Comparative transcriptome analysis of latex from rubber tree clone CATAS8-79 and PR107

Jinquan Chao, Yueyi Chen, Shaohua Wu, Wei-Min Tian *
Ministry of Agriculture Key Laboratory of Biology and Genetic Resources of Rubber Tree/State Key Laboratory Breeding Base of Cultivation and Physiology for Tropical Crops, Rubber Research Institute, Chinese Academy of Tropical Agricultural Sciences, Danzhou, Hainan 571737, PR China

Abstract
Rubber tree (Hevea brasiliensis Muell. Arg.) is the primary commercial source of natural rubber in the world. Latex regeneration and duration of latex flow after tapping are the two factors that determine the rubber yield of a rubber tree, and exhibit a huge variation between rubber tree clones CATAS8-79 and PR107. In the present paper, with the purpose of globally characterizing latex transcriptome, RNAs were extracted from CATAS8-79 and PR107 at first tapping and sequenced with Illumina paired-end sequencing technology individually. After excluding low-quality reads as empty adapters, 26 million clean reads were generated in both pools. Using SOAPdenovo software, 296,736 and 308,262 contigs ranging from 100 bp to more than 3000 bp were assembled in CATAS8-79 and PR107 (NCBI accession numbers: GSE59981). Based on paired-end and gap-filling, 53,571 and 57,806 unigenes were generated in CATAS8-79 and PR107 individually. With the help of unigenes from two pools, it is possible to identify the longer sequence derived from the same transcript as reference transcriptome. Thus, 51,829 all-unigenes were finally integrated using paired-end joining with an average length of 640 bp and a N50 of 526 bp.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

2. Experimental design, materials and methods

2.1. Plant materials
Seven-year-old virgin trees of rubber tree clone CATAS8-79 and PR107 were grown at the Experimental Station of the Rubber Research Institute of the Chinese Academy of Tropical Agricultural Sciences in Danzhou city, Hainan province. Virgin trees with the same circumference were selected for this study. For RNA-Seq, latex from five individual trees by the first tapping was pooled for each clone. The samples were immediately stored at −80 °C until RNA extraction. For real-time-PCR and determination of physiological parameters, latex was individually collected from another batch of five trees for each clone upon the first, second, third and forth tapping, respectively. All the selected virgin trees were tapped with a tapping system of S/2, d/2 (a half spiral pattern, every two days) at 6:00 am in August, 2013.

2.2. RNA isolation and sequencing
Total latex RNA was extracted as described [1] and RNA integrity was evaluated by NanoDrop (Thermo Scientific Inc., USA). The double strand cDNA was synthesized using SuperScript® Double-Stranded cDNA Synthesis Kit (Invitrogen Inc., USA), and purified and added single
nucleotide A (adenine) to the end with QiaQuick PCR extraction kit. Finally, sequencing adaptors were ligated to the cDNA fragments. The required fragments were purified by 2% agarose gel electrophoresis and enriched by PCR amplification. The library products were sequenced via Illumina HiSeq™ 2000 by Beijing Genomics Institute (Shenzhen, China). The original image datasets was transferred into sequence datasets by base calling. Clean reads were obtained by removing adaptor sequence, low quality sequences, empty tags, low complexity, and tags with only one copy. Finally, 26,266,670 and 26,266,670 clean reads were generated in CATAS8-79 and PR107 pool, respectively.

2.3. Transcriptome de novo assembly, annotation and classification

Transcriptome de novo assembly was carried out using a de Bruijn graph and the SOAPdenovo as previously described [2]. Under a certain overlap length (k-mer = 29), SOAPdenovo combined overlapping reads into contigs. Adjacent contigs were constructed into scaffolds by read mate pairs. Within the scaffold, the connected contigs used 'N' to represent unknown sequences and insert size information. Finally, paired-end information was used to fill the gap of scaffolds to obtain the extended sequences with fewer Ns, which were defined as unigenes for further analysis. The data for contig and unigene were listed in Table 1.

All unigenes were used for BLAST searches (E-value \( \leq 1E^{-5} \)) against databases as NCBI Nr (http://www.ncbi.nlm.nih.gov/), Swissprot (http://www.expasy.org/sprot/), KEGG (http://www.genome.jp/kegg/) and COG (http://www.ncbi.nlm.nih.gov/cog/). The best aligning results were chosen for unigene annotation. The aligning results were selected with an order of Nr, Swiss-Prot, KEGG and COG. To classify the unigenes, the Blast2GO program was used to get GO annotation based on molecular function, biological process and cellular component. All unigenes were also aligned to the COG database to predict possible functions and KEGG pathway database to perform pathway assignments.

2.4. Digital gene expression analysis

A rigorous algorithm was developed to identify differentially expressed genes between two different DGE libraries (CATAS8-79 versus PR107). Raw clean tags in each library were normalized to Tags Per Million (TPM) to obtain normalized gene expression level. Differential digital gene expression was deemed with FDR value \( \leq 0.001 \) and \( |\log_2 \text{Ratio}| \geq 1 \) in sequence counts across libraries. "Up-regulated" means the level of gene transcripts were higher in PR107 whereas "down-regulated" means the level of gene transcripts were higher in CATAS8-79. Based on the limit role listed above, a total of 6726 unigenes with differential expression patterns were detected between CATAS8-79 and PR107.

Conflict of interest

The authors declare that they have no competing interests.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (31170642) and the Special Program for Key Basic Research of the Ministry of Science and Technology, China (2012CB723005).

References

[1] C. Tang, D. Huang, J. Yang, S. Liu, S. Sakr, H. Li, Y. Zhou, Y. Qin, The sucrose transporter HbSUT3 plays an active role in sucrose loading to laticifer and rubber productivity in exploited trees of Hevea brasiliensis (para rubber tree). Plant Cell Environ. 33 (2010) 1708–1720.

[2] C.C. Mantello, C.B. Cardoso-Silva, C.C. da Silva, L.M. de Souza, E.J. Scaloppi Junior, G.P. de Souza, R. Vicentini, A.P. de Souza, De novo assembly and transcriptome analysis of the rubber tree (Hevea brasiliensis) and SNP markers development for rubber biosynthesis pathways. PLoS ONE 9 (2014) e102665.