Identification of Genes Expressed in Apical Cells of the Moss *Physcomitrella patens* Using Gene-trap and Enhancer-trap Systems

Yuji Hiwatashi

PhD Dissertation

2001

Department of Molecular Biomechanics
The Graduate University for Advanced Studies

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Postembryonic growth in land plants occurs from the meristem, a localized region that gives rise to all adult structures, such as a stem and leaves. Meristems control the continuous development of plant organs by balancing the maintenance and proliferation of undifferentiated stem cells, and directing their differentiation. Meristem establishment and maintenance is a fundamental question in plant development research.

Mosses have two types of meristems: a protonema apical cell and a gametophore apical cell. The gametophore apical cell is a single meristematic cell that is maintained through self-renewal, and gives rise to such organs as the stem and leaves. In the moss *Physcomitrella patens*, the developmental process of the apical cell is well defined at the cellular level, and gene targeting based on homologous recombination is feasible. Thus, apical cell differentiation in *P. patens* is used as a model system for studies of meristem development in land plants.

This study investigated apical cell differentiation of gametophores in *P. patens* by identifying the genes expressed during this differentiation. First, gene-trap and enhancer-trap systems were established in *P. patens*. These techniques are useful for cloning genes and enhancers that function in specific tissues or cells. In addition, the systems are convenient for obtaining molecular markers specific to certain developmental processes. Elements for the two systems were constructed using a *uidA* reporter gene with a splice acceptor in the case of the gene-trap system, and a minimal promoter for the enhancer-trap system. The homologous recombination method allowed a high rate of transformation, finding 235 gene-trap and 1073 enhancer-trap lines with variable expression patterns from a total of 5637 gene-trap and 3726 enhancer-trap transgenic lines.

To assess the feasibility of isolating a trapped gene, one gene-trap line, YH209 with rhizoid-specific GUS expression, was characterized. *UidA*-fused fragments were amplified by the 5' RACE method using *uidA*-specific primers. One of the amplified fragments was used to screen the mini-transposon-tagged genomic DNA library that was used to generate the *P. patens* gene-trap lines. A genomic fragment containing the sequence of 5' RACE fragments was obtained. This fragment was re-integrated into the *P. patens* genome by homologous recombination, confirming that the fragment-integrated transformants exhibited rhizoid-specific expression patterns observed in the YH209 line. In addition, a full-length cDNA was isolated by the 3' RACE method, and the gene was named *PpGLU*. *PpGLU* forms a clade with the acidic alpha-glucosidase genes of plants. The gene-trap and enhancer-trap systems should be useful for identifying cell-type and tissue-specific genes in *P. patens*.

From the 235 gene-trap lines and the 1073 enhancer-trap lines, three and four lines, respectively, were isolated. The isolated lines were exhibiting GUS activity preferentially in the apical cells of buds. One gene-trap line, *Apicar1*, showing GUS activity predominantly in the apical cell of caulonemata, rhizoids, and gametophores, was further characterized. The candidate trapped gene was isolated by both the 5' and 3' RACE methods, using the same approach as used with the YH209 line. A sequence analysis of the isolated cDNA revealed that the trapped gene encoded a kinesin-like protein, *API1* (*Apicar1*). According to a phylogenetic analysis of *API1* and kinesin superfamily genes, the *API1* gene formed a new family of kinesin-like proteins with one of the *Arabidopsis* kinesin-like genes. This suggests that *API1* may have a novel function that is different from those of kinesins of other subfamilies. *API1* will also be useful as a molecular marker in studies of the establishment and maintenance of the apical cell.

GENERAL INTRODUCTION

In land plants, the adult plant body is formed postembryonically by the continuous activity of small cell clusters known as the meristem. The meristem can be active over extended periods of plant development. As a result of this continued meristematic activity, the body plan of plants is indeterminate, whereas the body plan of metazoans is largely determinate (Sussex and Kerk, 2001). Plants retain a pool of undifferentiated cells by balancing the proliferation of meristematic cells with their subsequent incorporation into tissue or organ primordia. Undifferentiated stem cells in the meristem replenish the cells used for the initiation of primordia, and thus, are the ultimate source of almost the entire plant body (reviewed in Waites and Simon, 2000). One of the fundamental questions in plant development research is how is the meristem established and maintained?

In angiosperms, the meristems originate at the two distal ends, the shoot and root meristems, of the plant body during embryogenesis. The shoot meristem is located in the apex of a shoot and is responsible for generating the aboveground portion of the plant. The shoot meristem can be divided into two highly organized zones, each with a specialized cell type. The central zone contains meristematic cells in an undifferentiated state, and the peripheral zone contains cells that enter a specific developmental pathway toward a differentiated state (reviewed in Fletcher and Meyerowitz, 2000; Steeves and Sussex, 1989). Using *Arabidopsis thaliana* as a model angiosperm, several mutants with altered meristem development have been identified, and several genes have been isolated. This has added considerably to our understanding of such processes as the establishment and maintenance of the shoot meristem (reviewed in Clark, 2001; Fletcher and Meyerowiz, 2000; Lenhard and Laux, 1999).

Mosses, members of the Bryophyta, are a successful group of small non-vascular plants, which diverged from vascular plants more than 400 million years ago (Stewart and Rothwell, 1993). They are widely distributed around the world and can endure extreme environmental conditions (Birks *et al.*, 1998). Their suitability as model systems to study plant cellular and developmental processes has been recognized because they exhibit simple morphogenesis with a small number of cell types, enabling observation at the level of individual cells (reviewed in Cove *et al.*, 1997). Mosses have two types of meristems: a protonemal apical cell and a gametophore apical cell. In contrast to the highly organized shoot meristem of angiosperms, the gametophore apical cell is a single meristematic cell that generates primordial cells. These primordial cells develop into organs such as the stem and leaves, which are simpler than those of angiosperms. The gametophore apical cell is maintained through self-renewal. The process of gametophore apical cell development is well defined at the cellular level (reviewed in Schumaker and Dietrich, 1998), and is similar to the shoot meristem development of angiosperms (reviewed in Reski, 1998).

In animals, model organisms are useful for understanding developmental mechanisms because of broad similarities in development (Wolpert *et al.*, 1998). Furthermore, comparisons of developmental mechanisms among model organisms are indispensable for understanding animal diversity and evolution. Even comparisons of developmental mechanisms among phylogenetically diverse metazoans, for example, the mouse, frog, nematode, and fruit fly, have yielded significant insights on development and evolution (Wolpert *et al.*, 1998). The gametophore apical cell of mosses provides a unique opportunity to study the establishment and maintenance of the meristem in land plants through the use of a relatively simple organism. The ultimate aim of this study is to demonstrate the general principles of meristem development in land plants by clarifying the process of apical cell differentiation in mosses, and

comparing the mechanisms in mosses with those in angiosperms.

Among the mosses, *Physcomitrella patens* is particularly useful for the purpose of comparing developmental mechanisms with angiosperms because of four characteristics (reviewed in Cove *et al.*, 1997; Schumaker and Dietrich, 1998; Reski, 1998), which are described below.

(1) *P. patens* is easily cultured *in vitro* and has a simple life cycle.

P. patens can be propagated easily under axenic conditions. Its anatomical structures are simple, with relatively few cell types, and organs composed of either a single layer or only a few layers of cells.

The life cycle of *P. patens* can be completed in approximately 3 months *in vitro*. The earliest stage of vegetative development in *P. patens* is characterized by cellular differentiation during filament growth. The filamentous cells are called protonemata. Spore germination leads to the formation of a filament that consists of the apical cell and a linear array of subapical cells, which are produced by successive divisions of an apical cell (Figures 1A and 1B). These filamentous cells are known as the chloronema, and have disc-shaped chloroplasts and cross walls perpendicular to the filament axis.

The chloronema filament continues to grow until the appearance of the chloronemal apical cell begins to change. This cellular differentiation leads to the formation of the second filament cell type, the caulonema (Figure 1C). In comparison to a chloronemal apical cell, a caulonemal apical cell has smaller, spindle-like chloroplasts and cross walls that are oblique to the filament axis. A small swelling, termed a side branch initial cell, normally appears in the second subapical cell (the third cell of the filament) (Figure 1D). This cell is positioned at the end of the oblique cross wall, closest to the apical cell, and can follow one of four different developmental

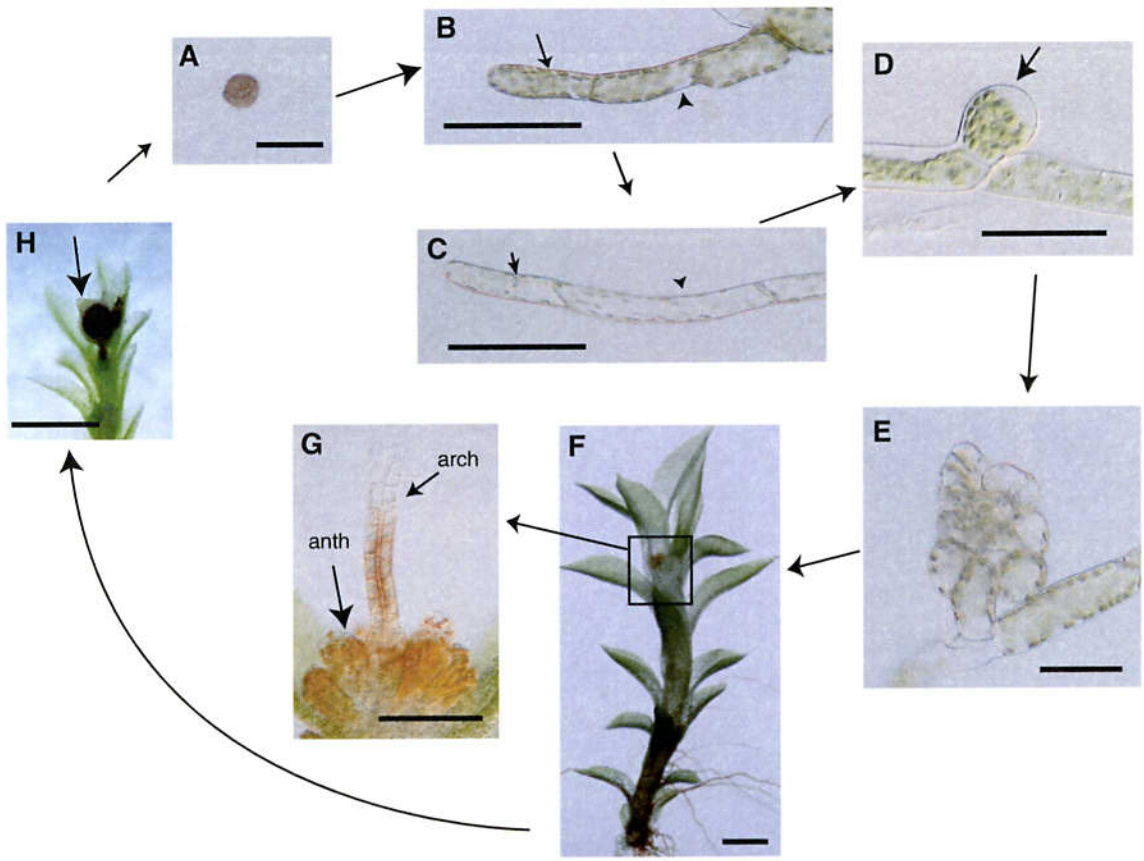


Figure 1. The life cycle of *Physcomitrella patens*.
(A) Spore. (B) Chloronemal cells. An arrow and an arrowhead indicate a chloronemal apical cell and a chloronemal subapical cell, respectively. (C) Caulonemal cells. An arrow and an arrowhead indicate a caulonemal apical cell and a caulonemal subapical cell, respectively. (D) A caulonemal side branch initial cell (an arrow). (E) A bud. (F) A gametophore. (G) An archegonia (arch) and antheridia (anth). (H) A sporophyte (an arrow). The bar in (A) = 30 μm, in (B), (C), and (G) = 100 μm, in (D) and (E) in 50 μm, in (F) and (H) = 500 μm.

pathways: quiescence, development of a secondary chloronemal cell, development of a new caulonemal cell, or development of a leafy shoot primordia, the bud (McClelland, 1987). The developmental step from side branch initial cell to bud corresponds to the transition from one-dimensional apical growth to three-dimensional growth (Figure 1E). The bud has a tetrahedral apical cell that produces a leafy shoot and develops into the gametophore. The gametophore is composed of a non-vascular stem and leaves that are formed by a single layer or a few layers of cells (Figure 1F). With cold treatment of the gametophore, cell division of the apical cell eventually gives rise to the gametangia, upon which are the archegonia and antheridia (Figure 1G). Following the production of gametes and fertilization, the zygote develops into a sporophyte (Figure 1H) with a single capsule containing thousands of haploid spores. The simple structure of the gametophores and protonemata allows us to investigate physiological and developmental processes at the cellular level.

(2) Genetic analysis is facilitated in *P. patens*.

The haploid gametophyte dominates the life cycle of *P. patens*, making mutagenesis and genetic analysis easy. The size of the haploid genome of *P. patens* has been estimated at around 480 Mb (Reski, 1998). Methods have been established for genetic analysis by sexual crossing (Ashton and Cove, 1977), and somatic hybridization of protoplasts for complementation analysis (Grimsley *et al.*, 1977; Watts *et al.*, 1985).

Chemical mutagenesis of *P. patens* has been achieved using spore tissue (Ashton and Cove, 1977), and somatic tissue (Boyd *et al.*, 1988). A large collection of *P. patens* mutants has been isolated (Knight *et al.*, 1988), including the following classes: (i) auxotrophs, (ii) analogue-resistant mutants, (iii) mutants with altered phototropic or gravitropic responses, (iv) morphological mutants altered at different stages of protonema and gametophore development, and (v) hormone (auxin or

cytokinin)-resistant mutants. Some of these classes have exhibited altered apical cell differentiation of the gametophore, indicating that they are useful for analyzing *P. patens* developmental mechanisms.

(3) Phytohormones and environmental factors regulate development of the gametophyte.

The development of *P. patens* is regulated by phytohormones (e.g., auxin and cytokinin) similar to those regulating angiosperm morphogenesis (reviewed in Schumaker and Dietrich, 1998). Developmental mutant analysis has shown that auxins and cytokinins regulate gametophyte development. Wild-type tissue grown on auxin-supplemented medium displays an increased production of caulonemal filaments. On the other hand, the presence of exogenous cytokinin induces an overproduction of buds, with up to 90% of caulonemal side branch initial cells developing a bud (McClelland, 1987). This indicates that cytokinin plays a critical role in the regulation of apical cell differentiation.

Light plays an important role in the normal development of *P. patens* and acts synergistically with phytohormones (Cove *et al.*, 1978; Jenkins and Cove, 1983a). Light is required for spore germination, chloronema and protoplast division, and bud formation. In addition, the direction from which light comes influences the growth of the gametophore apical cell (Jenkins and Cove, 1983a, 1983b). All of these processes can be used to study how phytohormones and light regulate the development of the apical cell.

(4) Molecular biology techniques for the study of *P. patens* are well established

The transformation of *P. patens* has been accomplished by both the PEG-mediated DNA transfer method that uses protoplasts (Schaefer *et al.*, 1991), and the particle

bombardment method (Sawahel *et al.*, 1992; Cho *et al.*, 1999). Recently, Schaefer and Zrÿd (1997) presented molecular evidence for the occurrence of highly efficient homologous recombination in the genomic DNA of *P. patens*. The gene-targeting technique, based on homologous recombination, was used successfully to analyze the biological function of as-yet unknown genes in *P. patens*. In particular, several genes were 'knocked out' by the gene-targeting technique, and their functions could be discussed according to the phenotype of the disruptants (Girke *et al.*, 1998; Girod *et al.*, 1999; Strepp *et al.*, 1998). To investigate gene function in detail, it is necessary to generate specific mutations of the gene, and this is only possible with use of the gene-targeting technique (Schaefer, 2001).

This study investigated the molecular mechanisms of gametophore apical cell differentiation in *P. patens* by identifying a gene that is expressed in the apical cell of the gametophore. Reverse genetics approaches, such as insertional mutagenesis, are useful for identifying genes that are involved in the shoot meristem. Gene-trap and enhancer-trap systems are the most powerful tools available. With these systems, a particular gene or enhancer that is responsible for cell-type or tissue-specific expression can be identified based on the expression of a reporter gene (Springer *et al.*, 2000). In *Arabidopsis*, these systems are highly exploited and have contributed to the understanding of developmental mechanisms by providing a number of cell-type or tissue-specific marker genes (Sundaresan *et al.*, 1995).

Currently, however, no methods exist for isolating a gene that is expressed in a particular cell-type or tissue of *P. patens*. Therefore, in this study, initial work was done to establish gene-trap and enhancer-trap systems that can isolate a gene based on its spatial expression pattern in *P. patens*. Due to the many advantages of using *P. patens* in developmental biology studies, as described above, the establishment of the

gene-trap and enhancer-trap systems is expected to be useful for studies of other developmental processes. The next step used the gene-trap and enhancer-trap systems to identify genes that are expressed in the apical cell of gametophores. In addition to the functional analysis of these genes in apical cell differentiation, the genes will be useful as molecular markers in other studies of apical cell differentiation. Chapter 1 details the establishment of the gene-trap and enhancer-trap systems, and Chapter 2 describes the identification of a gene that is predominantly expressed in the apical cell.

CHAPTER 1. ESTABLISHMENT OF GENE-TRAP AND ENHANCER-TRAP SYSTEMS IN THE MOSS *PHYSCOMITRELLA PATENS*

Introduction

The simple morphology of mosses makes them useful systems for studies in plant physiology and developmental biology (reviewed in Cove *et al.*, 1997; Schumaker and Dietrich, 1998; Reski, 1998). As haploids, mosses form a stem-leaf-like structure called a gametophore. This structure originates from a single apical cell at the gametophore apex, and its leaf is composed of a single or a few layer of cells that originate from a single leaf primordial cell. In gametophores, the development of multicellular organs, such as leaves and stems, results from the differentiation of a single cell, after asymmetric cell division of the apical cell occurs. Additionally, as in higher plants, gametophyte development is regulated by phytohormones (e.g., auxin and cytokinin) and by intracellular messengers (e.g., Ca^{2+}), thus providing an opportunity to analyze the actions of these substances in a simple system (reviewed in Schumaker and Dietrich, 1997).

Among mosses, *Physcomitrella patens* has been of particular interest recently, owing to the development of the gene targeting technique (Kammerer and Cove, 1996; Schaefer and Zrÿd, 1997; Scheafer 2001), which enables directed gene disruption and allele replacement (reviewed in Mengiste and Paszkowski, 1999). The ability to target gene disruption in *P. patens* is attributed to a high rate of homologous recombination that has not been observed in other land plants (Kammerer and Cove, 1996; Schaefer and Zrÿd, 1997). This technique has been used successfully to identify the functions of certain nuclear-encoded genes (Girke *et al.*, 1998; Girod *et al.*, 1999; Strepp *et al.*, 1998).

Gene-trap and enhancer-trap systems are useful tools for finding novel genes and unidentified regulatory elements based on the expression pattern of a reporter gene randomly integrated into the genome (reviewed in Bellen, 1999; Springer, 2000). In the gene-trap system, a reporter gene lacking promoter sequences is inserted randomly into the genome, and genes expressed in a tissue-specific manner are detected. Tissue-specific enhancers and promoters are obtained by the enhancer-trap system by cloning a region that borders an integrated reporter gene fused to a weak promoter.

Previously, shuttle mutagenesis (Seifert *et al.*, 1986; Ross-Macdonald *et al.*, 1997) was employed to insert DNA fragments randomly in the *P. patens* genome and a gene-tagging system was established (Nishiyama *et al.*, 2000). Using this system, it was also demonstrated that the splice acceptor-fused reporter gene was activated *in vivo* in certain tagged lines, indicating that shuttle mutagenesis is potentially useful in gene-trap and enhancer-trap systems. In this study, two gene-trap and two enhancer-trap elements for *P. patens* were constructed, and the utility of each element to generate large numbers of gene-trap and enhancer-trap lines was examined. Using the elements, 235 gene-trap and 1073 enhancer-trap lines from 9363 transgenic *P. patens* plants were obtained. In order to test the usefulness of the lines, it is demonstrated that a gene trapped using these elements can be easily identified.

Material and Methods

Culture conditions and transformation

Physcomitrella patens (Hedw.) Bruch and Schimp subsp. *Patens* Tan was derived from a wild type strain cultured from a single spore isolated in Gransden, Wood, Huntingdonshire, England, in 1962, by Dr. H.L.K. Whitehouse (Tan, 1979; Ashton and Cove, 1977).

