

The floral transcriptome of *Eucalyptus grandis*

Kelly J. Vining¹, Elisson Romanel², Rebecca C. Jones³, Amy Klocko¹, Marcio Alves-Ferreira⁴, Charles A. Hefer⁵, Vindhya Amarasinghe^{1,6}, Palitha Dharmawardhana⁶, Sushma Naithani⁶, Martin Ranik⁷, James Wesley-Smith⁸, Luke Solomon⁹, Pankaj Jaiswal⁶, Alexander A. Myburg⁶ and Steven H. Strauss¹⁰

¹Center for Genome Research and Biocomputing, Oregon State University, Corvallis, OR 97331, USA;

²Departamento de Biotecnologia, Escola de Engenharia de Lorena, Universidade de São Paulo (EEL-USP), CP 116, 12602-810 São Paulo, Brazil;

³School of Biological Sciences, University of Tasmania, Private Bag 55, Hobart, 7001 TAS, Australia;

⁴Laboratório de Genética

Molecular Vegetal (LGMV), Departamento de Genética, Universidade Federal do Rio de Janeiro (UFRJ), Av. Prof. Rodolpho Paulo Rocco, CCS 21949900 Rio de Janeiro, Brazil;

⁵Department of Botany, University of British Columbia, 3529-6270 University Blvd, Vancouver, BC V6T 1Z4, Canada; ⁶Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR 97331, USA;

⁷Department of Genetics, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Private Bag X20, Pretoria 0028, South Africa;

⁸Council for Scientific

and Industrial Research, 1 Meiring Naude Rd, Pretoria, South Africa;

⁹Seed Technology Programme, Sappi Forests Shaw Research Center, Howick 3290, South Africa;

¹⁰Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR 97331, USA

Summary

As a step toward functional annotation of genes required for floral initiation and development within the *Eucalyptus* genome, we used short read sequencing to analyze transcriptomes of floral buds from early and late developmental stages, and compared these with transcriptomes of diverse vegetative tissues, including leaves, roots, and stems. A subset of 4807 genes (13% of protein-coding genes) were differentially expressed between floral buds of either stage and vegetative tissues. A similar proportion of genes were differentially expressed among all tissues. A total of 479 genes were differentially expressed between early and late stages of floral development. Gene function enrichment identified 158 gene ontology classes that were overrepresented in floral tissues, including ‘pollen development’ and ‘aromatic compound biosynthetic process’. At least 40 floral-dominant genes lacked functional annotations and thus may be novel floral transcripts. We analyzed several genes and gene families in depth, including 49 putative biomarkers of floral development, the MADS-box transcription factors, ‘S-domain’-receptor-like kinases, and selected gene family members with phosphatidylethanolamine-binding protein domains. Expanded MADS-box gene subfamilies in *Eucalyptus grandis* included SUPPRESSOR OF OVEREXPRESSION OF CO 1 (SOC1), SEPALLATA (SEP) and SHORT VEGETATIVE PHASE (SVP) *Arabidopsis thaliana* homologs. These data provide a rich resource for functional and evolutionary analysis of genes controlling eucalypt floral development, and new tools for breeding and biotechnology.