of the lignin toolbox of *Eucalyptus grandis*. *New Phytologist* **206**: 1297–1313.
- Hefer CA, Mizrahi E, Myburg AA, Douglas CJ, Mansfield SD. 2015. Comparative interrogation of the developing xylem transcriptomes of two wood-forming species: *Populus trichocarpa* and *Eucalyptus grandis*. *New Phytologist* **206**: 1391–1405.
- Hudson CJ, Freeman JS, Myburg AA, Potts BM, Vaillancourt RE. 2015. Genomic patterns of species diversity and divergence in *Eucalyptus*. *New Phytologist* **206**: 1378–1390.
- Hussey SG, Saïdi MN, Hefer CA, Myburg AA, Grima-Pettenati J. 2015. Structural, evolutionary and functional analysis of the NAC domain protein family in *Eucalyptus*. *New Phytologist* **206**: 1337–1350.
- Kersting AR, Mizrahi E, Bornberg-Bauer E, Myburg AA. 2015. Protein domain evolution is associated with reproductive diversification and adaptive radiation in the genus *Eucalyptus*. *New Phytologist* **206**: 1328–1336.
- MacMillan CP, Taylor L, Bi Y, Southerton SG, Evans R, Spokevicius A. 2015. The fascilin-like arabinogalactan protein family of *Eucalyptus grandis* contains members that impact wood biology and biomechanics. *New Phytologist* **206**: 1314–1327.
- Mizrahi E, Maloney VJ, Silberbauer J, Hefer CA, Berger DK, Mansfield SD, Myburg AA. 2015. Investigating the molecular underpinnings underlying morphology and changes in carbon partitioning during tension wood formation in *Eucalyptus*. *New Phytologist* **206**: 1351–1363.
- Myburg AA, Grattapaglia D, Tuskan GA, Hellsten U, Hayes RD, Grimwood J, Jenkins J, Lindquist E, Tice H, Bauer D *et al.* 2014. The genome of *Eucalyptus grandis*. *Nature* **510**: 356–362.
- Plett JM, Kohler A, Khachane A, Keniry K, Plett KL, Martin F, Anderson IC. 2015. The effect of elevated carbon dioxide on the interaction between *Eucalyptus grandis* and diverse isolates of *Pisolithus* sp. is associated with a complex shift in the root transcriptome. *New Phytologist* **206**: 1423–1436.
- Silva-Junior OB, Faria DA, Grattapaglia D. 2015. A flexible multi-species genome-wide 60K SNP chip developed from pooled resequencing of 240 *Eucalyptus* tree genomes across 12 species. *New Phytologist* **206**: 1527–1540.
- Soler M, Camargo ELO, Carocha V, Cassan-Wang H, San Clemente H, Savelli B, Hefer CA, Paiva JAP, Myburg AA, Grima-Pettenati J. 2015. The *Eucalyptus grandis* R2R3-MYB transcription factor family: evidence for woody growth-related evolution and function. *New Phytologist* **206**: 1364–1377.
- Vining KJ, Romanel E, Jones RC, Klocko A, Alves-Ferreira M, Hefer CA, Amarasinghe V, Dharmawardhana P, Naithani S, Ranik M *et al.* 2015. The floral transcriptome of *Eucalyptus grandis*. *New Phytologist* **206**: 1406–1422.

Key words: adaptation, biotechnology, *Eucalyptus*, evolution, flowering, forests, lignin, wood.