P. patens was grown on solid medium overlaid with cellophane (Futamura Industry, Nagoya, Japan) in 9 cm-dishes and incubated at 25°C under continuous daylight supplied by FL40L fluorescent tubes (NEC, Tokyo, Japan) giving a light flux of 50 μmol m^{-2} sec^{-1}. The minimal medium, named BCD, in this study is a modified Knop medium (Ashton and Cove, 1977). BCD medium contains 1.8 mM KH_2PO_4 (pH6.4 adjusted with 4M KOH), 1.0 mM $MgSO_4$, 45 μM $FeSO_4$, 10 mM KNO_3, and 1 ml of alternative TES solution per liter. Alternative TES solution contains 0.22 mM $CuSO_4$, 10 mM H_3BO_4, 0.23 mM $CoCl_2$, 0.10 mM Na_2MoO_4, 0.19 mM $ZnSO_4$, 2.0 mM $MnCl_2$, and 0.17 mM KI. Routinely BCD medium plus 1 mM $CaCl_2$ was used. BCDAT medium is BCD medium plus 5 mM d-ammonium (+) tartrate and 1 mM $CaCl_2$. BCDATG medium is BCDAT medium plus 0.5% (w/v) glucose. All mediums were solidified with 0.8% (w/v) high gel strength agar (Sigma, St. Louis, USA) or Agar Powder (Nacalai tesque, Kyoto, Japan) and sterilized by autoclaving at 120°C for 20 min.

PEG-mediated transformation of *P. patens* was performed according to Schaefer *et al.* (1991). To isolate protoplasts, 5 days-old protonema (mainly chloronema) was incubated at 25°C for 30 min in a solution of 4% (w/v) Driserase (Kyowa Hakko Kugyo, Tokyo, Japan) and 8% (w/v) mannitol, and filtered through 50 μm nylon mesh. Freshly isolated protoplasts were centrifuged (2 min, 180 x g) and suspended in 8% (w/v) mannitol. This washing procedure was repeated twice. Finally suspended protoplasts

were counted with hemocytometor, and resuspended at 1.6 x 10^6 ml^{-1} in MMM solution [15 mM $MgCl_2$, 0.1% (w/v) MES, 8% (w/v) mannitol, pH 5.6]. Protoplast suspension (300 µl) and PEG solution [300 µl; 4% (w/v) PEG6000, 100 mM $Ca(NO_3)_2$, 100 mM Tris-HCl, pH 8.0] was added to Falcon 2057 tube (Becton Dickinson, Franklin Lakes, USA) containing 30 µl of plasmid DNA and mixed gently. The plasmids were purified with phenol/chloroform extraction, followed by PEG precipitation, and were resuspended in TE at a concentration of 1 µg $µl^{-1}$. The transformation mixture was incubated at 45°C for 5 min and then at 20°C for 10 min. Subsequently, the solution was diluted by adding 300 µl of protoplast liquid medium [5 ml $Ca(NO_3)_2$, 1 mM $MgSO_4$, 45 µM $FeSO_4$, 0.18 mM KH_2PO_4 (pH 6.5 adjusted with KOH), 1 ml of alternative TES per liter, 5 mM d-ammonium (+) tartrate, 6.6% (w/v) mannitol, and 0.5% (w/v) glucose] every 5 min at 5 times and then 1 ml of protoplast liquid medium every 5 min at 5 times. The diluted protoplast solution was poured into a 6 cm-dish and incubated at 25°C overnight in darkness. Protoplasts were centrifuged (2min, 180 x g), resuspended in 8 ml of PRM/T medium [BCD medium supplemented with 5 mM d-ammounium tartrate, 10 mM $CaCl_2$, and 8% (w/v) mannitol]. 2 ml of protoplast suspension was poured on a 9 cm-dish containing PRM/B medium [BCD medium supplemented with 5 mM d-ammonium tartrate, 6% (w/v) mannitol and 9 mM $CaCl_2$], and incubated at 25°C for 3 days under continuous daylight with a light flux of 50 µmol m^{-2} s^{-1}. Cultures were transferred to BCDAT medium supplemented with 50 mg l^{-1} or 20 mg l^{-1} G418 (GIBCO-BRL, Rockville, MD, USA) for the selection of transformants, and incubated at 25°C for 4 weeks under continuous daylight. Each transformant regenerated from a protoplast under G418 selection was transferred to antibiotic-free BCDAT medium and cultured at 25°C for 1 week. After an incubation of 1 week, each transformant was re-transferred to BCDAT medium supplemented with 50 mg l^{-1} or 20

mg l^{-1} G418 (GIBCO-BRL). Transformants that were able to survive on selection medium were selected as stable transformants.

Construction of gene-trap and enhancer-trap elements and mini-transposons

The NHI-GT and NHI-ET plasmids were constructed using pTN3, the pBluescript SKII+ plasmid (Stratagene, La Jolla, CA, USA) equipped with the NPTII expression cassette (Nishiyama *et al.*, 2000). As a splice acceptor in NHI-GT, the fourth intron of the *Arabidopsis* G protein gene, *GPA1*, was PCR-amplified using the int-F primer (5'-GCAAGCTTAGGGATGAAAACGGTCG-3') and the int-R primer (5'-GGACCTGCATATAACCTGCATATAAC-3'), and the pWS32 plasmid as template (Sundaresan *et al.*, 1995). A fragment containing the CaMV 35S promoter, *uidA* gene, and nos terminator was excised from pBI221 (Clontech Laboratories, Palo Alto, CA, USA) as an *Eco*RI/*Hin*dIII fragment, and cloned into the *Eco*RI and *Hin*dIII sites of pTN3. This plasmid was designated pCaMV35S. pCaMV35S was digested with *Hin*dIII and *Sma*I to remove the CaMV 35S promoter, blunt-ended, and religated to make pdeltaGT3. To construct NHI-GT, the PCR-amplified *GPA1* intron was digested with *Hin*dIII to cut one end of the product, and then ligated into *Hin*dIII/*Sma*I-digested pCaMV35S. A short core sequence (–46 to +1) of the CaMV 35S promoter was amplified by PCR using the 35S mini-F primer (5'-CTAAGCTTCGCAAGACCCTTCCTC-3') and the GUSseq primer (5'-TCACGGGTTGGGGTTTCTAC-3') and pBI221 as template. To make NHI-ET, the resulting PCR fragment was digested with *Hin*dIII and *Sma*I and ligated into the *Hin*dIII/*Sma*I-digested pCaMV35S.

The HI-ET mini-transposon was constructed using the pTn plasmid (Seifert *et al.*, 1986). A synthetic adapter containing *Hin*dIII, *Sac*II, *Kpn*I, *Cla*I, and *Sal*I cleavage sites was inserted between the *Hin*dIII and *Sal*I site of pTn to make pTn3-5.

A *Cla*I/*Sac*II fragment containing the CaMV 35S promoter core sequence, *uidA* gene, nos terminator, and NPTII expression cassette was excised from NHI-ET and ligated into the *Cla*I and *Sac*II sites of pTn3-5, to produce HI-ET. HI-ET was moved to the F derivative, pOX38, as described by Seifert *et al.* (1986).

Shuttle mutagenesis

Shuttle mutagenesis of *P. patens* was performed as described (Nishiyama *et al.*, 2000). *P. patens* genomic DNA was isolated from protonemata by the CTAB method (Murray and Thompson, 1980). The genomic DNA was partially digested with *Sau*3AI, and run on 0.6% (w/v) SeaKem GTG agarose (FMC BioProducts) gels. The 3 to 6 kb fragments were recovered by the electroelution method (Sambrook *et al.*, 1989), and their ends were partially filled with dATP and dGTP. These fragments were ligated with the *Sal*I-digested pHSS-Sal vector (Burns *et al.*, 1994) partially filled with dCTP and dTTP. The ligation solution was used to transform *E. coli* XL2 blue MRF' (Stratagene). Plasmids extracted from XL2 blue MRF' were subsequently used to transform *E. coli* DH5α (GIBCO-BRL). Plasmids extracted from DH5α were further used for the transformation of *E. coli* RDP146/pLB101. The *E. coli* RDP146/pLB101 containing the *P. patens* genomic DNA fragments was mated with *E. coli* RDP146/pOX38::HI-GT to form *E. coli* RDP146/pLB101/pOX38::HI-GT and to mutagenize the *P. patens* genomic library with the HI-GT mini-transposon, which was then mated with *E. coli* NG135 (Ross-Macdonald *et al.*, 1997). To mutagenize the genomic library using HI-ET, *E. coli* NS2114Sm (Seifert *et al.*, 1986) was used instead of NG135. The mutagenized library recovered from NG135 or NS2114Sm was amplified in *E. coli* XL2 blue MRF', because NG135 and NS2114Sm are *endA*$^+$ strains. The plasmids extracted from XL2 blue MRF' were digested with *Not*I, extracted with phenol and chloroform, precipitated with ethanol, and diluted with TE. The concentration of DNA was adjusted to 1.0 μg l^{-1}.

DNA and RNA extractions

For DNA extraction, propagated protonemata were cultured on BCDATG medium (Nishiyama *et al.*, 2000) under continuous light for two weeks, and then kept in the dark for two days. For RNA extraction, vegetatively propagated protonemata were grown on BCDATG medium for 13 days under continuous light. Tissue was ground to a powder in liquid nitrogen. Genomic DNA was extracted with a Nucleon Phytopure plant and fungal DNA extraction kit (Amersham-Pharmacia Biotech, Buckinghamshire, UK), and further purified by the CTAB method (Murray and Thompson, 1980). Total RNA was extracted according to Hasebe *et al.* (1998). The tissue powder was mixed with GuSCN buffer [4M GuSCN, 1M NH_4SCN, 100 mM Tris-HCl, 1% (w/v) sarcosine, 0.5% (w/v) polyvinylpolypyrrlidone, 1% (v/v) 2-mercaptoethanol] and extracted with chloroform/isoamylalcohol (24:1). After centrifugation, the supernant was extracted with chloroform/isoamylalcohol (24:1). The chloroform/isoamylalcohol extraction was carried out for totally 4 times. The aqueous phase was precipitated with ethanol and dissolved in CTAB buffer [2% (w/v) CTAB, 50 mM Tris-HCl, 5 mM NaCl, 0.84 M NaCl, pH 7.5] supplemented with 1% (v/v) 2-mercaptoethanol, and extracted with chloroform/isoamylalcohol (24:1). After centrifugation, the aqueous was added by 1/10 vol. of 10% CTAB solution [10% (w/v) CTAB, 0.7 M NaCl], extracted with chloroform/isoamylalcohol (24:1). After centrifugation, the aqueous was precipitated with CTAB precipitation buffer [1% (w/v) CTAB, 50 mM Tris-HCl, 5 mM EDTA, pH 7.5] supplemented with 0.5% (v/v) 2-mercaptoethanol. The precipitation was dissolved in high salt-TE (10 mM Tris-HCl, 1mM, EDTA, 1 M NaCl, pH 8.0) supplemented with 0.5% (v/v) 2-mercaptoethanol and mixed with ethanol to precipitate nucleic acids. The precipitated nucleic acids were dissolved in TE. Total RNA was then purified from the nucleic acids solution with ISOGEN-LS (Wako Pure chemical, Osaka, Japan) according

to manufacture's instructions. Poly(A)$^+$ RNA was further purified from total RNA using Dynabeads mRNA purification kit (DYNAL, OSLO, Norway) according to the manufacture's instructions.

Southern and Northern analyses

For Southern analysis, 2 μg of *P. patens* genomic DNA were digested, separated on 0.7% (w/v) SeaKemGTG agarose (BME, Rockland, ME, USA), and transferred to a Hybond N$^+$ nylon membrane (Amersham-Pharmacia Biotech). For Northern analysis, 0.5 μg of poly(A)$^+$ RNA was electrophoresed on 1.0% (w/v) agarose gels containing formaldehyde/MOPS and transferred to Hybond N$^+$ nylon membrane. The *uidA* DNA fragment was PCR-amplified with the 35S mini-F (5'-CTAAGCTTCGCAAGA CCCTTCCTC-3') and NOSter (5'-ATGTTTGAACGATCGGGGAAAT-3') primers using NHI-ET as a template. The amplified fragment was purified from an agarose gel using GENECLEAN III (BIO101, Vista, CA, USA) and labeled using a Random Primer DNA labeling kit ver. 2.0 (TAKARA, Tokyo, Japan). The 804 bp *Hin*dIII/*Xho*I fragment of the 209-7 DNA fragment was also excised and labeled as described. For both Southern and Northern analyses, hybridization was performed in 0.5 M Church's phosphate buffer (0.5 M Na_2HPO_4 adjusted to pH 7.2 with H_3PO_4; Church and Gilbert, 1984) at 65°C. Northern blots were reprobed with *P. patens* glyceraldehyde 3-phosphate dehydrogenase gene (*GAPDH* probe) to control for loading (Leech *et al.*, 1993); the *Pst*I-*Eco*RI fragment of the pPpGapC plasmid (provided by R. Kofuji) was excised, labeled, and hybridized as described above.

RACE

For 5' RACE of the YH209 gene-trap line, total RNA was treated with RNase-free DNaseI (GIBCO-BRL), and first strand cDNA was synthesized with 5' RACE System

using SuperScriptII reverse transcriptase (GIBCO-BRL) and the GUS R4 primer (5'-ATTGACCCACACTTTGCCGTAATGAGTGAC-3'). The synthesized cDNA was amplified with the anchor primer supplied by GIBCO-BRL and the GUS R3 primer (5'-TCTTGTAACGCGCTTTCCCACCAACGCTGA-3') using ExTaq DNA polymerase (TAKARA). PCR cycle conditions consisted of 5 min at 94°C, then 25 cycles of 94°C for 30 sec, 62°C for 30 sec, and 72°C for 3 min, followed by 10 min at 72°C. Further PCR was carried out with the UAP primer (5'-CUACUACUACUAGGCCACGCGTCGACTAGTAC-3') supplied by GIBCO-BRL and the 5'GUS1 primer (5'-CAUCAUCAUCAUTTGGCGATCCAGACTGAATG-3'). PCR conditions were as in the previous reaction, except the annealing temperature was lowered to 59°C. PCR products were purified with SUPREC-02 (TAKARA), and cloned into pIMAz2. pIMAz2 was constructed by PCR amplification of pGEM3z (Promega, Madison, WI, USA) using the prIMAz2-2- primer (5'-AUGAUGAUG AUGCCTCTAGAGTCGACCTGCAGGCA-3') and the prIMAz2-1+ primer (5'-UAGUAGUAGUAGCCCGGGTACCGAGCTCGAA-3') to add the sequences complementary to CAUCAUCAUCAU and CUACUACAUCAU to each end. The PCR fragment amplified with the UAP and 5'GUS1 primers was annealed with pIMAz2 using uracil DNA glycosylase (GIBCO-BRL).

For 3' RACE, to obtain the 3' region of the 1041 bp DNA fragment, cDNA was synthesized from 2 μg of total RNA using SuperScriptII reverse transcriptase with the adaptor primer containing polyT stretch (GIBCO-BRL). PCR was performed using this cDNA as template with the UAP primer and the 209-7p1 primer (5'-GGACATCAGAGAAGGAGGCG-3'). PCR cycle conditions consisted of 3 min at 94°C, 30 cycles of 94°C for 30 sec, 58°C for 30 sec, and 72°C for 3 min. Additional PCR was carried out with the UAP primer and the 209-7p2 primer (5'-CGCATGGAGTGTTGATGC-3'). PCR conditions were as in the previous PCR,

except the annealing temperature was 59°C. Finally, the entire cDNA was amplified with YH209LF190 (5'-CGGAGCGGTGGATTTTTTCTC-3') and YH209LR3030 (5'-GGATGAACACCCACTCAAGTACC-3') primers. PCR conditions were the same as in the previous PCR. The PCR product was isolated from an agarose gel using GENECLEAN III and cloned into pGEM-T (Promega).

DNA sequencing

DNA sequences were determined using a BigDye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems, Foster City, CA) and an ABI PRISM 377 DNA sequencer (Applied Biosystems). Analysis of DNA sequence was performed with GENETYX-MAC software, version 11.0 (Software Development, Tokyo, Japan).

Phylogenetic analysis

Related protein sequences were identified by performing a BLASTP search at NCBI against the nr dataset. The amino acid sequences were aligned using CLUSTAL W, version 1.81 (Thompson *et al.*, 1994; optimized for a multiprocessor by SGI) obtained at http://www.sgi.com/chembio/resources/clustalw/parallel_clustalw.html, and revised manually. To search for the Maximum-Likelihood (ML) tree, I used Neighbor-Joining (NJ) and most parsimonious (MP) trees as the starting trees for a local rearrangement search (Adachi and Hasegawa, 1996). ML distances under the JTT model (Jones *et al.*, 1992) were calculated using ProtML (Adachi and Hasegawa, 1996). MP trees were searched by the tree bisection reconnection algorithm using PAUP* version 4.0 beta 6 (Swofford, 2001) on a PowerPC-based Macintosh. The local bootstrap probability of each branch was estimated by the resampling-of-estimated-log-likelihood (RELL) method (Kishino *et al.*, 1990; Hasegawa and Kishino, 1994).

GUS assay

Histochemical detection of β-glucuronidase (GUS) activity in gametophyte tissues was performed according to Jefferson *et al.* (1987) with slight modifications. Tissues were cultured for 14 days on BCTAT medium under continuous illumination of 50 μmol m^{-2} sec^{-1}. Tissues were fixed in a solution of 0.3% (v/v) formalin, 0.3 M mannitol and 0.2% (w/v) MES (pH5.6) for 30 min at room temperature and then washed 3 times with 50 mM NaH_2PO_4 (pH7.0). Tissues were infiltrated for 30min in a substrate solution [50mM NaH_2PO_4 (pH7.0), 0.5 mM 5-bromo-4-chloro-3-indolyl β-D-glucurunic acid (X-Gluc, Wako Pure Chemical, Osaka, Japan) 0.5 mM $K_3Fe(CN)_6$, 0.5 mM $K_4Fe(CN)_6$ and 0.3 % Triton X-100] and then were incubated at 37°C for 36-48 h. After the incubation, tissues were fixed in 5% (v/v) formalin for 10min and soaked in 5% (v/v) acetic acid for 10min. Then the tissues were dehydrated through an ethanol series. Finally the stained tissues were observed under a Leica MZAPO microscope (Leica Mikroskopie und Systeme GmbH, Wetzlar, Germany) and a Leica DMLB microscope (Leica Mikroskopie und Systeme GmbH). Images of the stained tissues were digitized with a CCD Camera (FUJIX HC-300z, Fuji Film, Tokyo, Japan).

GUS transient assays in protoplasts were performed according to Jefferson *et al.* (1987) with minor modifications. Protoplasts that were cultured for 48 h after PEG-mediated transformation were collected by centrifugation, and suspended in extraction buffer [50 mM NaH_2PO_4 (pH7.4), 1 mM EDTA, 0.1% (v/v) Triton X-100, 0.1% (w/v) Sarkosyl and 10 mM 2-mercaptoethanol]. Protoplasts were lysed by 3 cycles of freeze and thaw and then centrifuged. The supernatant was recovered as a crude extract. GUS enzyme assays were performed in triplicates in microtiter plates. GUS activities were determined using the substrate 4-methylumbelliferone glucuronide (MUG, Wako chemical), and the fluorescent product (4-methyl-umbelliferone, MU) was quantified using a CORONA 130F fluorescence plate reader (CORONA ELECTRIC, Ibaraki, Japan), based on fluorescence with an excitation wavelength of 360 nm and emission

wavelength of 460 nm. Protein content was determined using a BioRad Protein Assay kit (BioRad Laboratories, Hercules, CA, USA). Specific GUS enzyme activities were calculated as pmol MU produced per min of reaction per mg protein.

Results

Gene-trap and enhancer-trap elements

Each gene-trap and enhancer-trap element was introduced into *P. patens* by two methods: (1) non-homologous integration and (2) homologous integration using shuttle mutagenesis (Seifert *et al.*, 1986; Nishiyama *et al.*, 2000). For the former method, protoplasts of *P. patens* were transformed with each trap element lacking any *P. patens* genomic DNA sequence (Schaefer *et al.*, 1991), and trap elements were integrated into the *P. patens* genome by non-homologous recombination. The latter method employed homologous recombination to integrate a trap element into the *P. patens* genome and had a transformation efficiency approximately 10-fold higher than non-homologous integration (Schaefer and Zrÿd, 1997). Shuttle mutagenesis consists of three main steps: (1) generation of a *P. patens* genomic DNA library in *Escherichia coli*, (2) introduction of a mini-transposon including each trap element to the *P. patens* genomic DNA in *E. coli*, and (3) transformation of *P. patens* protoplasts with *P. patens* genomic DNA containing a mini-transposon insertion.

