

Genome-wide classification and identification of transmembrane protein sequences in plant genomes

Yasuhito Inoue*

Department of Electrical and Computer Engineering, Maizuru National College of Technology, 234 Shiroya, Maizuru, Kyoto, 625-8511, Japan

Transmembrane proteins in plant genomes are largely responsible for signaling cascades and pathogen defense reactions, enzymes such as the apparatus for cell wall biosynthesis, and transporters responsible for the import and export of metals and alkali and the establishment of electrochemical gradients across membranes. In this study, we predicted the transmembrane protein sequences to classify/identify functional annotations based on sequence similarity searches or transmembrane topology patterns.

First, after we downloaded the open reading frame (ORF) sequences of 16 plant genomes (*A. lyrata*, 32,670; *A. thaliana*, 28,952; *B. distachyon*, 32,255; *Coccomyxa sp. C-169*, 9,629; *C. papaya*, 25,508; *C. reinhardtii*, 16,698; *C. sativus*, 32,527; *G. max*, 46,430; *M. esculenta*, 43,844; *M. guttatus*, 27,501; *O. sativa*, 36,120; *P. patens*, 36,937; *P. trichocarpa*, 44,482; *S. bicolor*, 36,336; *S. moellendorffii*, 34,697; *V. vinifera*, 26,346), we were segregated into transmembrane/soluble proteins by SOSUI (Hirokawa et al., 1998). Then, the transmembrane topology prediction was carried out by TMHMM 2.0 (Tusnady and Simon, 1998 and Tusnady and Simon, 2001) after removing any signal peptide, chloroplast transit peptide, and mitochondrial transit peptide regions detected by SignalP 3.0, ChloroP 1.0, and TargetP 1.0. transmembrane protein sequences predicted to be 7-transmembranes (*A. lyrata*, 9,852; *A. thaliana*, 9,117; *B. distachyon*, 10,064; *Coccomyxa sp. C-169*, 2,365; *C. papaya*, 6,377; *C. reinhardtii*, 3,777; *C. sativus*, 9,472; *G. max*, 15,685; *M. esculenta*, 11,074; *M. guttatus*, 8,956; *O. sativa*, 11,020; *P. patens*, 7,354; *P. trichocarpa*, 11,719; *S. bicolor*, 10,196; *S. moellendorffii*, 9,058; *V. vinifera*, 8,442) were submitted to the classification/identification procedure by the statistical analysis. We will provide for anonymous ftp download as a plain text file from <http://bioinfo.maizuru-ct.ac.jp/membrane/>.

*Corresponding author (yinoue@maizuru-ct.ac.jp)