RNA-Seq Analysis of Eucalyptus Genotypes that Differ in Carbon Allocation

Isabel R. Gerhardt, Poliana F. Giachetto, Michel E. B. Yamagishi, Francisco P. Lobo, Ricardo M. Penchel, Alexandre A. Missiaggia, Adhemar Zerlotini, Felipe Rodrigues da Silva
Embrapa Forestry, Embrapa Agriculture Informatics, Embrapa Agriculture Informatics, Fibria Celulose S.A., Fibria Celulose S.A., Embrapa Agriculture Informatics, Embrapa Agriculture Informatics

The global demand for wood combined with its diversified use requires the generation of trees that sequester and accumulate more carbon, providing differentiated raw material for the production of pulp, paper, charcoal and even cellulosic ethanol. Carbon allocation is the process of translocation of photosynthate from source to sink organs, and though its physiology is well known, the genetic mechanisms involved in its regulation are still poorly understood. Since differences in levels of gene expression may largely explain the observed phenotypic variation, and there is great variability among species of Eucalyptus, we decided to perform a gene expression analysis of four contrasting Eucalyptus genotypes to gain insight into the mechanisms that lead to differences in carbon allocation. Leaf, xylem and root samples from each genotype were used for transcriptome sequencing using Illumina Hi-Seq technology. After control quality analysis, the reads were mapped to the Eucalyptus grandis reference genome using TopHat, and the gene expression analysis was performed using CuffDiff. Pairwise comparisons were carried out between tissues and genotypes, and statistical tests were performed to assess differential expression. The analysis was executed within Galaxy, which also permitted us to visualize the reads mapping. CummeRbund was used to generate result tables and charts. Gene Ontology terms were assigned to the genes using InterProScan, and Bingo was used to identify the enriched terms. We generated a total of 89.3Gb reads and between 70.75% and 90.33% of them were aligned to the reference genome. We then computed the FPKM values, thus allowing us to identify 26,190 differentially expressed genes out of 44,974 predicted genes. We expect to unveil candidate genes for carbon allocation that will be further investigated by transgenic over- and down-regulated expression approaches.

Keywords: RNA-Seq, Eucalyptus
Concentration area: Transcrip Proteomic