The gene-trap and enhancer-trap elements used in this study are shown in Figure 2. The plasmids NHI-GT (Non-Homologous Integration-Gene-Trap) and NHI-ET (Non-Homologous Integration-Enhancer-Trap) were used for non-homologous integration and HI-GT (Homologous Integration-Gene-Trap) and HI-ET (Homologous Integration-Enhancer-Trap) were used for homologous integration. HI-GT under the name of mTn-3xHA/GUS2 (Nishiyama *et al.*, 2000), HI-ET, NHI-GT, and NHI-ET were constructed in this study. The NPTII expression cassette (nptII, Nishiyama *et al.*, 2000), containing the cauliflower mosaic virus (CaMV) 35S promoter (Odell *et al.*, 1985), *ntpII* gene (Beck *et al.*, 1982), and CaMV polyadenylation signal (Guerineau *et al.*, 1990), was used as a marker for selection of transformed *P. patens*. Each element

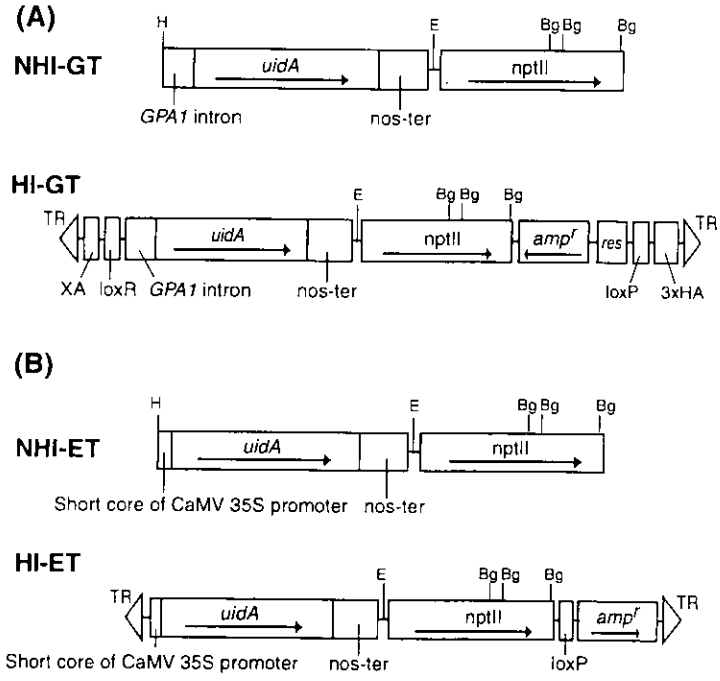


Figure 2. Schematic diagrams of gene-trap (A) and enhancer-trap (B) elements. Each element contains the *uidA* coding region (*uidA*), the nopaline synthase polyadenylation signal (nos-ter), and an NPTII expression cassette (nptII). Gene-trap elements contain the Arabidopsis *GPA1* intron (Sundaresan *et al.*, 1995). Enhancer-trap elements contain a short core promoter (–46 to +1) of the CaMV 35S promoter. HI-GT and HI-ET are mini-transposons containing 38-bp terminal repeats of the Tn3 transposon (TR), a loxP site (loxP), and an ampicillin-resistance gene (*amp*^{r}). A factor Xa protease cleavage site, a loxR stie, a Tn3 *res* site, and three HA epitopes are depicted as XA, loxR, *res*, and 3xHA, respectively. Arrows within *uidA*, nptII, and *amp*^{r} indicate their directions. *Eco*RI (E), *Bgl*II (Bg), and *Hin*dIII (H) sites are shown.

contains the NPTII expression cassette and results in resistance to G418 (Schaefer *et al.*, 1991). The *amp*^{r} gene was used as a selection marker in *E. coli*. The *uidA* gene, encoding β–glucuronidase (GUS, Jefferson *et al.*, 1987), was used as the reporter gene. The fourth intron of *Arabidopsis GPA1*, the alpha subunit of the G-protein gene (Sundaresan *et al.*, 1995), was fused to the 5' end of the reporter gene in the gene-trap elements (NHI-GT and HI-GT). This construction enables detection of gene expression when the gene-trap element inserts in either intron or exon sequences (Sundaresan *et al.*, 1995). Gene expression is limited to the right orientation of the inserted element and to the site of insertion with regard to the reading frame.

The enhancer-trap elements (NHI-ET and HI-ET) contain a short core sequence (–46 to +1) of the CaMV 35S promoter (Benfey *et al.*, 1989) fused to the 5' end of the *uidA* gene. To test whether the CaMV 35S core promoter functions as a weak promoter in *P. patens*, its promoter activity was assayed using the *P. patens* protoplasts into which each of the following gene-trap or enhancer-trap elements were transiently introduced: NHI-GT, NHI-GT lacking the *GPA1* intron (pdeltaGT3), NHI-ET, and pCaMV35S. NHI-GT and pdeltaGT3 were used to examine whether the *GPA1* intron confers enhancer activity in *P. patens*. In pCaMV35S, the short core promoter of NHI-ET was replaced with a longer region (–800 to +1) of the CaMV 35S promoter. All elements were cloned into pBluescript SKII+ (Stratagene, La Jolla, CA, USA). Protoplasts were prepared from protonemata and transformed by the polyethylene glycol (PEG) - mediated method (Schaefer *et al.*, 1991). After 48 hours of incubation in the dark at 25°C, GUS activity was measured (Table 1). Levels of GUS activity of pdeltaGT3 and NHI-GT were markedly lower than that of NHI-ET and p35CaMV35S. GUS activity of NHI-ET, containing the short core promoter, was approximately 10-fold lower than that of pCaMV35S, but approximately 5-fold higher than that of NHI-GT and pdeltaGT3.

Table 1. GUS activity of protoplasts transiently transformed with gene-trap and enhancer-trap elements

Constructs [a]	GUS activity [b] (pmol 4MU / min / mg protein)
No DNA	0.05 ± 0.01
NHI-GT	1.89 ± 2.58
pdeltaGT3	1.63 ± 0.26
NHI-ET	11.12 ± 5.60
pCaMV35S	128.98 ± 8.75

[a] The Schematic diagrams of NHI-GT and NHI-ET are shown in Figure 2. The pdeltaGT3 is a construct lacking a *GPA1* intron of NHI-GT. pCaMV35S contains a longer CaMV 35S promoter (–800 to +1) instead of its core promoter (–46 to +1) present in NHI-ET.

[b] Mean ± SD (n = 3)

Comparison of non-homologous and homologous integration methods

The efficiency of integration of the trap elements into the *P. patens* genome using the non-homologous and homologous integration methods was compared. *P. patens* transformants were divided into two categories: unstable and stable. Unstable transformants are resistant to antibiotics on the primary selection medium, but become sensitive on secondary selection after two weeks of culturing on non-selective medium. Stable transformants can grow on both primary and secondary selection media. Using non-homologous integration with 30 μg of NHI-GT or NHI-ET, 5 ± 3 (n = 4) and 7 ± 6 (n = 4) stable transformants were obtained, respectively. Homologous integration, using the HI-GT or HI-ET mini-transposons and the shuttle mutagenesis method, resulted in 94 ± 47 (n = 4) and 102 ± 48 (n = 4) stable transformants, respectively.

Expression patterns and copy number of the* uidA *gene in trap lines

All transformants were screened for GUS activity and it was detected histochemically in 8 of 267 stable transformants (3%) carrying NHI-GT, 18 of 146 stable transformants (12%) carrying NHI-ET, 227 of 5370 stable transformants (4%) carrying the HI-GT mini-transposon, and 1055 of 3580 stable transformants (30%) carrying the HI-ET mini-transposon. GUS expression patterns in protonemata and gametophores varied (Table 2); representative lines are shown in Figure 3.

To estimate the copy number of each trap element in a trap line, genomic Southern analysis of gene-trap lines generated using the HI-GT mini-transposon was performed. Genomic DNA preparations from 32 randomly selected lines were digested with *Eco*RI. The mini-transposon contains one *Eco*RI site (Figure 2). A PCR-amplified fragment of the mini-transposon *uidA* gene, from start to stop codons, was used as a *uidA*-specific probe. Based on hybridization patterns, the copy number of the mini-

Table 2. Expression patterns of *uidA* gene in gene-trap and enhancer-trap lines

Stained tissue	Non-homologous integration		Homologous integration	
	Gene-trap	Enhancer-trap	Gene-trap	Enhancer-trap
Gametophyte [a]	8 (100%)	18 (100%)	227 (100%)	1055 (100%)
Protonema				
Chloronema	5 (63%)	8 (44%)	49 (21%)	471 (44%)
Caulonema	3 (38%)	8 (44%)	24 (11%)	165 (16%)
Gametophore				
Bud	1 (13%)	1 (6%)	13 (6%)	55 (5%)
Leaf	5 (62%)	12 (67%)	110 (49%)	330 (31%)
Stem	1 (13%)	3 (17%)	17 (8%)	31 (3%)
Rhizoid	2 (25%)	2 (11%)	28 (12%)	57 (5%)
Apex [b]	2 (25%)	1 (6%)	14 (6%)	24 (2%)
Axillary hair	6 (75%)	12 (67%)	149 (66%)	884 (84%)

[a] Numbers of lines exhibiting GUS staining in listed tissues are given. The number in gametophyte is set as 100%. Since most of lines showed GUS positive staining in several different tissues, the sum of percentage became more than 100%.

[b]Apex indicates the number of lines showing GUS staining in an apical cell and leaf-primordial cells.

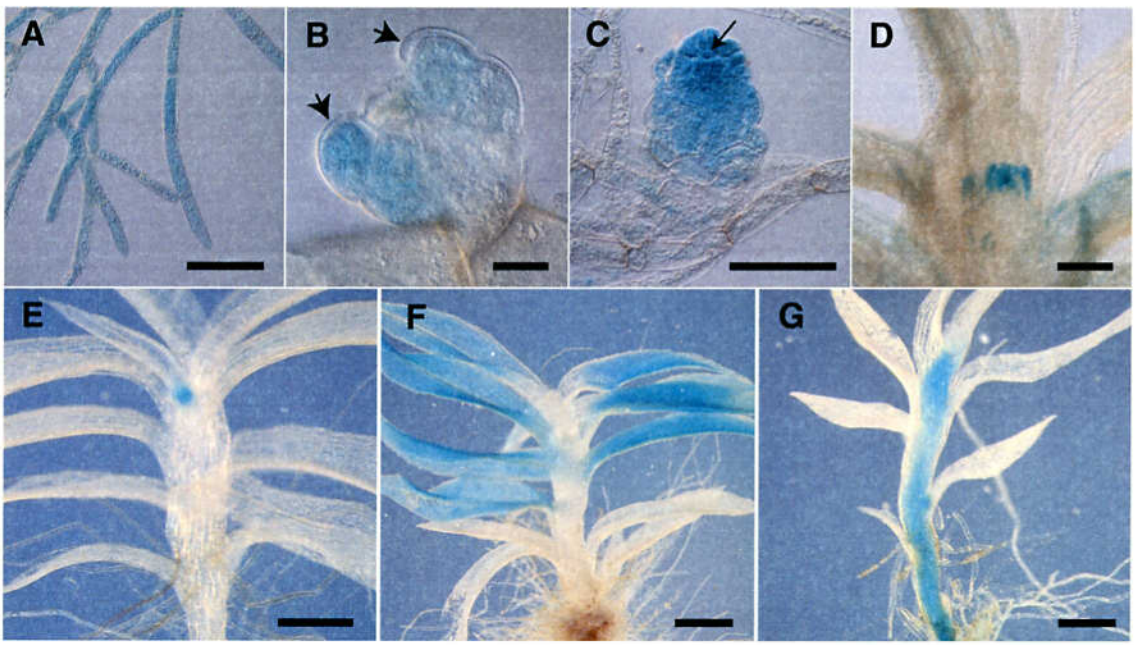


Figure 3. Histochemical detection of GUS activity in representative trap lines.
(A) Chloronema cells of gene-trap line YH261. (B) A bud of enhancer-trap line ET77. Leaf-primordial cells (arrows) are predominantly stained. (C) A bud of gene-trap line YH727. An apical cell (arrow) and its surrounding cells are stained. (D) Apical portion of a gametophore in gene-trap line YH8. Axillary hairs that differentiate at the adaxial base of the leaf are stained. (E) A gametophore of enhancer-trap line ET63. The apical portion of the gametophore is predominantly stained. (F) A gametophore of gene-trap line YH560. Leaves are stained. (G) A gametophore of enhancer-trap line ET21. A stem is stained. LinesYH261, YH727, YH8, and YH560 were generated using HI-GT, lines ET63 and ET21 were generated using HI-ET, and line ET77 was generated using NHI-ET. Bars in (A), (C), and (D) = 100 μm, in (B) = 20 μm, and in (E), (F), and (G) = 300 μm.

transposon in each trap line was estimated. Thirty trap lines contained multiple insertions, and two lines contained only an insertion of single copy (no. 19 and 25 in Figure 4).

Analysis of gene-trap line, YH209

To assess the feasibility of isolating a trapped gene, one line, YH209, generated using HI-GT, was studied further. In this line, GUS activity was observed in the basal part of the gametophore (Figure 5A). At high magnification, GUS activity could be seen predominantly in mature rhizoid cells proximal to the stem (Figure 5B). A rhizoid is a slender, multicellular filament attached to the stem of a gametophore. Rhizoid differentiation begins when an epidermal cell of a gametophore stem forms a protrusion, into which a nucleus moves. Then, cell division forms a rhizoid basal cell and a rhizoid cell (Figure 5C). The rhizoid cell divides several times to form filamentous cells. While the rhizoid is composed of four or less cells, the cells are colorless. When the rhizoid consists of more than approximately four cells, the basal cell and several rhizoid cells proximal to the stem have a brown pigmentation. GUS activity was observed in the pigmented rhizoid basal cell and in two or three pigmented rhizoid cells attached to the basal cell (Figure 5B), although pigmented rhizoid cells were not always stained. GUS activity was not observed in colorless rhizoid cells (Figure 5C).

Genomic DNA isolated from the line YH209 was digested with *Bgl*II and hybridized with the *uidA*–specific probe. Four *Bgl*II fragments were detected (Figure 6A), indicating that the line YH209 contains at least four copies of the HI-GT mini-transposon. Hybridization of poly(A)$^+$ RNA from the line YH209 with the *uidA*-specific probe detected a broad band of approximately 2.4 kb (Figure 6B). The size of *uidA* mRNA from the start to the stop codon is 1.9 kb long with 0.1 kb 3' untranslated

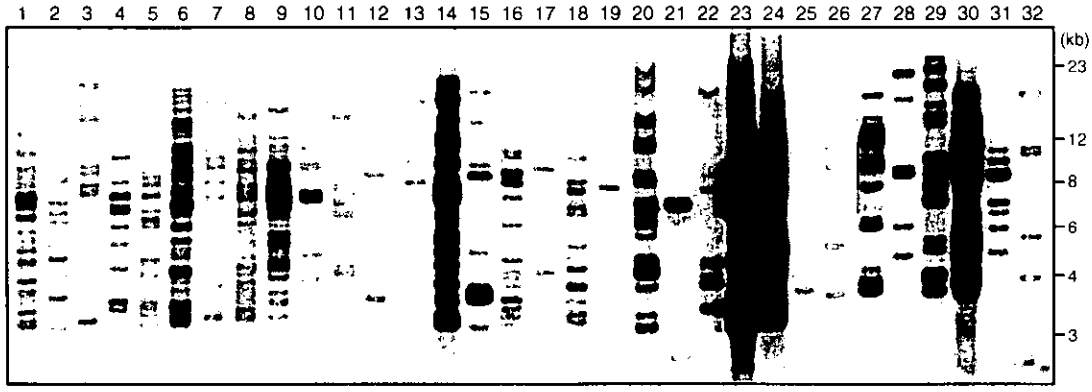


Figure 4. Estimation of the copy number of the *uidA* gene in gene-trap lines.
Genomic DNA from 32 gene-trap lines was digested with *Eco*RI and hybridized with *uidA*-specific probe. Lane 19 and 25 shows a single band, indicating that these lines contain a single copy of the *uidA* gene.

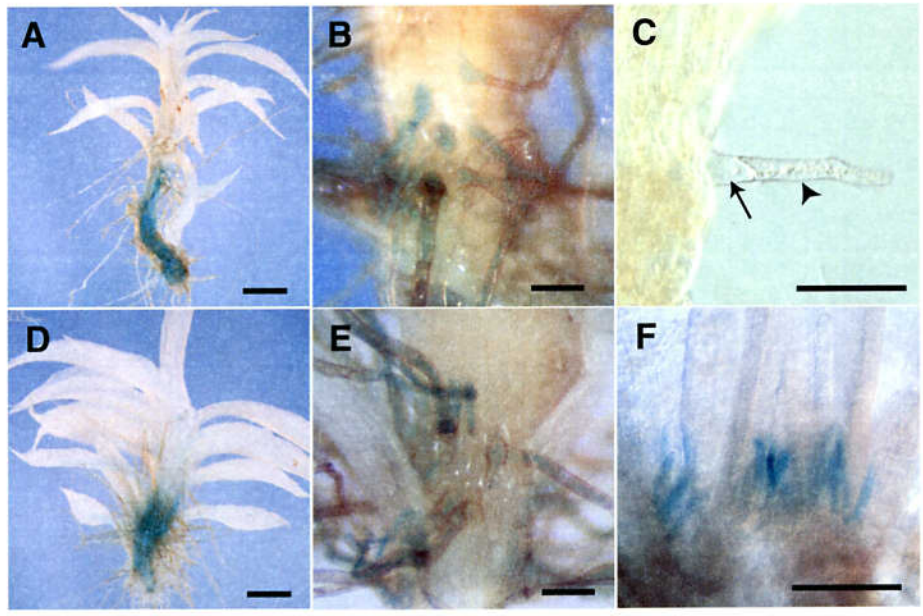


Figure 5. Histochemical detection of GUS activity in a gene-trap line YH209 (A-C) and the transformant line 0404-1 (D-F), in which the 209-8SC DNA fragment replaces its homologous site.

(A) A gametophore of the lineYH209. (B) A magnified view of the basal part of the YH209 gametophore, demonstrating staining of the pigmented rhizoid basal cells and rhizoid cells. (C) A young rhizoid of the line YH209. A basal cell (arrow) and a rhizoid cell (arrowhead) that lack pigmentation are not stained. (D) A gametophore of the 0404-1 line. (E) A magnified view of the basal part of the line 0404-1 gametophore showing that rhizoid basal cells and rhizoid cells are stained. (F) A magnified view of the apical portion of the line 0404-1, showing stained axillary hairs. Bars in (A) and (D) = 500 μm, in (B) and (E) = 200 μm, in (C) and (F) = 100 μm.

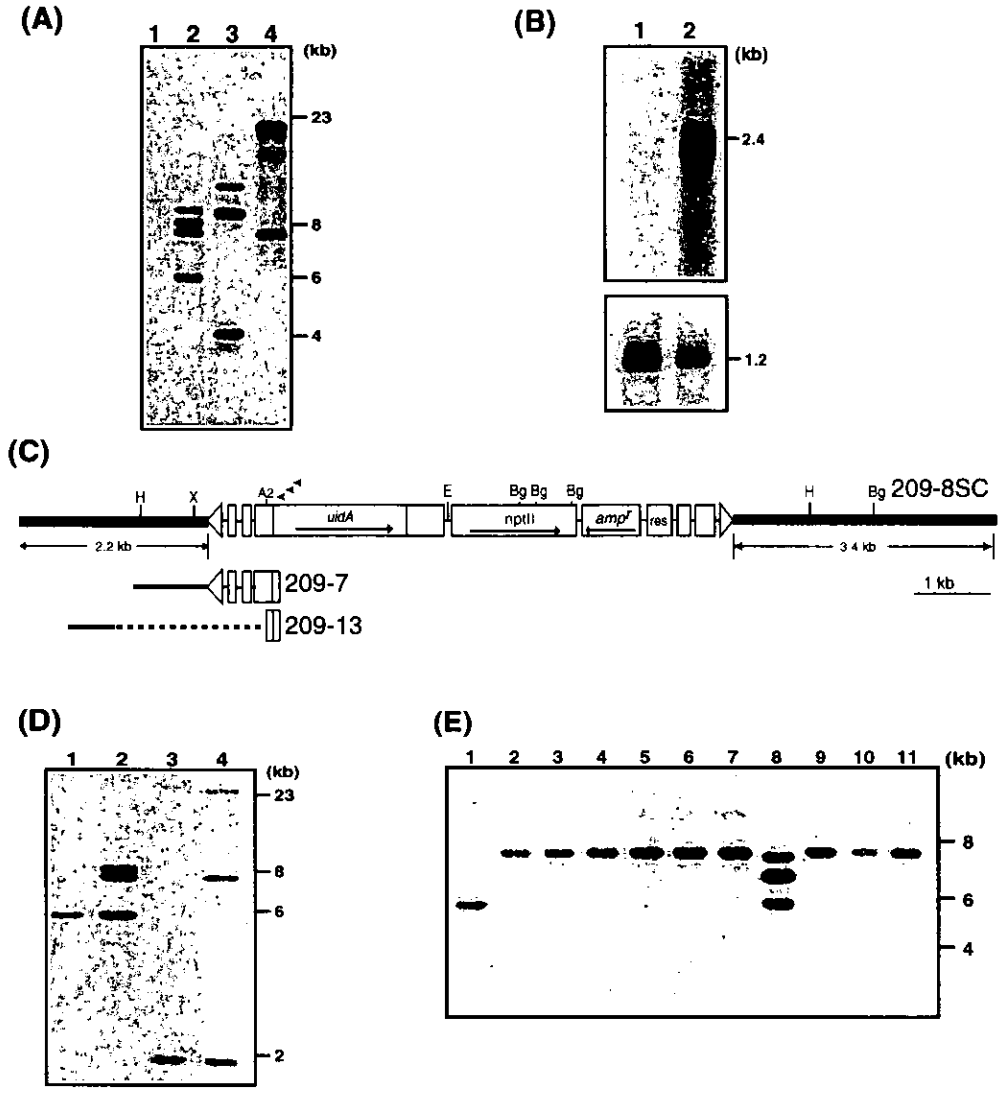


Figure 6. Molecular analysis of gene-trap line YH209, generated by the HI-GT mini-transposon.

