

Identification of putative arabinogalactan protein genes from *Physcomitrella patens*

Yoichi Sakata, Ralph S. Quatrano

Department of Biology, Washington University, St. Louis, MO 63130-4899, USA

Contact email: naokisat@molbiol.saitama-u.ac.jp

Arabinogalactan protein (AGP) is a widely distributed proteoglycan in the plant kingdom. Although AGP is suggested to be involved in differentiation, development and cell-cell communication in plants, its function remains unclear. A reverse genetics approach will give us clues to understand the molecular function of AGP. *Physcomitrella patens* is ideal for this purpose because gene targeting is available with high efficiency in this plant.

In *P. patens* little is known about AGP, and no sequence information has been reported. We performed a BLAST search of the moss EST database by BLAST program using *Arabidopsis* AGP (AtAGP1-17 and AtFLA1-14) genes as queries. So far, we have retrieved 6 cDNAs encoding putative AGP proteins. From the analyses of deduced amino acid sequences, all proteins encoded by these cDNAs appear to be classical AGPs. Classical AGPs are defined by; 1) an N-terminal secretion sequence that is removed from the mature protein, 2) the mature protein contains Pro/Hyp (hydroxyproline), Ala, Ser, Thr and Gly as the major amino acids and 3) the presence of GPI-anchor signal in the C-terminal portion of the ORF.

On a structural basis, *P. patens* AGPs are classified into three subgroups, which are also found in *Arabidopsis*. There is one AGP with a mature protein of 117 amino acids, and three AGP cDNAs encode peptides with short mature proteins of 11, 16 and 17 amino acids, respectively. Another group consists of two AGPs that contain one or two fasciclin domains. Fasciclin domains were first found in the *Drosophila* cell adhesion molecule (CAM) fasciclin I, and it was also reported that algal CAM which contains this domain plays an important role in adhesion of neighboring cells during embryogenesis in *Volvox*.

Southern blot analysis of *P. patens* genomic DNA showed each cDNA exist as a single copy gene. We have started to make transgenic moss lines for over-expression and knockout of these AGPs. Some preliminary results will be presented.