PB-13

Gene Prediction and Functional Annotation in *Perilla citriodora*

Hong-Il Ahn[1], Seon-Hwa Bae[1], Ye-Ji Lee[1], Myoung-Hee Lee[2], Jeong-Hee Lee[3], Keunpyo Lee[1], Tae-Ho Kim[1]*

[1]Genomics Division, National Institute of Agricultural Sciences, RDA, Jeonju 54874, Korea
[2]Department of Southern Area Crop Science, National Institute of Crop Science, Miryang 50424, Korea
[3]SEEDERS, Daeduk Industry—Academic Cooperation Building, Daejeon 30415, Korea

[Introduction]

Perilla has been cultivated as annual herbaceous oil crop of the Laminaceae family in East Asia, and is well known for its high α-linolenic acid accumulation in seeds. *Perilla* composes high ALA almost 60% of total fatty acid in seed (common edible oils contained ALA less than 3%) and is endowed with various benefits to human health. Recently, as the demand for functional natural nutrients increased, the structural and functional characteristics of *Perilla* genome could strengthen the use of this crop as a beneficial food supplement. In this study, to accelerate the improvement of breeding efforts for *Perilla*, we present here the diploid *Perilla* genome analysis.

[Materials and Methods]

Gene prediction analysis was performed using major transcript sequence generated by merging all transcript assembly results, de-novo repeat sequence of perilla produced by repeatModeler, plant RefSeq protein sequence of NCBI, and GyDB. In addition, coding region was found using TransDecoder for transcript base sequences obtained from TACO and Seqping, and BLAST and hmmsearch were performed using NCBI RefSeq plant protein and Pfam DB. The BLAST and HMMER were performed to calculate the statistical significance and to deduce the function by comparing with the gene database which was already known about the selected genes.

[Results and Discussion]

Of the 32,867 genes in *Perilla*, 32,279 proteins were found in NCBI, and 24,445 functional proteins were detected in Pfam. In addition, the BLAST2GO analysis showed the highest ratio in a cellular process, binding, and cell sections. Using the BUSCO and plant lineage dataset, the *Perilla* pseudomolecule sequence and gene set were evaluated to be as high as 91.7%.

[Acknowledgement]

*Corresponding author: Tel. +82—63—238—4563, E-mail. thkim@korea.kr