Figure 6. (continued)
(A) Genomic Southern analysis of wild type (lane 1) and the line YH209 (lane 2 to 4). Blots were probed with a *uidA* gene-specific probe. Genomic DNA was digested with *Eco*RI (lane 1 and 3), *Bgl*II (lane 2) and *Hin*dIII (lane 4). (B) Northern analysis of wild type (lane 1) and the line YH209 (lane 2). Blots were probed with a *uidA* gene-specific probe (upper panel) and glyceraldehyde 3-phosphate dehydrogenase gene (*GAPDH* probe) as a control (lower panel). (C) Schematic representation of a transposon-tagged genomic fragment (209-8SC) and two 5' RACE products (209-7 and 209-13). Thick lines represent *P. patens* genomic DNA fragments. Locations of the *uidA* gene-specific primers used for 5' RACE are indicated with arrowheads. A2 indicates the second acceptor site of the *GPA1* intron. For the 5' RACE products, the 1041 bp and 433 bp of cDNA fragments corresponding to the genomic region in 209-8SC are indicated as black lines in 209-7 and 209-13, respectively. A dotted line in the 209-13 fragment shows the region, which may be alternatively spliced out unlike in the 209-7 fragment. *Eco*RI (E), *Bgl*II (Bg), *Hin*dIII (H), and *Xho*I (X) sites are shown. (D) Genomic Southern analysis of wild type and the line YH209. Genomic DNA from wild type (lanes 1 and 3) and the line YH209 (lanes 2 and 4) was digested with *Bgl*II (lanes 1 and 2) or *Hin*dIII (lanes 3 and 4). The 804 bp DNA fragment of the 209-7 product was used as a probe. (E) Southern analysis of wild type (lane 1) and stable transformants (lanes 2-11) with 209-8SC fragment. Genomic DNA was digested with *Bgl*II and probed with the 804 bp of DNA fragment of the 209-7. Bands of lane 2-11 were shifted, indicating that the wild type sequence was homologously replaced with the 209-8SC fragment in these lines. Multiple 209-8SC fragments should be integrated in the line of lane 8. Lane 6 corresponds to the line 0404-1, shown in Figures 5D – 5F.

region, indicating that a part of *P. patens* chromosomal genes was fused to the *uidA* gene.

Using *uidA* gene-specific primers and cDNA derived from gametophyte poly(A)$^+$ RNA of the line YH209, 5' RACE was performed and two candidate DNA fragments 209-7 and 209-13 were obtained (Figure 6C). Fragment 209-7 contained 1041 bp of *P. patens* DNA adjacent to the 5' end of the HI-GT mini-transposon. Fragment 209-13 contained 433 bp of *P. patens* DNA adjacent to the 5' end of a partial HI-GT mini-transposon that begins with the second acceptor site of the *GPA1* intron. The 804 bp fragment generated by *Hin*dIII and *Xho*I digestion of *P. patens* DNA in 209-7 was used as a probe for colony hybridization screening of the tagged genomic library used for *P. patens* transformation. Sequencing of a positive clone, containing an 11.6 kb mini-transposon-tagged fragment named 209-8SC, revealed that HI-GT separated a 5.6 kb *P. patens* genomic DNA fragment into 2.2 and 3.4 kb fragments (Figure 6C). The *P. patens* DNA fragments found in both the 209-7 and 209-13 clones were present in the 2.2 kb fragment (Figure 6C).

To see whether the 209-8SC fragment integrated into homologous or non-homologous sites in line YH209, *P. patens* genomic DNA digested with *Bgl*II or *Hin*dIII was hybridized with the 804 bp DNA fragment of 209-7. A single band was detected in wild-type genomic DNA digested with either enzyme, while two extra bands were detected in digests of line YH209 genomic DNA (Figure 6D). *Bgl*II and *Hin*dIII do not cut within the 804 bp fragment used as a probe. Thus, it was concluded that two copies of 209-8SC integrated into non-homologous sites.

To determine whether the GUS activity of the line YH209 reflects the endogenous expression pattern of a gene encoded on 209-8SC, 209-8SC was introduced to *P. patens* by homologous recombination. Genomic DNA was extracted from ten stable transformants, digested with *Bgl*II, and hybridized with the 804 bp probe

fragment. The band hybridizing to the wild type DNA was shifted in nine out of ten lines, clearly indicating that 209-8SC is integrated in a homologous site in these nine lines (Figure 6D). GUS activity in the nine lines was examined (Figures 5D and 5E). The observed expression patterns were similar to those of the line YH209, except for additional activity at axillary hairs (Figure 5F).

The 1041 bp *P. patens* DNA fragment of 209-7 contained an expected start codon, as deduced from comparisons with the alpha-xylosidase and acidic alpha-glucosidase genes. To clone the 3' region of the 1041 bp DNA fragment, 3' RACE was performed, and a 2198 bp fragment was obtained. According to the nucleotide sequence of the 1041 bp DNA fragment and the 3' RACE product, primers were designed to match the 5' and 3' ends of the fragments and were used to amplify the entire cDNA. The isolated cDNA was sequenced and found to encode a putative protein of 916 amino acids with high similarity to members of the glycoside hydrolase family 31 (Frandsen and Svensson, 1998) (Figure 7A). The gene was named *Physcomitrella patens alpha-glucosidase-like gene (PpGLU).* The protein encoded by *PpGLU* contains two conserved signature regions of the glycoside hydrolase family 31, including the putative catalytic aspartic acid residues (Frandsen and Svensson, 1998) (Figure 7B).

To determine the relationship between *PpGLU* and other glycoside hydrolase family 31genes, a phylogenetic analysis was performed. A BLAST search identified 85 genes with high similarity (*E*-value < $1e^{-5}$), and the sequences were aligned with *PpGLU*. After exclusion of short or redundant sequences, the phylogeny was inferred using 175 amino acid sites of 64 sequences by the maximum-likelihood method. A distinct clade of eukaryotic genes with 100% local bootstrap support was found; in this clade, 12 land

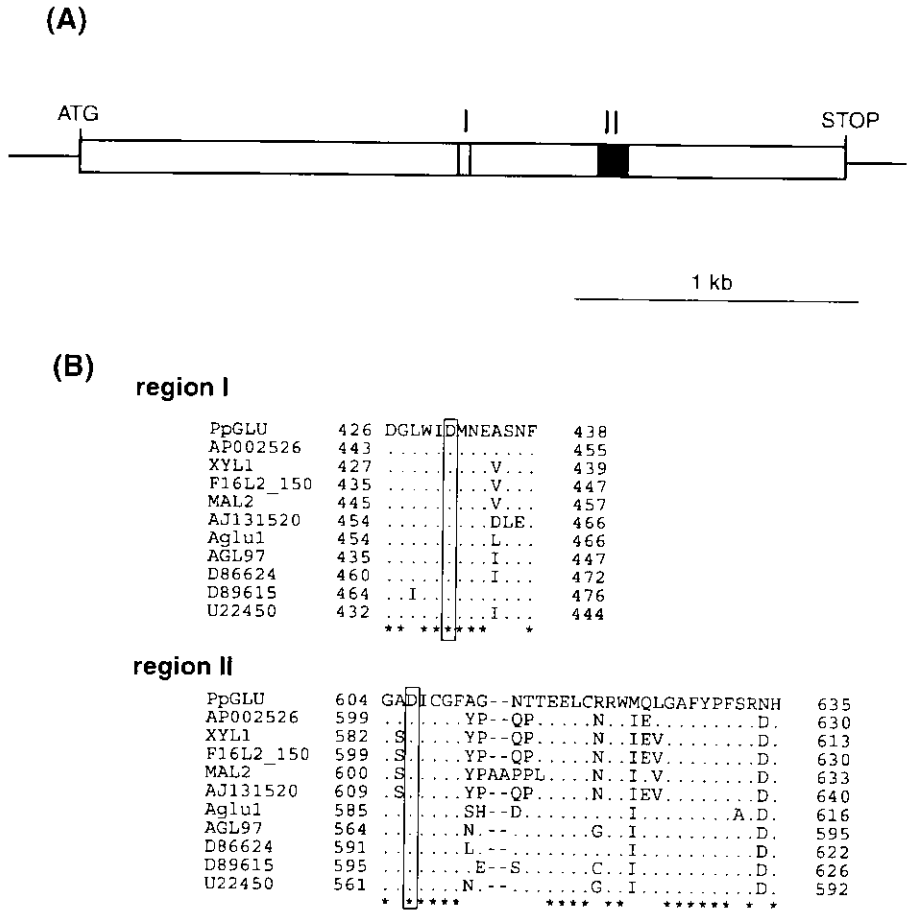


Figure 7. Structure of *PpGLU*.

(A) Schematic representation of *PpGLU* cDNA. Lines indicate 5' or 3' untranslated regions, and boxes represent coding regions. The conserved signature regions I and II are shown with a dark box and a black box, respectively. (B) Alignment of amino acid sequences in the conserved signature regions I and II of the glycoside hydrolase family 31 (Frandsen and Svensson, 1998). Asterisks indicate identical amino acids in all members and the putative catalytic site is boxed. Identical amino acid residues to PpGLU are shown as dots and gaps as dashes, respectively. Accession numbers of the aligned genes are: *PpGLU* (AB057452), *XYL1* (AF087483), F16L2_150 (AL162459), *MAL2* (AJ277244), *Aglu1* (AF014806), and *AGL97* (AF118226). Other accession numbers show the following: AP002526 (EST of *Oryza sativa*), AJ131520 (alpha-xylosidase of *Tropaelim majus*), D86624 (alpha-glucosidase of *Spinacia oleacea*), D89615 (alpha-glucosidase of *Beta vulgaris*), and U22450 (alpha-glucosidase of *Hordeum vulgare*).

plant, 15 metazoan, and 9 fungal genes each formed a clade. The *PpGLU* gene was positioned in the land plant clade with acidic alpha-glucosdiase. To obtain better resolution, further analysis of the relationships within the plant gene clade was performed using 442 amino acid sites of the 36 eukaryotic sequences, with the *Tetrahymena pyriformis* gene (accession number D83384) as an outgroup. After local rearrangements starting from the Neighbor-Joining tree and the 5 most parsimonious trees, which contained 3517 steps, a tree with a log likelihood of – 18911.96 ± 476.56 was obtained as the Maximum-Likelihood tree (Figure 8). In this tree, flowering plant alpha-glucosidase genes were divided into two clades supported with high local bootstrap support (89 and 100%). Each clade included both eudicot and monocot genes. *PpGLU* was positioned at the base of a clade that included the *Arabidopsis* alpha-xylosidase gene *XYL1* and the potato alpha-glucosidase gene *MAL2*, with 67% local bootstrap support. Characterization of the *XYL1* and *MAL2* genes has not been published.

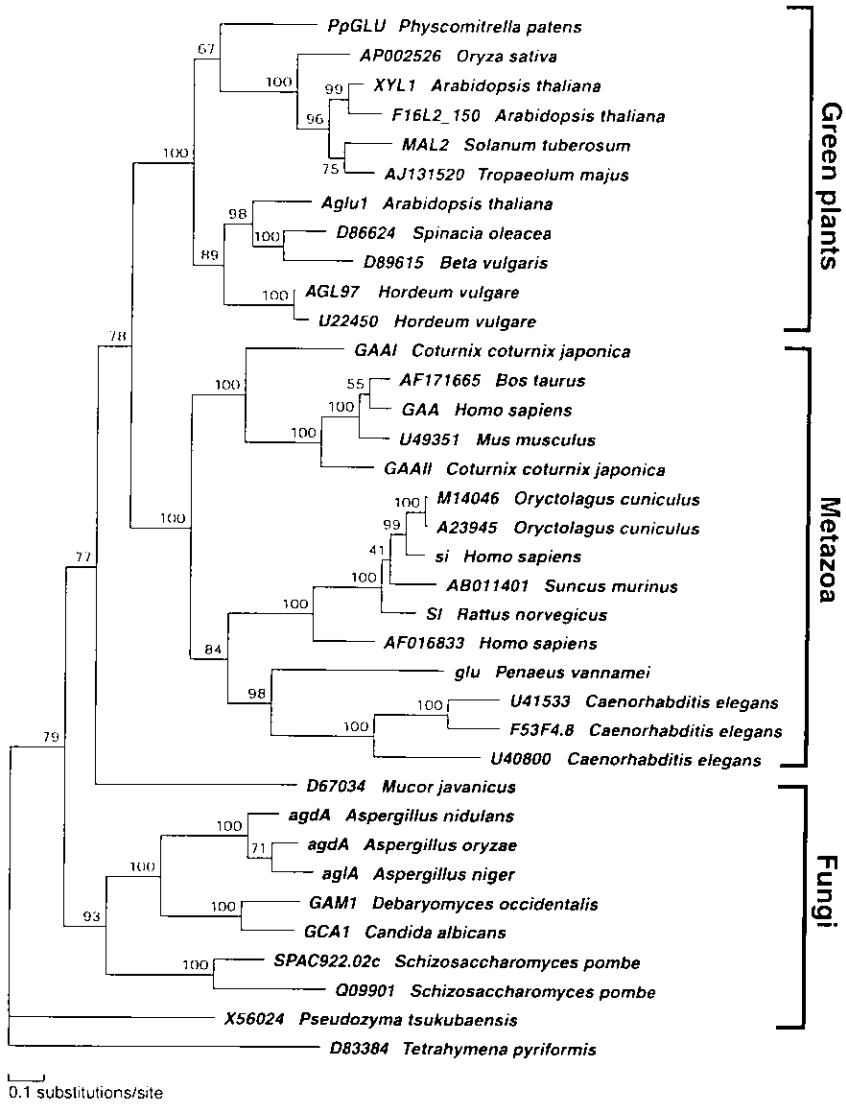


Figure 8. Phylogenetic analysis of *PpGLU*.
Maximum-Likelihood tree of 36 genes, rooted by the *Tetrahymena* gene. Local bootstrap values are indicated on or below branches. Horizontal branch length is proportional to the estimated evolutionary distance. Accession numbers of genes are *PpGLU* (AB057452), *XYL1* (AF087483), F16L21_50 (AL162459), *MAL2* (AJ277244), *Aglu1* (AF014806), *AGL97* (AF118226), *GAAI* (AB000967), *GAA* (Y00839), *GAAII* (AB006754), *SI* (L25926), *si* (X63597), *glu* (AJ250828), F53F4.8 (Z77663), *agdA* of *Aspergillus nidulans* (AF208225), *agdA* of *Aspergillus oryzae* (D45179), *aglA* (D45356), *GAM1* (M60207), *GCA1* (AF082188), SAPC9922.02c (Q09901).

Discussion

Large-scale generation of trap lines by the homologous integration method

Gene-trap and enhancer-trap elements were constructed for both non-homologous and homologous integration methods (Figure 2). Transient GUS assays of enhancer-trap elements indicated that the core sequence of the CaMV 35S promoter has weaker promoter activity than the longer form and that the former version of the promoter is suitable for an enhancer-trap system in *P. patens*. Transformation efficiencies of non-homologous and homologous integration methods were compared; homologous integration methods generated approximately 10-fold more stable transformants than non-homologous methods. Schaefer (2001) reported that targeting of the *apt* gene mediated by homologous recombination events occurring within stretches of continuous homology as short as 53 and191 bp have been observed and stretches of homology ranging from 500-700 bp are sufficient to target the *apt* locus at high frequency (up to one clone per μg DNA). High transformation frequency of homologous integration may be dependant on homologous recombination events based on stretches of long homology (3-6 kb) of mini-transposon-inserted genomic DNA. The results indicate that homologous integration using shuttle mutagenesis is more suitable for generating large numbers of trap lines. Molecular analysis of one gene-trap line, YH209, revealed that the DNA fragment 209-8SC did not insert into its homologous site. In addition to the line YH209, three other gene-trap lines were preliminarily analyzed, of which all had tagged genomic DNA at non-homologous sites. This was unexpected, since the rate of homologous recombination is 10-fold higher than the non-homologous recombination rate (Schaefer and Zrÿd, 1997), and homologous DNA fragments are expected to insert in their homologous sites. Previous analysis of *Physcomitrella* mutant lines generated by shuttle mutagenesis suggested that arrays of inserted DNA fragments are formed in *Physcomitrella* cells (Nishiyama *et al.*, 2000). Extrachromosomal homologous

recombination may explain how the large fragments were formed and why the 209-8SC DNA fragment was inserted into non-homologous sites in the line YH209. Shuttle mutagenesis results in tagging of *Physcomitrella* genomic DNA fragments by mini-transposons. Typically, one DNA fragment is tagged by a single mini-transposon, and the insertion site is random (Nishiyama *et al.*, 2000). Thus, in a tagged genomic library, there are multiple fragments derived from the same *Physcomitrella* DNA fragment; however, each is tagged in a different position by the mini-transposon. When these differently tagged DNA fragments containing the same *Physcomitrella* DNA sequences are introduced into a protoplast, two of these fragments themselves may undergo homologous recombination to form a DNA fragment having two copies of the mini-transposon and a DNA fragment without any mini-transposon. Furthermore, DNA fragments containing different *Physcomitrella* genomic DNA sequences may be subject to homologous recombination based on a homology in the mini-transposon. Successive homologous recombination of these two kinds will generate a large DNA fragment consisting of an array of homologous and non-homologous *P. patens* genomic sequences and multiple copies of the mini-transposon. Finally, this large DNA fragment may be integrated into the *P. patens* genome by homologous recombination based on the homology of the terminal *P. patens* sequence of the large DNA fragment, but the internal part of the large DNA fragment may contain different *P. patens* DNA sequences from the terminal sequence. Formation of a large DNA fragment before integration into the genome by extrachromosomal ligation has been reported in rice (Kohli *et al.*, 1998). Homologous recombination is expected to be more frequent than ligation in *P. patens*. If extrachromosomal ligation occurred frequently, multiple insertions of the 209-8SC fragment should have been observed.

In addition to the high transformation rate achieved, integration in a non-homologous site provides the advantage of trapping a gene or enhancer without

disrupting its original site. A gene-trap element does not need to be integrated at the homologous site to drive expression of the reporter gene as long as regulatory elements are present in the tagged genomic DNA fragment, which necessary and sufficient to regulate expression of the reporter gene. This is likely to be important when using a moss system. *P. patens* has an autotrophic haploid generation that differs from the epiphytic seed plant haploid generation, and the majority of its life cycle is haploid. Thus, a greater number of lethal transformants may result from disruption of a gene or regulatory element responsible for cell viability. Integration by non-homologous recombination will lower the possibility of such deleterious effects.

GUS expression patterns of trap lines

In this study, 235 and 1073 lines of gene-trap and enhancer-trap lines were generated, respectively. GUS staining of these lines revealed that genes or enhancers with various expression patterns were trapped (Table 2 and Figure 3). It is noteworthy that axillary hairs were stained in 155 of 235 (66%) gene-trap lines and 896 of 1073 (84%) enhancer-trap lines, much higher percentages than for other cells or tissues (Table 2). Axillary hairs develop at the adaxial base of a leaf and consist of a few cells arranged linearly. These hairs are reported to produce mucilage in some mosses (Ligrone, 1986), although this has not been examined in *P. patens*. It is possible that a gene expressed in axillary hairs or its enhancer element, harbors a hot spot for the integration of trap elements. Alternatively, it is possible that many genes are expressed in axillary hairs compared to other cells and tissues. Nishiyama *et al.*, (2000) reported that mini-transposon-tagged DNA fragments insert into different sites, supporting the latter explanation. However, the alternate possibility cannot be formally excluded. Further analyses of the expression patterns of *P. patens* genes and analyses of trap lines will provide insights into this issue.

