Transcriptome analysis and SSR molecular marker exploration in early flower development of *Camellia azalea*

Xiao-fei Liu, Ying-bo Sun, Li-li Huang, Chao-yi Zhao, and Bo Yu*

Environmental Horticulture Institute, Guangdong Academy of Agricultural Sciences, Guangdong Key Lab of Ornamental Plant Germplasm Innovation and Utilization, Key Laboratory of Urban Agriculture in South China, Ministry of Agriculture, Guangzhou, 510640, China.

*Corresponding author. Email: yubo423@163.com

Background

Camellia azalea (Theaceae) is an evergreen shrub or small tree naturally restricted to the Erhuangzhang Nature Reserve in Yangchun, Guangdong Province, China.

Methods

In this study, inflorescences with diameters of 0.5 cm (S1), 1.0 cm (S2), and 1.5 cm (S3) were selected, respectively. And their differences were compared at the transcription level, meanwhile, MISA (default parameters) was used to perform SSR detection on unigenes.

Results

Correlation analysis between 3 samples (Fig.1) showed that there was a greater correlation between S2 and S3. On the contrary, the correlation between S1 and the other two samples was relatively low.

Analysis of differences expression genes (DEGs) between samples (Fig.2) suggested that there was 7,753 DEGs between S1 and S3, and the expression of up-regulated genes was 5,153, while the down-regulated ones was 2,600. The DEGs between S1 and S2 was 4,880, and with 2,267 DEGs was up-regulated genes, while the down-regulated ones was 2,613. There was 3,213 DEGs between S2 and S3, and the expression of up-regulated genes was 817, while the down-regulated ones was 2,396.

By predicting transcription factors in differentially expressed genes, a total of 1,005 transcription factors were obtained, which can be attributed to 23 families. The top two are MYB_superfamily and bHLH families, with 162 and 94 predicted respectively. Combined with enrichment analysis results of DEGs (Fig.3) suggested that plant hormone signal transduction, especially zeatin biosynthesis-related TFs may play an important role in the early flower development in *C. azalea*.

Total number of identified SSRs was 35,626. The two SSR types, Di-nucleotide repeats and Mono-nucleotide repeats, have the largest number (Fig.4).

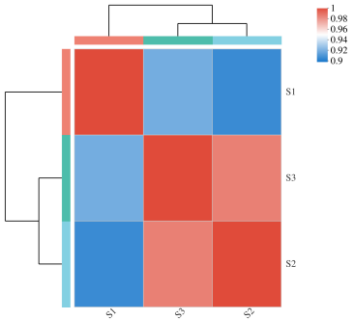

Fig. 1 Correlation analysis between samples.

Fig. 2 Statistics of differences expression genes between samples

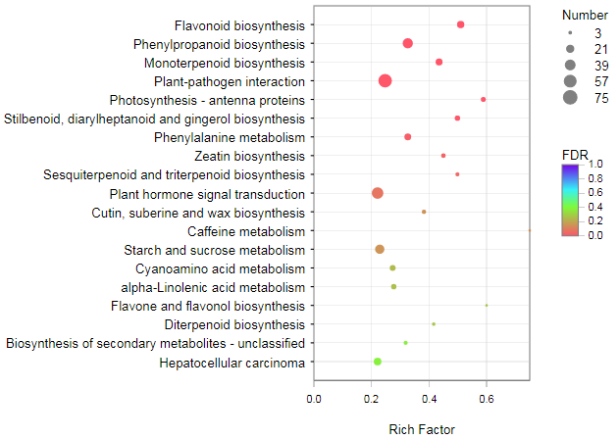

Fig. 3 KEGG extremely significant enrichment analysis (S1_vs_S3_G).

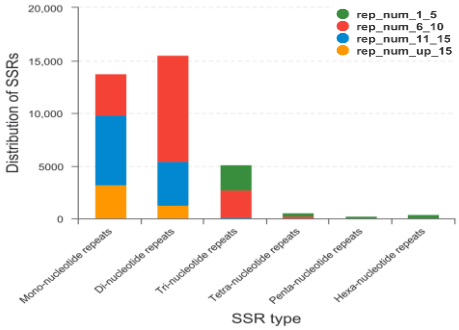

Fig. 4 Statistical results of SSR type

Conclusion

The RNA-seq technique was used to compare the differences in the transcription level of the early flower development of *Camellia azalea* at different stages. Subsequently, on this basis, SSR molecular marker exploration was performed by MISA (default parameters). The results of statistics of differences expression genes between samples suggested that there was a big difference between S1 and S3, and the expression of up-regulated genes was significantly more than down-regulated ones. Analysis on TF families and enrichment of DEGs suggested that plant hormone signal transduction, especially zeatin biosynthesis-related TFs may play an important role in the early flower development in *C. azalea*. Total number of identified SSRs was 35,626. The two SSR types, Di-nucleotide repeats and Mono-nucleotide repeats, have the largest number.

Funding

This study was supported by the Key-Area Research and Development Program of Guangdong Province (2018B020202002, 2020B020220005).