Identification of a trapped gene, PpGLU

As it was demonstrated in the analysis of the line YH209, a trapped gene can be identified efficiently by 5' RACE, using *uidA*-specific primers, despite the presence of multiple insertions in a trap line. The isolation of the trapped gene should be more difficult, if the reporter gene was expressed at all insertion sites. The isolated trapped gene, *PpGLU* encodes a member of the glycoside hydrolase family 31 (Frandsen and Svensson, 1998). In the tree, *PpGLU* clusters with flowering plant acidic alpha-glucosidase genes (Figure 8). In plant, alpha-glucosidases have been classified into two classes based on the optimal pH for activity (Taylor *et al.*, 1998). One is neutral alpha-glucosidase (pH optimum 6-7), which was reported to be involved in starch degradation as a chloroplastic emzyme (Sun *et al.*, 1995) or complex control of glycoprotein folding and oligomer assembly in endoplasmic reticulum (Taylor *et al.*, 2000), and another is acid alpha-glucosidase (pH optimum 4-5). Two different functions have been assigned to plant acidic alpha-glucosidases. Acidic alpha-glucosidases of barley (U22450, Tibbot *et al.*, 1998; *AGL97*, Frandsen *et al.*, 2000) and spinach (D86624, Sugimoto *et al.*, 1995) are involved in the degradation of seed starch. Additionally, acidic alpha-glucosidase activity has been found in fractions that include cell walls (Klis, 1971; Parr and Edelman, 1975; Yamasaki and Konno, 1987, 1992; Beers *et al.*, 1990; Monroe *et al.*, 1999), although the respective genes have not been cloned. Based on their cell wall-specific localization, Taylor *et al.* (2000) speculated that acidic alpha-glucosidases are involved in xyloglucan oligosaccharide metabolism within cell walls, although possible substrates were not specified. Rhizoids contain reduced chloroplasts, and their starch granules are not visible. Thus, a rhizoid is not likely to be a photosynthetic organ, suggesting that the protein encoded by *PpGLU* is indeed involved in xyloglucan oligosaccharide metabolism, rather than starch degradation. *PpGLU* is expressed in

mature, pigmented rhizoid cells, including the rhizoid basal cell. However, the gene is not expressed in these cells before they mature. The metabolism of the pigment is unknown, and it is difficult to speculate on the relationship between pigmentation and *PpGLU* function.

Many genes have been isolated based on similarity to known genes, using PCR or hybridization techniques, and characterized in *P. patens* (Andreeva and Kutuzov, 1999; Girke *et al.*, 1998; Girod *et al.*, 1999; Hara *et al.*, 2001; Krogan and Ashton, 2000; Leech *et al.*, 1993; Sakakibara *et al.*, 2001; Strepp *et al.*, 1998). However, no genes have been identified based on their spatial expression patterns in *P. patens*. In addition to the example of the line YH209, two other genes specifically expressed in meristem tissue have been identified, one of which will be discussed in the next chapter. In conclusion, gene-trap and enhancer-trap systems established in this study should be useful to isolate genes with specific expression patterns.

Summary

Owing to its simple body plan and ease of gene knockout and allele replacement, the moss *Physcomitrella patens* is often used as a model system for studies in plant physiology and developmental biology. Gene-trap and enhancer-trap systems are useful techniques for cloning genes and enhancers that function in specific tissues or cells. Additionally, these systems are convenient for obtaining molecular markers specific for certain developmental processes. Elements for gene-trap and enhancer-trap systems were constructed using the *uidA* reporter gene with either a splice acceptor or a minimal promoter, respectively. Through a high rate of transformation conferred by a method utilizing homologous recombination, 235 gene-trap and 1073 enhancer-trap lines were obtained from 5637 and 3726 transgenic lines, respectively. Expression patterns of these trap lines in the moss gametophyte varied. The candidate gene trapped in a gene-trap line YH209, which shows rhizoid-specific expression, was obtained by 5' and 3' RACE. This gene was named *PpGLU*, and forms a clade with plant acidic alpha-glucosidase genes. Thus, these gene-trap and enhancer-trap systems should prove useful to identify tissue- and cell-specific genes in *Physcomitrella*.

CHAPTER 2. IDENTIFICATION OF THE *API1* GENE ENCODING A KINESIN-LIKE PROTEIN, PREDOMINANTLY EXPRESSED IN APICAL CELLS OF THE MOSS *PHYSCOMITRELLA PATENS*

Introduction

Plant body is formed by the continuous activity of small cell clusters. In angiosperms, the cell clusters locate at the shoot and the root meristem. Cells produced by these meristems serve two purposes: the formation of organs and the self-perpetuation of the meristem (reviewed in Steeves and Sussex, 1989). The shoot meristem arises during embryogenesis. During development of *Arabidopsis thaliana*, the shoot meristem is highly organized in the zones (e.g., the central zone and the peripheral zone) and gives rise to a stem, leaves and axillary meristems in a repetitive indeterminate fashion while being maintained by self-renewal. Recently, the roles of several genes in the shoot meristem development were investigated, giving us ideas for meristem formation (reviewed in Clark, 2001; Fletcher and Meyerowitz, 2000; Lenhard and Laux,1999).

Mosses have meristematic cells that are termed apical cells. During the course of moss development, an apical cell produces a multicellular protonemata and gametophores with different specialized cells. The gametophore apical cell differentiates from a caluonemal side branch initial cell during bud formation and gives rise to a stem and leaves in a repetitive indeterminate manner, while the apical cell is maintained by self-renewal.

In the moss, *Physcomitrella patens*, the gene-targeting technique based on homologous recombination enables to carry out the functional analysis of a gene (Scheafer, 2001). The simplicity of apical cell differentiation makes the study of the establishment and maintenance of the meristem more feasible. Thus, apical cell

differentiation of a gametophore is an excellent model for the plant meristem development, which displays three-dimensional apical growth. Comparisons of developmental mechanism between *P. patens* and *A. thaliana* will provide new insights into the understanding of the meristem development in land plants.

For the first step to analyze the molecular mechanism of apical cell differentiation of a gametophore, a gene was isolated, which was expressed in an apical cell of a bud using gene-trap and enhancer-trap systems. In this chapter, it is demonstrated that the isolated gene, *API1*, encoding a novel kinesin, is predominantly expressed around the apical cell of a bud as well as of a caulonemal and rhizoid cell. The *API1* gene is useful for the analysis of apical cell differentiation as a molecular marker of the apical cell.

Material and Method

Culture conditions and transformation

The wild-type strain of *Physcomitrella patens* subsp. *patens* (Ashton and Cove, 1977) was used. Culture conditions and transformation were as described in the culture and transformation in the chapter 1.

RACE

Procedures for 5' RACE of YH229 gene-trap line, *Apicar1*, was performed as described in the RACE section in the chapter 1. For 3' RACE to obtain the 3' region of *API1* cDNA, cDNA was synthesized from 2 μg of total RNA using SuperScriptII reverse transcriptase (GIBCO-BRL) with the adaptor primer. The PCR was performed using this cDNA as template with the UAP primer (5'-CUACUACUACUAGGCCACGCGTCGACTAGTAC-3') and the 229-2p1 primer (5'-CTGGTGAGGACGAGGATTCC-3'). PCR cycle conditions consisted of 3 min at 94°C, 30 cycles of 94°C for 30 sec, 58°C for 30 sec, and 72°C for 3 min. Additional PCR was carried out in the same condition as of the above PCR cycle with the UAP primer and the 229-2p2 primer (5'-ATGTCGTCTTGCTTATGCAG-3'). PCR conditions were the same as in the previous PCR. The amplified PCR fragment was isolated from an agarose gel using GENECLEAN III (BIO101) and cloned into pGEM-T (Promega). This plasmid was named p229-4-2.

DNA and RNA extractions

Procedures for nucleic acids extraction were described in the DNA and RNA extractions section in the chapter 1.

Southern and Northern analyses

Southern analysis was performed according to the procedure described in the Southern and Northern analyses section in the chapter 1 with slight modification. Under high-stringency hybridization, the blots were hybridized and washed at 65°C. On the other hand, the blots were hybridized at 55°C and washed at 53°C under low-stringency hybridization. A 229-2 fragment was PCR-amplified with the M13-21 primer (5'-TGTAAAACGACGGCCAGT-3') and the AIDA primer (5'-GGACCTGCATATAACCTGCATATAAC-3') using 229-2 clone as a template. The amplified fragment was labeled and used for colon screening. The 1013 bp *Hin*cII-*Bst*EII fragment and the 680 bp *Bam*HI fragment were excised from the 229-9SC fragment, labeled and then used for Southern analysis of *Apicar1* and 229-9SC-integrated transgenic lines, respectively. For a probe (termed N-terminal probe) corresponding to N terminal region of the API1 protein, the 1044 bp *Nco*I-*Sac*II fragment was excised from p229-2-4, while for a probe (termed C-terminal probe) corresponding to its C terminal region, the 1722 bp *Nco*I fragment was excised from the same plasmid. Each fragment was labeled and used for Southern analysis of wild type.

The method of Northern analysis is described in the same section as described above in the chapter 1, except that 0.7 μg of poly(A)$^+$ RNA was used. A fragment was PCR-amplified with p229-4-2 as template using KOD+ DNA polymerase (TOYOBO) with the api1F300SaI primer (5'-GAGCCGTCGACTGCGATGACAACGATG-3') and the api1R3940Cla primer (5'-AGAAATGTATCGATGCAGCGAG-3'), digested with *Sal*I and *Cla*I, and cloned in the *Sal*I and *Cla*I site of pBluescript KSII (+) (Stratagene). This fragment was used as a *API1*-coding probe for detection of *API1* transcripts in wild type. Northern blots were reprobed with *P. patens* glyceraldehyde 3-phosphate dehydrogenase gene (*GAPDH* probe) to control for loading (Leech *et al.*, 1993); the

*Pst*I-*Eco*RI fragment of the pPpGapC plasmid (provide by R. Kofuji) was excised, labeled, and hybridized as described in the chapter 1.

Amplification of genomic fragments coding the API1 *gene*

A genomic region covering the sequence from 5'-untranslated region to the stop codon of *API1* cDNA was divided into four regions, and the corresponding fragments were separately PCR-amplified. PCR was performed in 20 μl reaction containing 1 μl of genomic DNA in 1 x PCR buffer [10 mM Tris-HCl (pH8.3), 50 mM KCl, 1.5 mM $MgCl_2$], 1 unit of TAKARA ExTaq polymerase (TAKARA), 10 pmol of each primer, and 0.2 mM dNTP. PCR cycle conditions consisted of 5 min at 94°C, then 30 cycles of 94°C for 30 sec, 58°C for 30 sec, and 72°C for 3 min. PCR reactions were performed with the following primer sets: (1) the api1F300Sal (5'-GAGCCGTCGACTGCGATGACAACGATG-3') primer and the api1R1200 primer (5'-CCTGTCTGCGTCTTAAGTTCTC-3'), (2) the api1F1100 primer (5'-TGCATGGTCACGCTTGATG-3') and the api1R2130 primer (5'-GCTGCTGTTGGGACCAGG-3'), (3) the api1F2000 primer (5'-CCAGAAGCACATTAATGATCGC-3') and api1R3180 primer (5'-CCGCGAACTGTTGAGACC-3'), and (4) the api1F2820Sal primer (5'-GGACGCACAGTCGACTCACAGC-3') and the api1R3940Cla primer. Each amplified fragment was cloned into pGEM-T Easy (Promega) and then sequenced. The obtained sequences were assembled by AutoAssembler, ver.2.1.1 (ABI/Perkin-Elmer, Foster City, CA, USA).

To obtain the fragment corresponding to a genomic region after the stop codon of the *API1* gene, TAIL-PCR was performed (Liu *et al.*, 1995). TAIL-PCR cycles were performed using the GeneAmp9600 Thermal Cycler (ABI/Perkin-Elmer) and TAKARA

Ex Taq DNA polymerase (TAKARA). The gene-specific primers were designed to the sequence adjacent to the stop codon of the *API1* gene. First PCR was performed using the A1 primer (5'-NGTCGASWGANAWGAA-3') and the api1F3850 primer (5'-GGGAGCCAGATCCAGTAATGG-3') with 40 ng of genomic DNA as a template. PCR cycle conditions consisted of 1 min at 94°C, 1 min at 95°C, 5 cycles of 94°C for 1 min, 65°C for 1 min, and 72°C for 3 min, 1 cycle of 94°C for 30 sec, 30°C for 1 min, and 72°C for 3 min, 13 cycles of 94°C for 30 sec, 68°C for 1 min, 72°C for 3 min, 94°C for 30 sec, 68°C for 1 min, 72°C for 3 min, 94°C for 30 sec, 44°C for 1 min, and 72°C for 3 min, and 5 min at 72°C. Second PCR was performed using the A2 primer (5'-GTNCGASWCANAWGTT-3') and the api1F3895 primer (5'-GCCTTGCACACATTGTCAACTC-3') with a 50-fold dilution of the 1st PCR products. PCR conditions consisted of 1 min at 94°C, 1 min at 95°C, 10 cycles of 94°C for 30 sec, 68°C for 1 min, 72°C for 3 min, 94°C for 30 sec, 68°C for 1 min, 72°C for 3 min, 94°C for 30 sec, 44°C for 1 min, and 72°C for 3 min, and 5 min at 72°C. Third PCR was performed as in the 2nd PCR, using the A3 primer (5'-WGTGNAGWANCANAGA-3') and the api1F3950 primer (5'- GAGCTAGTTCTATTACGCCGACTTG-3') with a 50-fold dilution of 2nd PCR products. The PCR products were separated on agarose gel and the fragment of 1.1 kb was cloned into pGEM-T (Promega) and sequenced.

GUS assay

GUS staining was performed as described in the GUS assay section in the chapter 1, except for that tissues of *Apciar1* and 229-9SC-integrated transformants were incubated in a substrate solution for 3 h.

Phylogenetic analysis

Phylogenetic analysis was performed as described in the phylogenetic analysis section in the chapter 1.

Results

Screening of gene-trap and enhancer-trap line with GUS activity in apical cells

To search genes associated with apical cell differentiation during bud formation, gene-trap and enhancer-trap lines predominantly showing GUS activity at buds, were screened from a large pool of gene-trap and enhancer-trap lines, which were generated by both the non-homologous integration method and the homologous integration method (see the chapter 1). Three gene-trap lines (YH78, YH229, and YH727) and four enhancer-trap lines (ET21, ET77, ET63, and ET246) were isolated from 5637 gene-trap lines and 3726 enhancer-trap lines, respectively (Figure 9). The ET77 line was generated by the non-homologous recombination method, and the others were obtained by the homologous recombination method. Of them, the YH229 line showed one of the strongest GUS activities around an apical cell in a bud. Therefore, the YH229 line was selected for further analysis.

GUS expression pattern of YH229 line

During bud formation in the YH229 line, GUS activity was first detected at a side branch initial cell of a caulonemal cell (Figure 10A). A side branch initial cell obliquely divides with respect to its long axis and produces a small, densely cytoplasmic apical cell and a large, highly vacuolated subapical cell. GUS activity was detected in both an apical and a subapical cell (Figure 10B). Subsequent cell division of an apical cell gives rise to a tetrahedral apical cell that continues to divide in three planes to form a bud. At this stage, GUS activity was predominantly detected in an apical cell and its surrounding cells (Figures 10C and 10D). GUS activity in an apical cell and its surrounding cells was detected in the various stage of a gametophore

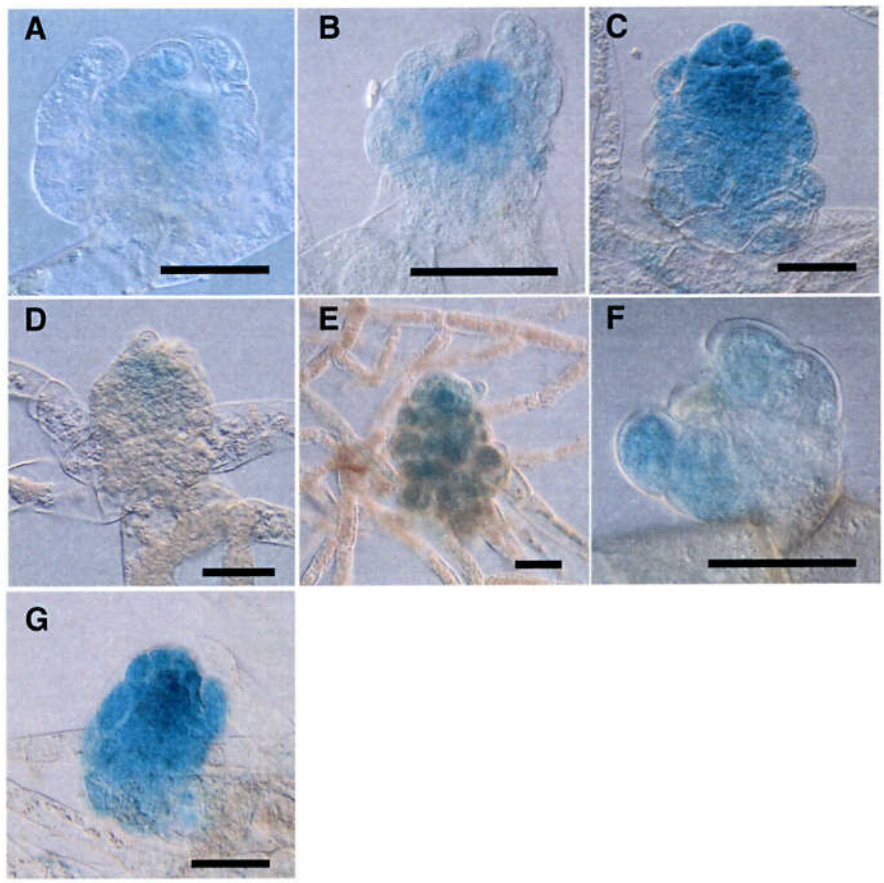


Figure 9. GUS activity in a bud of gene-trap and enhancer-trap lines.
(A) A gene-trap line, YH78. (B) A gene-trap line, YH229. (C) A gene-trap line, YH727. (D) An enhancer-trap line, ET21. (E) An enhancer-trap line, ET63. (F) An enhancer-trap line, ET77. (G) An enhancer-trap line, ET246. Bars = 50 μm.

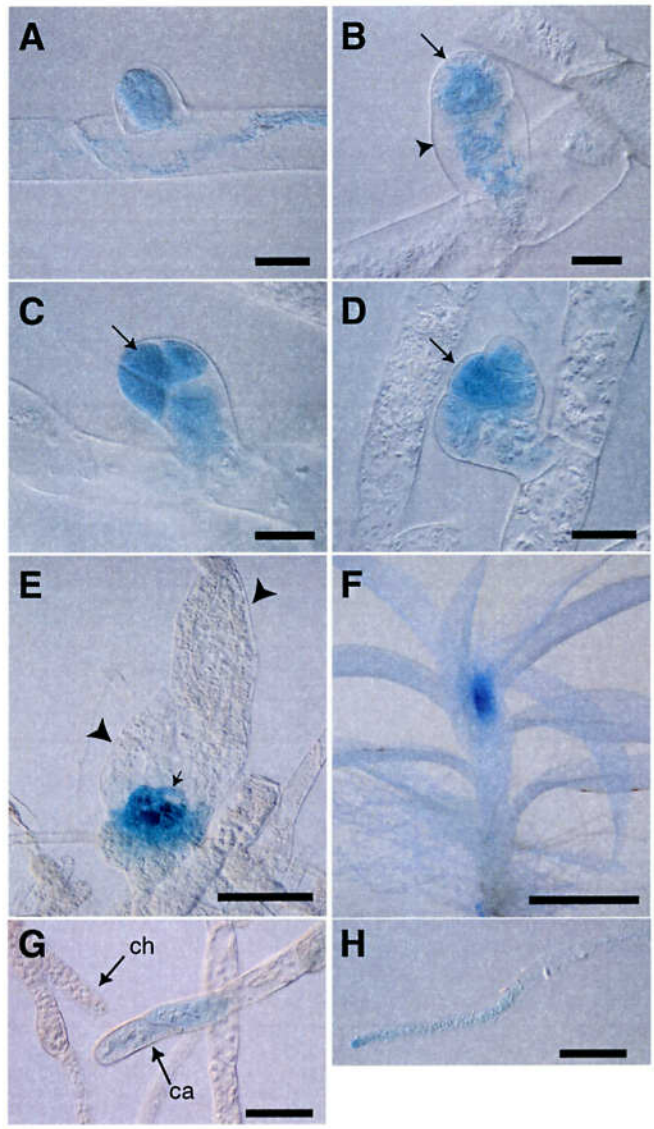


Figure 10. The GUS activity pattern of a gene-trap line, *Apicar1*.
(A) A side branch initial on a second subapical cell. The GUS activity was detected in a side branch initial cell as well as in a second subapical cell. (B) to (D) Bud at the different developing stages. Arrows indicate an apical cell. An arrowhead represents a subapical cell. (E) A gemetophore with two leaves (arrowheads). An arrow indicates an axillary hair. (F) A gametophore with several leaves. (G) Protonemal cells. Note that GUS enzyme activity is detected in caulonemal cells (ca), but not in chloronemal cells (ch). (H) Rhizoid cells. The bar in (A), (B), and (D) = 20 μm, in (E), (H), and (G) = 50 μm, in (F) = 0.5 mm.

(Figures 10E and 10F). Axillary hairs, which are filamentous cells at the base of leaves, showed GUS activity (Figure 10E). In protonemata, two cell-types exist; chloronemal cells and caulonemal cells. Chloronemal cells have large chloroplasts and a cross wall between adjacent cells that are perpendicular to the filament axis. Caulonemal cells have spindle-shaped chloroplasts and the cross wall that is oblique. GUS activity was predominantly detected in apical, subapical and second subapical cells of caulonemal cells (Figures 10A and 10G) but not chloronemal cells (Figure 10G). The apical cells exhibited much higher GUS activity than the subapical and second subapical cells did. Rhizoid cells are filamentous cells that are produced from the basal part of a gametophore and contain only few chloroplasts and possess oblique cross walls. The YH229 line predominantly showed GUS activity in apical and subapical cells of rhizoid cells in the same manner as detected in caulonemal cells (Figure 10H). Thus the YH229 line was designated *Apicar1* for exhibiting GUS activity in an apical part of gametophores, caulonemata and rhizoids 1.

***Isolation of the trapped gene in* Apicar1**

Genomic DNA isolated from the *Apicar1* line was digested with *Bgl*II, *Eco*RI, *Hin*dIII and *Pst*I, and hybridized with the *uidA*-specific probe. A PCR-amplified fragment of the mini-transposon *uidA* gene, from start to stop codons, was used as a *uidA*-specific probe. Multiple fragments were detected in each lane (Figure 11A), indicating that *Apicar1* contains multiple copies of the HI-GT mini-transposon. By northern hybridization of *Apicar1* poly(A)$^+$ RNA using the *uidA*-specific probe, several band of 5.1, 3.5, 2.4, 0.8, and 0.4 kb were detected (Figure 11B). The size of *uidA* transcripts is approximately 2.0 kb long, indicating that there are at least three *P. patens* chromosomal genes, a part of which was fused to the *uidA* gene.

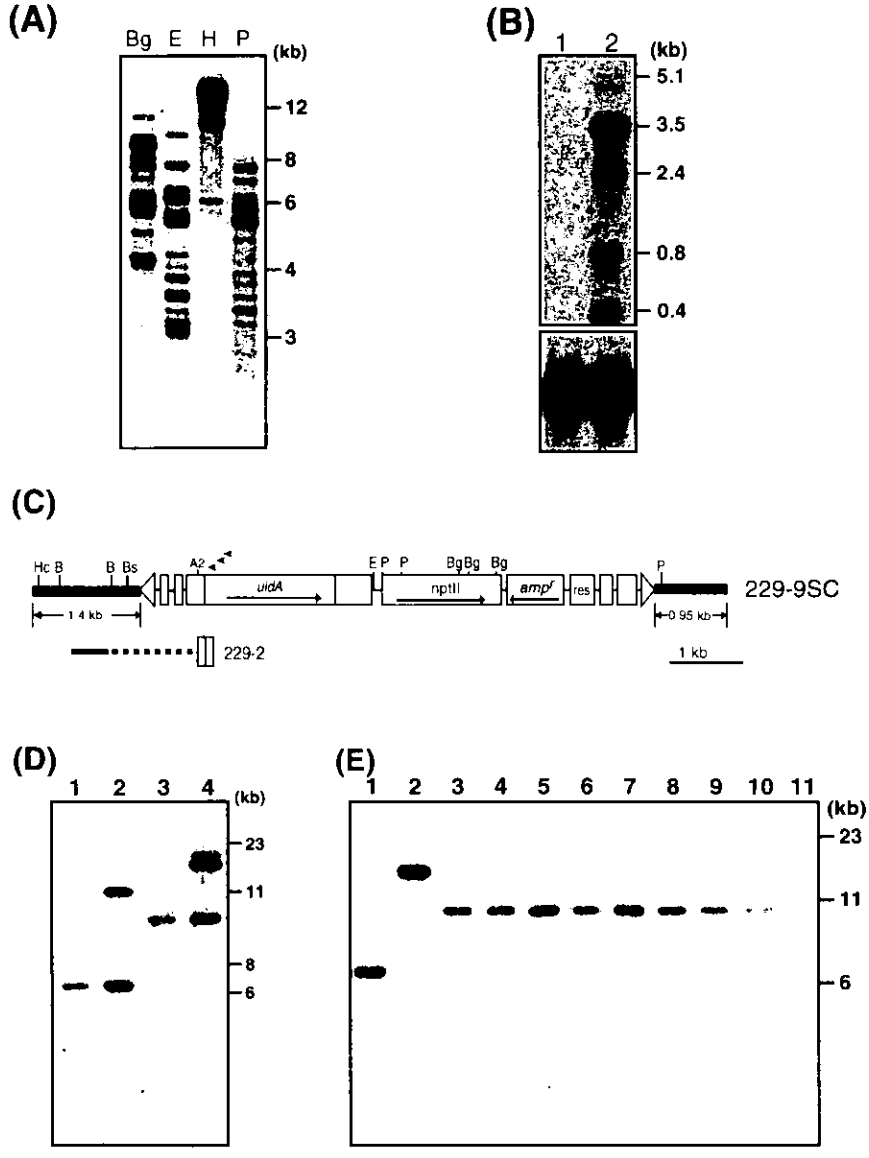


Figure 11. Molecular analysis of *Apicar1*, generated by the HI-GT mini-transposon

Figure 11. (continued)
(A) Genomic Southern analysis of *Apicar1*. The blot of genomic DNA digested with *Bgl*II (Bg), *Eco*RI (E), *Hin*dIII (H), and *Pst*I (P) was probed with *uidA* gene-specific probe. (B) Northern analysis of wild type (lane 1) and YH209 (lane 2). Blots were probed with a *uid*A gene-specific probe (upper lanel) and glyceraldehyde 3-phosphate dehydrogenase gene (*GAPDH* probe) gene as a control (lower lanel). (C) Schematic representation of a transposon-tagged genomic fragment (229-9SC). Locations of the *uid*A gene-specific primers used for 5' RACE are indicated with arrowheads. A2 indicates the second acceptor site of the *GPA1* intron. *Bam*HI (B), *Bgl*II (Bg), *Bst*HII (Bs), *Eco*RI (E), *Hin*cII (Hc), and *Pst*I (P) sites are shown. (D) Genomic Southern analysis of wild type and *Apicar1*. Genomic DNA from wild type (lanes 1 and 3) and *Apicar1* (lanes 2 and 4) was digested with *Bgl*II (lanes 1 and 2) or *Hin*dIII (lanes 3 and 4). The 1013 bp *Hin*cII-*Bst*EII DNA fragment of the 229-9SC was used as a probe. (E) Southern analysis of wild type (lane 1) and stable transformants (lanes 2-11) with 229-9SC fragment. Genomic DNA was digested with *Bgl*II and probed with the 680 bp *Bam*HI DNA fragment of the 229-9SC. Lane 3 and 5 correspond to line 0208-37 (shown in Figures 12A, 12B, 12C and 12D) and 0208-68 (Figure 12E), respectively.

Using *uidA* gene-specific primers and cDNA derived from gametophyte poly(A)$^+$ RNA from *Apicar1*, 5' RACE was performed and 52 DNA fragments were obtained. These clones were sequenced and sorted into nine groups. A 229-2 clone was first analyzed since the clone represents a largest group containing 25 clones out of 52 clones. The 229-2 clone contains 603 bp fragment in length and contained 527 bp of *P. patens* DNA adjacent to the 5' end of a partial HI-GT mini-transposon that begins with the second acceptor site of the *GPA1* intron. The 229-2 fragment was PCR-amplified and used as a probe for colony hybridization screening of the tagged genomic library used for *P. patens* transformation. Sequencing of a positive clone, an 8.35 kb mini-transposon-tagged fragment named 229-9SC, revealed that HI-GT separated a 2.35 kb *P. patens* genomic DNA fragment into 1.4 and 0.95 kb fragments (Figure 11C). The *P. patens* DNA sequence of the 229-2 clone was included in the 1.4 kb fragment (Figure 11C).

To see whether the 229-9SC fragment was integrated into homologous or non-homologous sites in *Apicar1*, *P. patens* genomic DNA digested with *Bgl*II or *Hin*dIII was hybridized with the 1013 bp *Hin*cII/*Bst*EII fragment of 229-9SC. A single band was detected in wild-type genomic DNA digested with either enzyme, while one extra band was detected in digests of *Apicar1* genomic DNA (Figure 11D). Thus, it was concluded that 229-9SC was integrated into a non-homologous site.

To determine whether the GUS activity of *Apicar1* reflects the endogenous expression pattern of a gene encoded on 229-9SC, the 229-9SC fragment was introduced to *P. patens* by homologous recombination. Genomic DNA of ten stable transformants was extracted, digested with *Bgl*II, and hybridized with the 680 bp *Bam*HI fragment of 229-9SC. An expected single band of approximately 10 kb in length was detected in nine out of ten lines, indicating that 229-9SC is singly integrated in a homologous site in these nine lines (Figure 11E). GUS activity in these nine lines

was examined (Figure12). GUS activity was detected in apical cells of caulonemata (Figure 12A) and rhizoids (Figure 12B) but not in chloronemal apical cells (Figure 12C) and apical cells of buds and gametophores (Figures 12D and 12E), indicating that the observed expression patterns were similar to those of *Apicar1*.

By a BLASTX search at NCBI against the nr dataset with the sequence of 1.4 kb fragment of 229-9SC, no similar sequences were found. However, a BLASTX search with the sequence of 0.95 kb fragment of 229-9SC indicated that the fragment contained the sequence similar to that of kinesin motor domain. To clone the 3' region extending from the 527 bp DNA fragment of the 229-2 clone, 3' RACE was performed using wild type cDNA as a template, and a 3981 bp fragment was obtained. The 3981 bp fragment, termed the 229-2-4 clone, contained a sequence of 44 bp which was identical to the 3' end of the 527 bp DNA fragment of 229-2 and poly(A)$^+$ tail. The 229-2-4 clone had a long open reading frame and was designated *API1* cDNA for *Apicar1*.

Sequence analysis of the API1 cDNA and the deduced amino acids

API1 cDNA contained an open reading frame of 3618 bp, and the predicted API1 protein consisted of 1206 amino acid with a calculated molecular mass of 132, 346 Da (Figure 13A). Genomic fragments corresponding to the *API1* gene was amplified by PCR with genomic DNA as template, and sequenced. Comparison of the nucleotide sequences of the genomic DNA and cDNA clones revealed that the coding sequence was separated by seven introns, ranging in size from 91 to 397 bp (Figure 13A). All of the introns contained GT and AG dinucleotides at their 5' and 3' ends, respectively, thereby confirming to the conserved GT-AG rule for 5' and 3' splice sites (Brown, 1986).

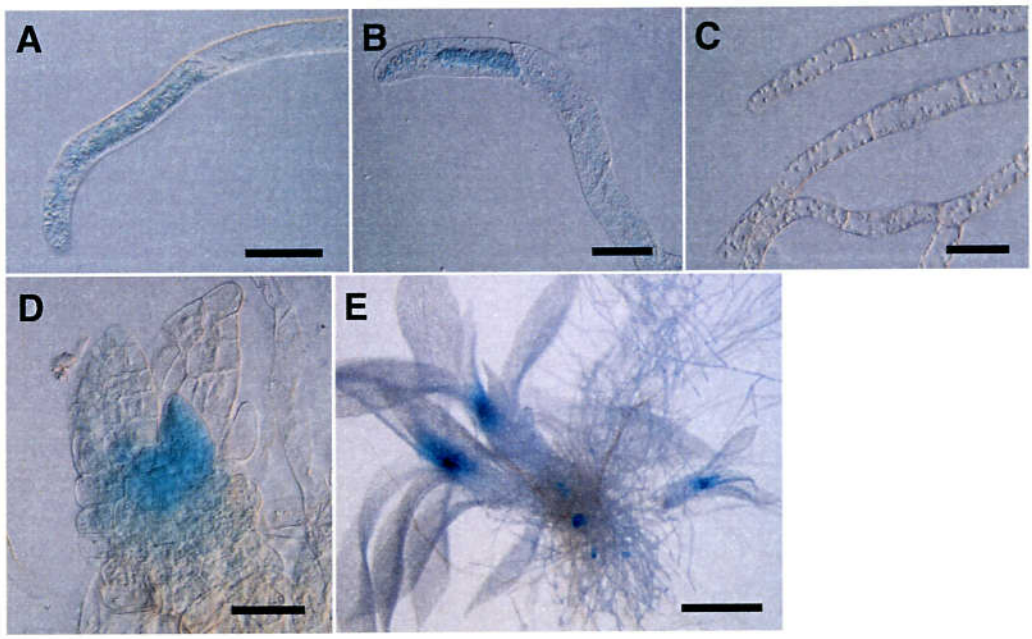
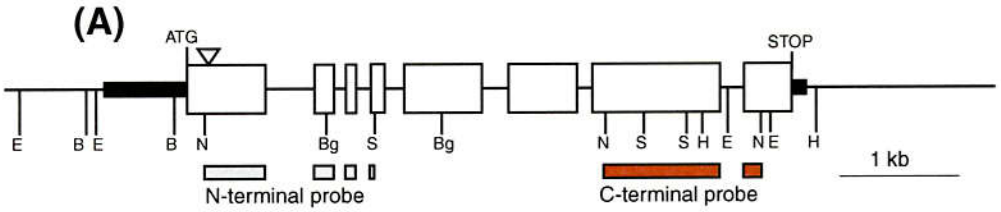


Figure 12 . GUS expression pattern of 229-9SC-integrated transformants.
(A) Caulonemal cells of the line 0208-37. (B) Rhizoid cells of the line 0208-37. (C) Chloronemal cells of the line 0208-37. (D) A gametophore with several leaves of the line 0208-37. (E) Protonemal cells and gametophores of the line 0208-68. Bars in (A), (B), (C) and (D) = 50 μm, and in (E) = 0.5 mm.



(B)

```
   1  MTTMMSPKPR EKRPSGNVGG NTPTGKARSM QPFLDRQASR TGSHPVEVVG RIREHPEGND
  61  KESAIRVLPH SRVAVRAEGM GNGCREFSLD GVSLAAMENL QAFYGRYVES RVEDVKAGGR
 121  CTIMMYGPTG AGKSYTMFGA AHEKGVAYHA LSQLMTRKAS DGGDDFSDKS IEVRATVWEI
 181  YNEEIYDLLA SVSAPKSGFG TLFKMSGSSS GRLEVMGKKV KTCMSISGTD PQKLLKEISK
 241  VEGRRVVKST NCNDRSSRSH CMVTLDVPDV GGKLALVDMA GSENVEQAGL GRELKTQTGK
 301  INQGNGALKR VVEAIAYGDS YIPYRDSKLT MLLQDSFEDD RAKILMILCA SPDLRDLHKT
 361  ICTLEYGAKA KCIVRLPNPP GKDQITATRA EQLQVLEARL TKKDAYIEAL RKQNELKSKE

 421  NEEREKELER KDRELADLKE KAKMWNSERV PQKKRKSLHD GEAEDLNSSP TRKRKIEELT
 481  SLVLQQQREI DLMRVRAEKA EAELLHLQRS YNLNNNARPG NTRGEXSDAK EGAGPESMDK
 541  RVPSFNQSPG ATNQRFQDIQ KHINDRIHAT QAECDRIKAQ FIPEQAVGER NGQRNPHSDL
 601  QAKPGPNSSM NDSFRHDRDS GSSSPSEVMS LTANSPSLLL LKNLVKEKER RLSLGTEEAT
 661  QSQSPSDAVN PSVQPRKDIE SEEWMGERLC SKGWLPIILE EDERIAQSLE SRKSVTPPQG
 721  SSAKDYSTLW SQVTPDNTIA SKEEAEGGDS IPSLTFPSPE EQFNTAMASY PSNSGLGGSP
 781  LFPWGDSRAK NDPSPIDNFA GVKLCTASSE FVFGVSKEWA LEEDQHDNRL AAPSLRTHSP
 841  LTALAEPAQN SLKNRADLDD PVSSMRSLDE SAGEEYSEVS PAETELLNLG TDLFDGGQSG
 901  GSDAEAARRA RIETIFMLCG ERREPARKTV SPFSKESEDT VRDVPSGRYS PTSSSYGSQQ
 961  FAENQLDQSI SDFSTSPSYS EAMYNTDVSV SSLPSLMPAV SPQRFPSSRR PTRRMSASIF
1021  ANENAQPKQE GEKTDVYVKW EKASDPADGK LICVVNVPKT SSLLELREEV EAHIPVSLRN
1081  FKFLFLGETG GAPVDREFES DVRVGSLPAS RENYGSRLAC LRPPSTESLF VAPRVESAPF
1141  RSLENQLGAS DEETTVAMGK LKGIPLTSHG RAGSPLGHAK GARSSNGETV RPKSGLAHIV
1201  NSTRCI
```

Figure 13. The *API1* gene structure and the predicted amino acid sequence.
(A) Schematic representation of the *API1* gene. White boxes and thick lines show the coding region and untranslated region, respectively. A triangle shows the integrated position of mini-transposon HI-GT. *Bam*HI (B), *Bgl*II (Bg), *Eco*RI (E), *Nco*I (N), and *Sac*I (S) are shown. Gray bars (N-terminal probe) and red bars (C-terminal probe) indicate the regions for use of Southern analysis. (B) Amino acid sequence deduced from *API1* cDNA. A box indicates kinesin motor motif. The solid arrows and the dashed arrows show putative coiled coil structure and putative nucleus localizaiton signals, respectively.

A BLASTX search with the nucleotide sequence of *API1* cDNA indicated that the API1 protein showed the similarity to members of kinesin superfamily proteins (Hirokawa, 1998). By a motif-search for the API1 protein by the SMART program (http://smart.embl-heidelberg.de/), the kinesin-motor motif was found in the N-terminal region of API1 (residues 43 to 384) and two putative coiled-coil domains were found in C-terminal region adjacent to motor domain (residues 396 to 450 and residues 473 to 514) (Figure 13B). The putative nuclear localization signals KKRK (residues 453 to 456) and RKRK (residues 472 to 475) were identified using the P-sort program (http://psort.nibb.ac.jp/index.html) (Figure 13B).

To compare the motor domain of API1 with other kinesin superfamily proteins, the corresponding amino acid sequence of API1 was aligned with four kinesin-like proteins that showed high similarity (E-value < $1e^{-24}$) with API1 (Figure 14). This region of API1 shares 57.9% amino acid sequence identity with the corresponding region of putative kinesin-like protein (B71450) of *Arabidopsis thaliana* (residues 46 to 409). The same region of API1 exhibits 30 to 32% sequence identity to the N-terminal region of bimC (residues 79 to 454) of *Aspergillus nidulans*, putative kinesin heavy chain (residues 96 to 458) of *Arabidopsis thaliana*, and cut7+ (residues 58 to 447) of *Schizosacchromyces pombe* (Figure 14). The API1 contained the consensus motif of motor domain, ExYxxxxxDL (residues 179 to 189 of API1), NxxSSRSH (residues 253 to 260 of API1), and DxAGSE (residues 277 to 283 of API1). The ATP-binding consensus motif (P-loop; [AG]-x(4)-G-K-[ST]) was highly conserved (residues 127 to 134 of API1) (Figure 14). On the C-terminal side of the motor domain, API1 protein contained six of eleven consensus amino acid residues normally found in the neck region of kinesin with N-terminal motor domain and an insertion of 4 amino acids in the region (Figure 14).

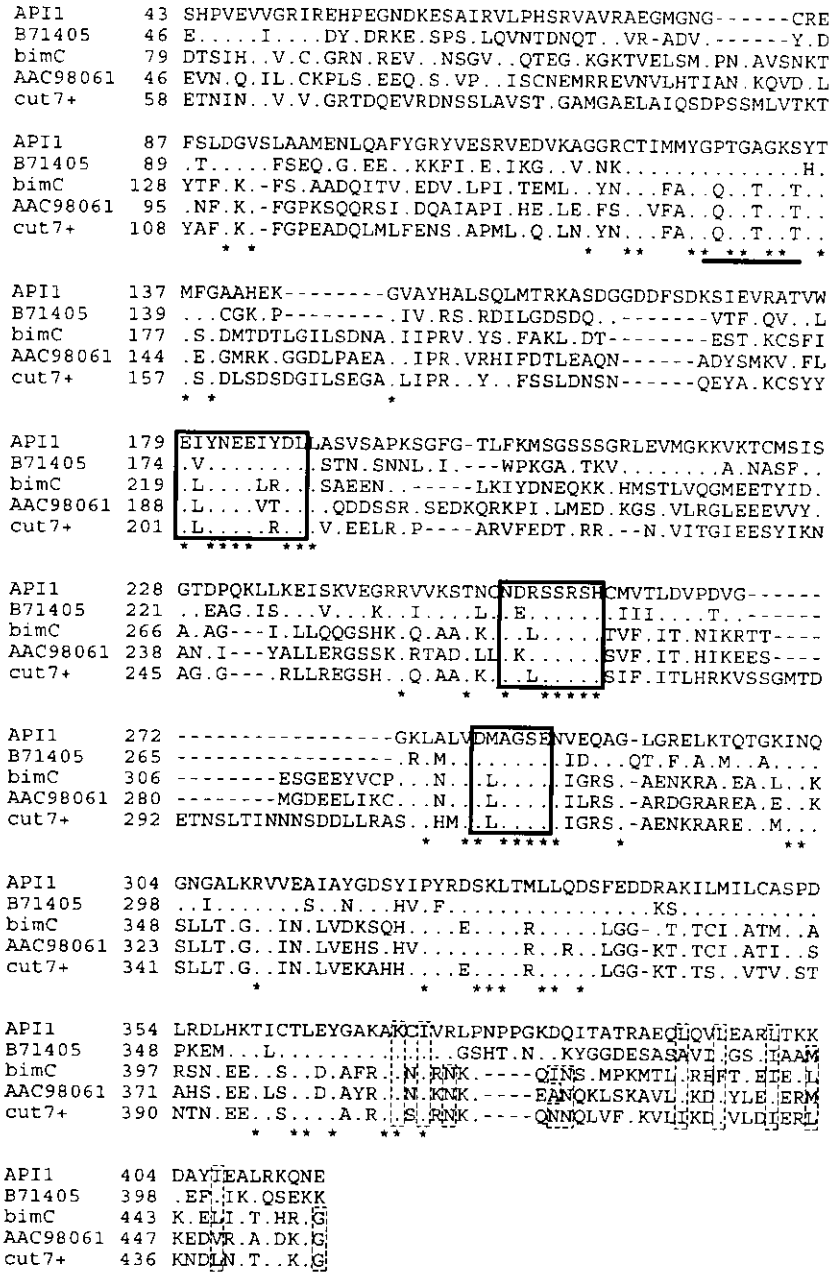


Figure 14. Comparison of the putative motor domain and neck region of the API1 protein with the corresponding domains of other kinesin superfamily proteins.
Dots and asterisks represent amino acids identical to the API1 sequence and conserved to the all sequences, respectively. Dashes represent gaps that have been introduced to improve the alignment. The underline indicates the putative ATP-binding motif. Thick boxes indicated conserved motifs in the motor domains. Dotted boxes represent conserved motifs in the neck region. B71405, putative kinesin-like protein of *Arabidposis thaliana*; bimC (AAA33298), kinesin-like protein of *Aspergillus nidulans*; AAC98061, putative kinesin heavy chain of *Arabidopsis thaliana;* cut7+ (CAA40738), putative kinesin of *Schizosaccharomyces pombe.*

Detection of API1 *transcripts and estimation of the copy number of* API1

To examine the size of *API1* transcripts, Northern analysis was performed with poly(A)$^+$ RNA from wild type. A fragment covering from start to stop codons of the *API1* gene was used as an *API1*-coding probe. The *API1* gene was transcribed into a single 4.5 kb mRNA in wild type (Figure 15A).

The copy number of the *API1* gene was examined by genomic Southern analysis. Genomic DNA extracted from wild type was digested by *Bgl*II, *Eco*RI, and *Hin*dIII, and blotted. Blots were probed using N-terminal probe or C-terminal probe (Figure 13A). Since the N-terminal probe corresponds to the sequence containing a conserved motor domain, this probe could hybridize with other member of kinesin superfamily along with the *API1* gene under the low-stringency conditions. Under the high-stringency conditions, N-terminal probe hybridized with a single fragment in each digest, and C-terminal probes did with a single fragment in the *Bgl*II digest and two fragments in both the *Eco*RI and the *Hin*dIII digests. Under the low-stringency conditions, C-terminal probe hybridized with only the same bands, as under the high stringency conditions, while N-terminal probe hybridized with an extra fragment (Figure 15B).

Phylogenetic analysis with the API1 *gene and kinesin superfamily genes*

To determine the relationship between the *API1* gene and kinesin superfamily genes, a phylogenetic analysis was performed. A BLAST search with kinesin-motor domain of API1 identified 454 genes with significant similarity (E-value < $1e^{-4}$), and the sequences were aligned with *API1*. After exclusion of short or redundant sequences, the phylogeny was inferred using 89 amino acid sites of 226 sequences by the maximum-likelihood method. After local rearrangements starting from the Neighbor-Joining tree, a tree with a log likelihood of -14722.16 ± 1280.63 was obtained as the Maximum-

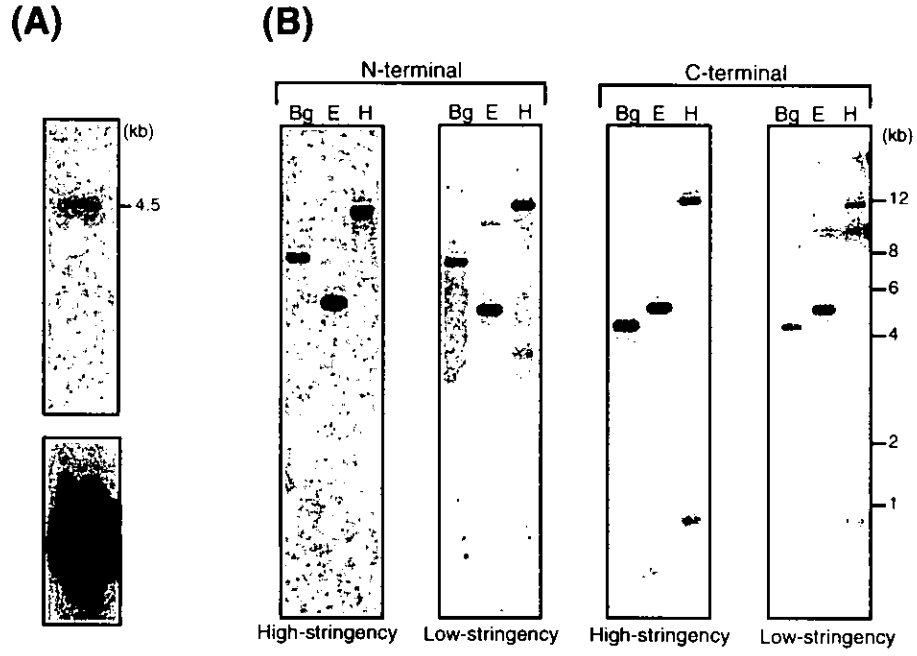


Figure 15. Northern and genomic Southern analyses of *API1*.
(A) The blot of poly(A)$^+$ RNA from wild type was hybridized with *API1*-coding probe (upper panel) or *GAPDH* probe as a contol (lower panel). (B) Genomic DNA digested with *Bgl*II (Bg), *Eco*RI (E), and *Hin*dIII (H) was hybridized with the N-terminal probe or C-terminal probe. Blots were washed in the high-stringency (High-stringency) or low-stringency (Low-stringency) condition .

Likelihood tree (Figure 16). In this tree, the *API1* gene formed a clade with a *Arabidopsis* putative kinesin-like gene (B71450) with 100% local bootstrap support. This clade was not included any gene family previously reported (Hirokawa, 1998; Moore and Endow, 1996). The close gene was resolved as a *Drosophila* kinesin-like gene (AAF57799), but separated with a large distance with a moderate local bootstrap probability of 82%.

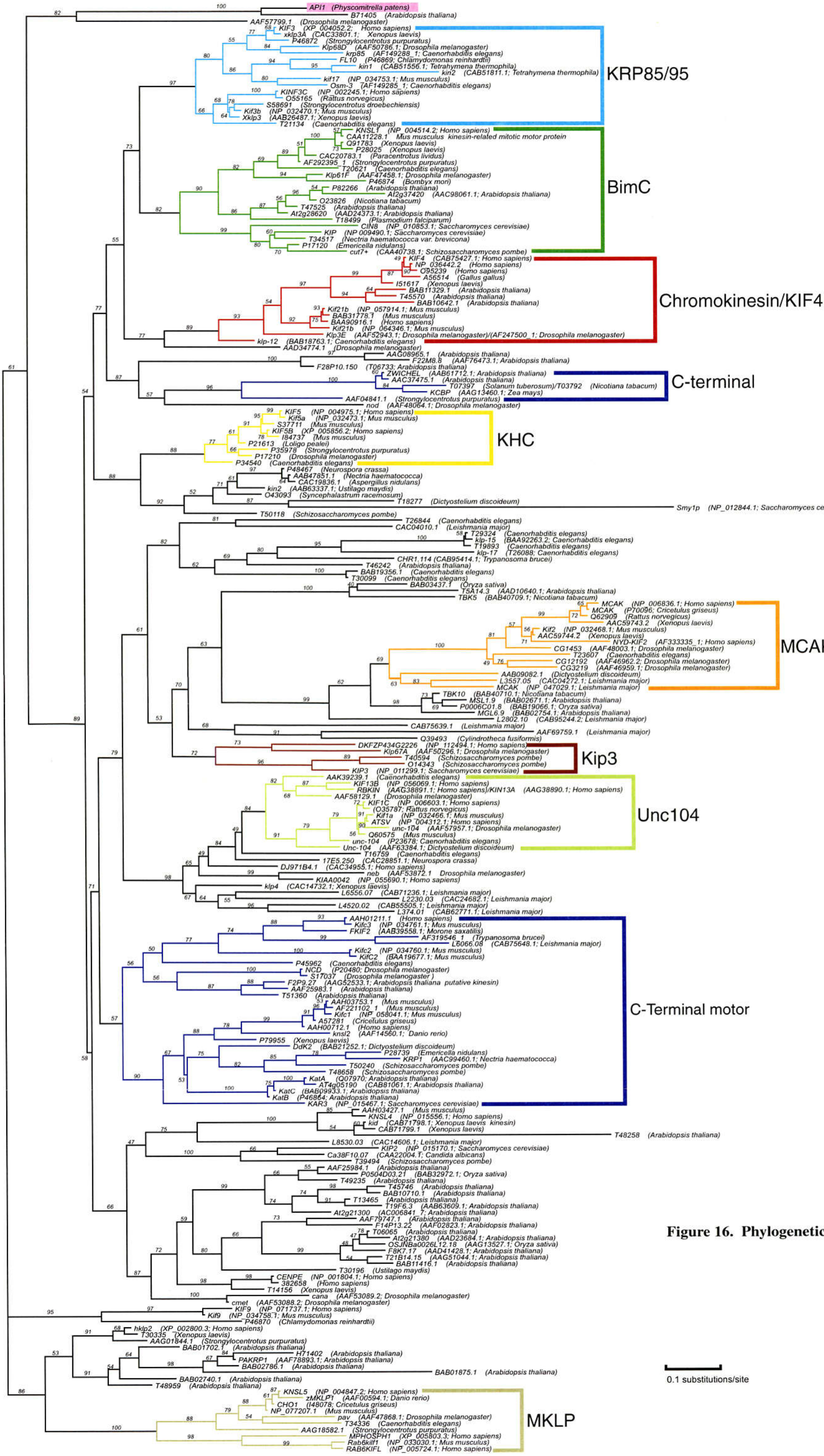


Figure 16. Phylogenetic analysis of *API1*.

Figure 16. (continued)
The maximum-likelihood tree of kinesin superfamily genes including *APII* obtained by local rearrangement search. This tree is unrooted tree. Local bootstrap values are shown on or below branches. Horizontal branch length is proportional to the estimated evolutionary distance. The nomenclature of each family follows by Moore and Endow (1996) and names of gene families are shown in the right side. *APII* is shown in a pink box. Genes with identical amino acid sequence within the region used for the phylogenetic analysis are described in parallel. Genes having no assigned name are represented by their accession numbers before parantheses. Accession numbers of genes with assigned names are shown in parentheses before the species name. Species are showed in parentheses.

Discussion

The API1 *gene from a gene-trap line Apciar1 encodes a novel kinesin-like protein*

Kinesin is a ATP-hydrolyzing motor protein that relies on microtubules as a track for movement, and mediates many cellular force-generating processes (Goldstein and Philp, 1999; Vale and Fletterick, 1997). Kinesin is characterized structurally by three distinct regions (Goldstein and Philp, 1999; Vale and Fletterick, 1997). The first region is a highly conserved head region with the motor domain, an approximately 350 amino acid sequence that contains the ATP and microtubule binding sites. The motor domain refers to the force-producing element of the protein, which is itself divided into two major parts: one part is the catalytic core and the other part termed the neck region is an adjacent ~40 amino acids found on either the N or C terminus of the catalytic core. The neck appears to work in concert with the catalytic core to produce movement and control the direction of movement. Beyond the motor domain, the second region is an alpha-helical coiled stalk region responsible for dimer formation. The third region is a non-conserved tail region that recognizes and binds the appropriate cargo. Kinesins and kinesin-like proteins showing similarity with the motor domain have been isolated from widely organisms, and constitute a kinesin superfamily (Hirokawa, 1998; Golstein and Philp, 1999). The predicted API1protein contains a putative motor domain that shares substantial sequence similarity with the motor domains of kinesin (Figure 14), which is followed by the putative coiled coil domains (Figure 13B). These features of the API1 protein indicate that the *API1* gene encodes a kinesin-like protein with N-terminal motor domain.

Multiple kinesin superfamily proteins are present in each organism (Hirokawa *et al.*, 1998; Miki *et al.*, 2001). These proteins were grouped into families based on their position of the motor domain and on sequence and structural similarity in the catalytic

motor and tail domains (Goodson *et al.*, 1994; Hirokawa, 1998; Moore and Endow, 1996). Eight to ten major families have been identified, and each group has its specific function, for example, movement of cargo along microtubules, zippering and sliding movement of microtubules, and depolymerizing or stabilizing microtubules (Goldstein and Philp, 1999; Moore and Endow, 1996). Phylogenetic analysis of the *API1* gene reveals that the *API1* gene forms a new family with one *Arabidopsis* gene encoding a kinesin-like protein that is deposited in the genomic sequence database, suggesting that these genes may have a novel function that is different from those of other kinesin-like proteins.

The API1 protein possesses structural features that distinguish it from kinesins with N-terminal motor domain, which are often called KIN N group (Vale and Fletterick, 1997). In KIN N group, its neck region of the motor domain locates on the C terminus of the motor. The neck region of KIN N group shows a consensus sequence of K/RxIxNxxxV/IN at the beginning of the its region. This neck region is followed by a hydorophobic repeat pattern of ϕ -xx (x) -ϕ -xxx-ϕ -xx- ϕ (ϕ represents a highly conserved, hydrophobic residue), which is then followed by a conserved glycine (Vale and Fletterick, 1997). Although a hydrophobic repeat pattern was found in the sequence of the API1 protein, the API1 protein has an incomplete K/RxIxNxxxV/IN motif and lack the conserved glycine (Figure 14), suggesting that the neck region of the API1 protein may have novel structure in this region.

The* API1 *gene is useful as a molecular marker for an apical cell

Through the analysis of a gene-trap line, *Apicar1*, the *API1* gene was identified, which is expressed predominantly in an apical cell in both protonema and gametophore. The restricted expression of the *API1* gene at the initial stage of the differentiation indicates that the *API1* gene is useful as a molecular marker for both establishment and

maintenance stage of the apical cell. The GUS expression patterns of *Apicar1* was consistent with those of the transgenic line with 229-9SC integrated into the homologous site, although the tagged genomic DNA fragment, 229-9SC was integrated into the non-homologous site of the genome in *Apicar1*. This suggests that 229-9SC contains the regulatory element, probably promoter located upstream of 5' coding region, which makes the *uidA* gene expressed in a cell-specific manner. Thus 229-9SC can be used for identification of regulatory elements responsible for apical cell-specific expression.

Kinesin superfamily proteins mediate various subcellular events that involve transport, including movement of vesicles or other organelles, membrane motility, chromosome segregation, microtubule dynamics, and intracellular signaling as a microtubule-based motor (Brady, 1995). The expression patterns of the *API1* gene suggest that its protein may be involved in the establishment and maintenance of meristematic cell through cellular force-generating processes. In plant cells, kinesins are shown to be involved in spindle formation (Vos *et al.*, 2000), phragmoplast organization (Asada *et al.*, 1997; Lee and Liu, 2000) or cell plate formation (Nishihama and Machida, 2000) during cell division. Thus, possibly, the function of API1 protein may be associated with those events during cell division in a novel microtubule-associated manner. Alternatively, the putative nuclear localization signals of the API1 protein suggest the API1 protein is likely to function in a nucleus. A kinesin that was isolated from chinese hamster and localized in the interphase nucleus was shown that it may participate in organization of chromosome structure through the process of moving along DNA or chromatin (Ohta *et al.*, 1996). However, the function of plant kinesin localized in the interphase nucleus has not yet been reported. Further characterization of the API1 protein, such as subcellular localization and gene disruption, is required to understand the function of the motor protein to apical cell differentiation. These

analysis of the *API* gene and *Arabidopsis* orthologoue will provide new insights into the establishment and maintenance of the apical meristem in land plants.

Summary

To investigate apical cell differentiation of a *Physcomitrella patens* gametophore, gene-trap and enhancer-trap lines were screened from a large collection of 9363 transgenic lines. Three gene-trap lines and four enhancer-trap lines, which preferentially expressed GUS activity in a bud, were isolated. One particular gene-trap line, *Apicar1*, which predominantly showed GUS activity in the apical cell of caulonemata, rhizoids, and gametophores, was characterized. *UidA*-fused fragments were amplified by the 5' RACE method using *uidA*-specific primers. One of the amplified fragments was used to screen the mini-transposon-tagged genomic DNA library that was used to generate the *P. patens* gene-trap lines. A genomic fragment, 229-9SC, which contained the sequence of the 5' RACE fragments, was then obtained. The 229-9SC genomic fragment was re-integrated into the *P. patens* genome by homologous recombination, confirming that the fragment-integrated transformants exhibited predominantly apical cell expression patterns observed in *Apicar1*. A full-length cDNA was isolated by the 3' RACE method. Sequence analysis of the isolated cDNA revealed that the trapped gene encoded a kinesin-like protein, named *API1* (*Apicar1*). According to a phylogenetic analysis of *API1* and kinesin superfamily genes, the *API1* gene formed a new family of kinesin-like proteins with one of the *Arabidopsis* kinesin-like genes. This suggests that *API1* may have a novel function that is different from those of kinesin-like proteins in other families. *API1* will be a useful molecular marker for studying the establishment and maintenance of the apical cell.

GENERAL DISCUSSION

As a first step toward understanding apical cell differentiation in the *Physcomitrella patens* gametophore, this study identified a gene expressed in apical cell differentiation, using gene-trap and enhancer-trap systems. Due to the simplicity of moss systems used to studying plant development, the gene-trap and enhancer-trap systems established in this study will allow the identification of specific development markers. These markers are useful in studies of developmental processes, including protonema differentiation regulated by phytohormones, and apical differentiation of gametophores.

Functional analysis of the *API1* gene provides an understanding of the developmental mechanisms of the apical cell. The gene-targeting technique, based on homologous recombination, is useful for determining the biological function of a gene (Schaefer, 2001). Since *API1* encodes a kinesin-like protein, the API1 protein is likely to be a microtubule-based motor protein that carries 'cargo' to particular destinations along the microtubule surface within cells (Goldstein and Philp, 1999; Vale and Fletterick, 1997). The intracellular placement of the API1 protein provides information for this cargo. Placement of the API1 protein might be determined by the insertion of the *GFP* gene just before the stop codon of the *API1* gene, using homologous recombination, and the generation of a transgenic line expressing the API1-GFP fusion protein. In addition to the functional analysis of the *API1* gene using the gene-targeting approach, it is necessary to characterize the biochemical properties of the API1 protein, for example, ATPase activity, and motor motility.

To fully understand apical cell differentiation, more molecular markers expressed in this differentiation must be identified. In addition to *Apicar1*, this study isolated two gene-trap lines and four enhancer-trap lines, which exhibit reporter gene expression in the apical cell of a bud. Using these trap lines, apical cell specific markers can be

identified. One gene-trap line, YH78, and one enhancer-trap line, ET21, are currently being characterized.

Two steps are necessary to attain a complete understanding of the establishment and maintenance of the meristem in plants. First, the mechanism involved in gametophore apical cell differentiation of *P. patens* must be clarified. Second, angiosperms must be investigated to determine whether the establishment and maintenance of the meristem in these organisms involves the *P. patens* mechanism. The complete genomic sequence of *Arabidopsis thaliana* makes it possible to find orthologues of genes isolated from different organisms (Lin *et al.*, 1999; Mayer *et al.*, 1999; Salanoubat *et al.*, 2000; Tabata *et al.*, 2000; Theologis *et al.*, 2000). Comparisons of developmental mechanisms in mosses and angiosperms will provide new insights into the general principles of meristem development in land plants.

ACKNOWLEDGEMENTS

I am deeply indebted to Professor Mitsuyasu Hasebe (National Institute for Basic Biology) for his generous support, advice and helpful discussion. I am grateful to Dr. Tomomichi Fujita (National Institute for Basic Biology) for helpful discussion and help on my thesis preparation. I also thank Dr. Takashi Murata (National Institute for Basic Biology) for helpful discussion.

I sincerely wish to thank Dr. Tomoaki Nishiyama (National Institute for Basic Biology) for collaboration in this work. I would like to thank Dr. Jo Ann Banks (Purdue Univ.) and Dr. Rumiko Kofuji (Kanasawa Univ.) for critical reading of this manuscript.

I would like to thank Dr. Robert Martienssen (Cold spring harbor lab.) for plasmid pWS32; Dr. Rumiko Kofuji for pPpGapC; Futamura Chemical Industries Co., Ltd. for providing the cellophane used to culture *P. patens*. I also thank Ms. Masae Umeda for moss transformation; Ms.Makiko Kondo, Ms.Chigusa Ono, and Ms.Yoshimi Bitoh for technical assistance. I thank all the past and present members of Dr. Hasebe's laboratory for giving me many help.

The computations were partly done on a SGI Origin 2000 in the Computer Lab of NIBB. The NIBB Center for Analytical Instruments provided sequence facilities.

LISTS OF REFERENCES

Adachi, J. and Hasegawa, M. (1996) MOLPHY version 2.3; programs for molecular phylogenetics based on maximum likelihood. *Comput. Sci. Monogr.* **28**, 1-150.

Andreeva, A.V. and Kutuzov, M.A. (1999) *Physcomitrella patens* gene/cDNA fragments related to genes encoding protein ser/thr phosphatases. *J. Plant Physiol.* **155**, 153-158.

Asada, T., Kuriyama, R. and Sibaoka, H. (1997) TKRP125, a kinesin-related protein involved in the centrosome-independent organization of the cytokinetic apparatus in tobacco BY-2. *J. Cell. Sci.* **110**, 179-189.

Ashton, N.W. and Cove, D.J. (1977) The isolation and preliminary characterization of auxotrophic and analogue resistant mutants in the moss *Physcomitrella patens*. *Mol. Gen. Genet.* **154**, 87-95.

Beck, E., Ludwig, G., Auerswald, E.A., Reiss, B. and Schaller, H. (1982) Nucleotide sequence and exact localization of the neomycin phosphotransferase gene from transposon Tn5. *Gene,* **19**, 327-336.

Beers, E.P., Duke, S.H. and Henson, C.A. (1990) Partial characterization a subcellular localization of three alpha-glucosidase isoforms pea (*Pisum sativum* L.) seedlings. *Plant Physiol.* **94**, 738-744.

Bellen, H. J. (1999) Ten years of enhancer detection: Lessons from the fly. *Plant Cell,* **11**, 2271-2281.

Benfey, P.N., Ren, L. and Chua, N.-H. (1989) The CaMV 35S enhancer contains at least two domains which can confer different developmental and tissue-specific expression patterns. *EMBO J.* **8**, 2195-2202.

Birks, H.J.B., Heegaard, E., Birks, H.H. and Jonsgard, B. (1998) Quantifying bryophyte-environment relationships. In: Bates, J.W., Ashton, N.M. and Duckeet, J.G. (eds), Bryology for the twenty-first century. Maney Publishing British Bryol. Society, Leeds, pp.305-320.

Boyd, P.J., Hall, J. and Cove, D.J. (1988) Somatic mutagenesis of the moss *Physcomitrella patens*. *Mol. Gen. Genet.* **211**, 545-546.

Brady, S.T. (1995) A kinesin medley: biochemical and functional heterogeneity. *Trends Cell Biol.* **5,** 159-164.

Brown, J.W. (1986) A catalogue of splice junction and putative branch point sequences from plant introns. *Nucleic Acid Res.* **32**, 531-535.

Burns, N., *et al.* (1994) Large-scale analysis of gene expression, protein localization, and gene disruption in *Saccharomyces cerevisiae*. *Genes Dev.* **8**, 1087-1105.

Cho, S.H., Chung, Y.S., Cho, S.K., Rim, Y.W. and Shin, J.S. (1999) Particle bombardment mediated transformation and GFP expression in the moss *Physcomitrella patens*. *Mols. Cell*, **9**, 14-19.

Church, G.M. and Gilbert, W. (1984) Genomic sequencing. *Proc. Natl. Acad. Sci. USA*, **81**, 1991-1995.

Clark, S.E. (2001) Meristems: start your signaling. *Curr. Opin. Plant Biol.* **4**, 28-32.

Cove, D.J., Knight, C.D. and Lamparter, T. (1997) Mosses as model systems. *Trends Plant Sci.* **2**, 99-105.

Cove, D.J., Schild, A., Ashton, N.W. and Hartmann, E. (1978) Genetic and physiological studies of the effect of light on the development of the moss *Physcomitrella patens*. *Photochem. Photobiol.* **27**, 249-254.

Fletcher, J.C. and Meyerowitz, E.M. (2000) Cell signaling within the shoot meristem. *Curr. Opin. Plant Biol.* **3**, 23-30.

Frandsen, T.P., Lok, F., Mirgorodskaya, E., Roepstorff, P. and Svensson, B. (2000) Purification, enzymatic characterization, and nucleotide sequence of a high-isoelectric-point alpha-glucosidase from barley malt. *Plant Physiol.* **123**, 275-286.

Frandsen, T.P. and Svensson, B. (1998) Plant alpha-glucosidases of the glycoside hydrolase family 31: molecular properties, substrate specificity, reaction mechanism, and comparison with family members of different origin. *Plant Mol. Biol.* **37**, 1-13.

Girke, T., Schmidt, H., Zähringer, U., Reski, R. and Heinz, E. (1998) Identification of a novel Δ6-acyl-group desaturase by targeted gene disruption in *Physcomitrella patens*. *Plant J.* **15**, 39-48.

Girod, P.-A., Fu, H., Zryd, J.-P. and Vierstra, R.D. (1999) Multiubiquitin chain binding subunit MCB1 (RPN10) of the 26S proteasome is essential for developmental

progression in *Physcomitrella patens*. *Plant Cell*, **11**, 1457-1472.

Goldstein, L.S.B. and Philp, A.V. (1999) The road less travelled: emerging principles of kinesin motor utilization. *Annu. Rev. Cell Dev. Biol.* **15**, 141-183.

Goodson, H.V., Kang, S.J. and Endow, S.A. (1994) Molecular phylogeny of the kinesin family of microtuble motor proteins. *J. Cell Sci.* **107**, 1875-1884.

Grimsley, N.H., Ashton, N.W. and Cove, D.J. (1977) Complementation analysis of auxotrophic mutants of the moss, *Physcomitrella patens*, using protoplast fusion. *Mol. Gen. Genet.* 155, 103-107.

Guerineau, F., Brooks, L., Meadows, J., Lucy, A., Robinson, C. and Mullineaux, P. (1990) Sulfonamide resistance gene for plant transformation. *Plant Mol. Biol.* **15**, 127-136.

Hara, K., Sugita, M. and Aoki, S. (2001) Cloning and characterization of the cDNA for a plastid sigma factor from the moss *Physcomitrella patens*. *Biochim. Biophys. Acta*, **1517**, 302-306.

Hasebe, M., Wen, C.-K., Kato, M. and Banks, J.A. (1998) Characterization of MADS homeotic genes in the fern *Ceratopteris richardii*. *Proc. Natl. Acad. Sci. USA*, **95**, 6222-6227.

Hasegawa, M and Kishino, H. (1994) Accuracies of the simple methods for estimating the bootstrap probability of a maximum-likelihood tree. *Mol. Biol. Evol.* **11**, 142-145.

Hirokawa, N. (1998) Kinesin and dynein superfamily proteins and the mechanism of organelle transport. *Science*, **279**, 519-26.

Hirokawa, N., Noda, Y. and Okada, Y. (1998) Kinesin and dynein superfamily proteins in organelle transport and cell division. *Curr. Opin. Cell Biol.* **10**, 60-73.

Jefferson, R.A., Kawanagh, T.A. and Bevan, M.W. (1987) GUS fusions: ß-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* 6, 3901-3907.

Jenkins, G.I. and Cove, D.J. (1983a) Light requirement for the regeneration of protoplasts of the moss *Physcomitrella patens*. *Planta*, **157**, 39-45.

Jenkins, G.I. and Cove, D.J. (1983b) Phototropism and polartropism of primary chloronemata of the moss *Physcomitrella patens*: responses of mutant strains. *Planta*, **159**, 432-438.

Jones, D.T., Taylor, W.R. and Thornton, J.M. (1992) The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **8**, 275-282.

Kammerer, W. and Cove, D.J. (1996) Genetic analysis of the effects of re-transformation of transgenic lines of the moss *Physcomitrella patens*. *Mol. Gen. Genet.* **250**, 380-382.

Kishino, H., Miyata, T. and Hasegawa, M. (1990) Maximum likelihood inference of protein phylogeny and the origin of chloroplasts. *J. Mol. Evol.* **30**, 151-160.

Klis, F.M. (1971) Alpha-glucosidase activity located at the cell surface in callus of *Convolvulus arvenisis*. *Physiol. Plant.* **25**, 253-257.

Knight, C.D., Cove, D.J., Boyd, P.J. and Ashton, N.W. (1988) The isolation of biochemical and developmental mutants in the moss *Physcomitrella patens*. In: Glime, J.M. (eds), Methods in bryology. Proc. Bryol. Meth. Workshop, Mainz. Hattori Botanical Laboratory, Tokyo, pp.47-58.

Kohli, A., Leech, M., Vain, P., Laurie, D. and Christou, P. (1998) Transgene organization in rice engineering through direct DNA transfer supports a two-phase integration mechanism mediated by the establishment of integration hot spots. *Proc. Natl. Acad. Sci. USA*, **95**, 7203-7208.

Krogan, N.T. and Ashton, N.W. (2000) Ancestry of plant MADS-box genes revealed by bryophyte (*Physcomitrella patens*) homologues. *New Phytol.* **147**, 945-950.

Lee, Y.-R.J. and Liu, B. (2000) Identification of a phragmoplast-associated kinesin-related protein in higher plants. *Cur. Biol.* **10**, 797-800.

Leech, M.L., Kammerer, W., Cove, D.L., Martin, C. and Wang, T.L. (1993) Expression of myb-related genes in the moss, *Physcomitrella patens*. *Plant J.* **3**, 51-61.

Lenhard, M. and Laux, T. (1999) Shoot meristem formation and maintenance. *Curr. Opin. Plant Biol.* **2**, 44-50.

Ligrone, R. (1986) Structure, development and cytochemistry of mucilage-secreting hairs in the moss *Timmiella barbuloides* (Brid.) Moenk. *Annal. Bot.* **58**, 859-568.

Lin, X. *et al.* (1999) Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. *Nature*, **402**, 761-768.

Liu, Y.G., Mitsukawa, N., Oosumi, T. and Whttier, R.F. (1995) Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal

asymmetric interlaced PCR. *Plant J.* **8**, 457-463.

Mayer, K. *et al.* (1999) Sequence and analysis of chromsome 4 of the plant *Arabidopsis thaliana*. *Nature*, **402**, 769-777.

McClelland, D. J. (1987) Protonemal branching patterns in *Physcomitrella patens*. *Br. Bryol. Soc. Bull.* **49**, 20-22.

Mengiste, T. and Paszkowski, J. (1999) Prospects for the precise engineering of plant genomes by homologous recombination. *Biol. Chem.* **380**, 749-758.

Miki, H., Setou, M., Kaneshiro, K. and Hirokawa, N. (2001) All kinesin superfamily protein, KIF, genes in mouse and human. *Proc. Natl. Acad. Sci. USA*, **98**, 7004-7011.

Monroe, J.D., Gough, C.M., Chandler, L.E., Loch, C.M., Ferrante, J.L. and Wright, P.W. (1999) Structure, properties and tissue localization of apoplastic alpha-glucosidase in crucifers. *Plant Physiol.* **119**, 385-397.

Moore, D.J. and Endow, S.A. (1996) Kinesin proteins: a phylum of motors for microtubule-based motility. *BioEssays*, **18**, 207-219:

Murray, M.G. and Thompson, W.F. (1980) Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* **8**, 4321-4325.

Nishihama, R. and Machida, Y. (2000) The MAP kinase cascade that includes MAPKKK-Related protein kinase NPK1 controls a mitotic process in plant cells. *Results Problem Cell Differ.* **27**, 119-130.

Nishiyama, T., Hiwatashi, Y., Sakakibara, K., Kato, M. and Hasebe, M. (2000) Tagged mutagenesis and gene-trap in the moss, *Physcomitrella patens* by shuttle mutagenesis. *DNA Res.* **7**, 9-17.

Odell, J.T., Nagy, F. and Chua, N.-H. (1985) Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. *Nature*, **313**, 810-812.

Ohta T, Kimble M, Essner R, Kofron M. and Kuriyama R. (1996) Cell cycle-dependent expression of the CHO2 antigen, a minus-end directed kinesin-like motor in mammalian cells. *Protoplasma,* **190**, 131-140.

Parr, D.R. and Edelman, J. (1975) Release of hydrolytic enzymes from the cell walls of intact and disrupted carrot callus tissue. *Planta*, **127**, 111-119.

Reski, R. (1998) Development, genetics and molecular biology of mosses. *Bot. Act.*

111, 1-15.

Ross-Macdonald, P., Sheehan, A., Roeder, G.S. and Snyder, M. (1997). A multipurpose transposon system for analyzing protein production, localization, and function in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA*, **94**, 190-195.

Sakakibara, K., Nishiyama, T., Kato, M. and Hasebe, M. (2001) Isolation of homeodomain-leucine zipper genes from the moss *Physcomitrella patens* and the evolution of homeodomain-leucine zipper genes in land plants. *Mol. Biol. Evol.* **18**, 491-502.

Salanoubat, M. *et al.* (2000) Sequence and analysis of chromosome 3 of the plant *Arabidopsis thaliana*. *Nature*, **408**, 820-822.

Sambrook, J., Fritch, E.F. and Maniatis, T. (1989) Molecular cloning: A Laboratory Manual. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Sawahel, W., Onde, S., Knight, C.D. and Cove, D.J. (1992) Transfer of foreign DNA into *Physcomitrella patens* protonemal tissue by using the gene gun. *Plant Mol. Biol. Rep.* **10**, 315-316.

Schaefer, D. (2001) Gene targeting in *Physcomitrella patens*. *Curr. Opin. Plant Biol.* **4**, 143-150.

Schaefer, D. G. and Zrÿd, J.-P. (1997) Efficient gene targeting in the moss *Physcomitrella patens*. *Plant J.* **11**, 1196-1206.

Schaefer, D., Zrÿd, J.-P., Knight, C. D. and Cove, C. D. (1991) Stable transformation of the moss, *Physcomirella patens*. *Mol. Gen. Genet.* **226**, 418-424.

Schumaker, K.D. and Dietrich, M.A. (1997) Programmed changes in form during moss development. *Plant Cell*, **9**, 1099-1107.

Schumaker, K.D. and Dietrich, M.A. (1998) Hormone-induced signaling during moss development. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **49**, 501-523.

Siefert, H.S., Chen, E.Y., So, M., and Heffron, F. (1986) Shuttle mutagenesis: A method of transposon mutagenesis for *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA*, **83**, 735-739.

Springer, P.S. (2000) Gene traps: Tools for plant development and genomics. *Plant Cell*, **12**, 1007-1020.

Steeves, T.A. and Sussex, I.M. (1989) Patterns in Plant Development. Cambridge

University Press, Cambridge.

Stewart, W. and Rothwell, G. (1993) Paleobotany and the evolution of plants. 2nd ed., Cambridge University Press, Cambridge.

Strepp, R., Scholz, S., Kruse, S., Speth, V. and Reski, R. (1998) Plant nuclear gene knockout reveals a role in plastid division for the homolog of the bacterial cell division protein FtsZ, an ancestral tubulin. *Proc. Natl. Acad. Sci. USA*, **95**, 735-739.

Sugimoto, M., Furui, S. and Suzuki, Y. (1995) Mutiple molecular forms of alpha-glucosidase from spinach seeds, *Spinacia oleracea* L. *Biosci. Biotechnol. Biochem.* **59**, 673-677.

Sun. Z., Duke, S.H. and Henson, C.A. (1995) The role of pea chloroplast alpha-glucosidase in transitory starch degradation. *Plant Physiol.* **108**, 211-217.

Sundaresan, V., Springer, P., Volpe, T., Haward, S., Jones, J.D.G., Dean, C., Ma, H. and Martienssen, R. (1995) Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements. *Genes Dev.* **3**, 1797-1810.

Sussex, I.M. and Kerk, N.M. (2001) The evolution of plant architecture. *Curr. Opin. Plant Biol.* **4**, 33-37.

Swofford, D. L. (2001) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sunderland, Massachusetts: Sinauer Associates.

Tabata S. *et al.* (2000) Sequence and analysis of chromosome 5 of the plant *Arabidopsis thaliana*. *Nature*, **408**, 823-826.

Tan, B. C. (1979) A new classification for the genus *Physcomitrella patens* B.S.G. *J. Hattori Bot. Lab.* **46**, 327-336.

Taylor, M.A., Ross, H.A., McRae, D., Stewart, D., Roberts, I., Duncan, G., Wright, F., Millam, S. and Davies, H.V. (2000) A potato alpha-glucosidase gene encoded a glycoprotein-processing alpha-glucosidase II-like activity. Demonstration of enzyme activity and effects of down-regulation in transgenic plants. *Plant J.* **24**, 305-316.

Theologis, A. *et al.* (2000) Sequence and analysis of chromosome 1 of the plant *Arabidopsis thaliana*. *Nature*, **408**, 816-820.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the

sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acid Res.* **22**, 4672-4680.

Tibbot, B.K. Henson, C.A. and Skadsen, R.W. (1998) Expression of enzymatically active, recombinant barley alpha-glucosidase in yeast and immunological detection of alpha-glucosidase gene from seed tissue. *Plant Mol. Biol.* **38**, 379-391.

Vale, R.D., and Fletterick, K. (1997) The design plan of kinesin motors. *Annu. Rev. Cell Dev. Biol.* **13**, 745-777.

Vos, J.W. Safadi, F., Reddy, A.S.N. and Hepler, P. (2000) The kinesin-like calmodulin binding protein differentially involved in cell division. *Plant Cell*, **12**, 979-990.

Yamasaki, Y. and Konno, H. (1987) Wall-bound alpha-glucosidase of suspension-cultured rice cells. *Phytochemistry*, **26**, 711-713.

Yamasaki, Y. and Konno, H. (1992) Wall-bound alpha-glucosidase of suspension-cultured sugar-beet cells. *Phytochemistry*, **31**, 2605-2607.

Waites, R. and Simon, R. (2000) Signaling cell fate in plant meristems. Three clubs on one tousle. *Cell.* **103**, 835-8.

Wolpert, L., Beddington, R., Brockes, J., Jessell, T., Lawrence, P. and Meyerowitz, E. (1998) Principles of development. Oxford University Press, London.

Watts, J.W., Doonan, J.H., Cove, D.J. and King, J.M. (1985) Production of somatic hybrids of moss by electrofusion. *Mol. Gen. Genet.* **199**, 349-351.