

University of New Hampshire

University of New Hampshire Scholars' Repository

Master's Theses and Capstones

Student Scholarship

Winter 2006

Calcium-dependent protein kinases in *Physcomitrella patens*

Kathryn Chervincky

University of New Hampshire, Durham

Follow this and additional works at: <https://scholars.unh.edu/thesis>

Recommended Citation

Chervincky, Kathryn, "Calcium-dependent protein kinases in *Physcomitrella patens*" (2006). *Master's Theses and Capstones*. 222.

<https://scholars.unh.edu/thesis/222>

This Thesis is brought to you for free and open access by the Student Scholarship at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in Master's Theses and Capstones by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact Scholarly.Communication@unh.edu.

CALCIUM-DEPENDENT PROTEIN KINASES IN

PHYSCOMITRELLA PATENS

BY

KATHRYN CHERVINCKY

Baccalaureate Degree, University of New Hampshire, 2004

THESIS

Submitted to the University of New Hampshire

in Partial Fulfillment of

the Requirements for the Degree of

Master of Science

in

Genetics

December, 2006

UMI Number: 1439266

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

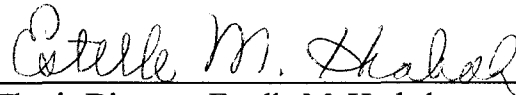
UMI Microform 1439266

Copyright 2007 by ProQuest Information and Learning Company.

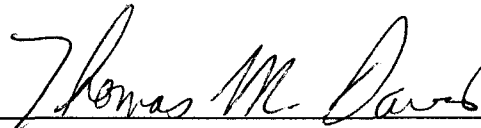
All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

This thesis has been examined and approved.



Thesis Director, Estelle M. Hrabak
Associate Professor of Plant Biology and Genetics



Thomas M. Davis,
Professor of Plant Biology and Genetics



John J. Collins, Associate Professor of
Biochemistry and Molecular Biology and Genetics

12-14-06

Date

TABLE OF CONTENTS

LIST OF TABLES	v
LIST OF FIGURES	vi
ABSTRACT	viii
CHAPTER	
I. INTRODUCTION	1
<i>Physcomitrella patens</i>	1
Calcium-Dependent Protein Kinases	3
The <i>GUS</i> Reporter Gene	6
Intron Effects on Gene Expression	8
RNA Interference (RNAi).....	13
II. MATERIALS AND METHODS.....	16
Identification of CDPKs.....	16
Expression of CDPK Proteins in <i>Escherichia coli</i>	17
CDPK Immunodetection.....	18
Kinase Assay	19
<i>CPK6</i> Promoter- <i>GUS</i> Constructs.....	20
<i>CPK6</i> Promoter- <i>GUS</i> Constructs Lacking the 5' UTR Intron.....	21
CDPK RNAi Construct.....	22

	<i>Physcomitrella</i> Growth Conditions.....	24
	<i>Physcomitrella</i> Transformation.....	24
	Selecting for Stable Transformants.....	25
	Histochemical GUS Assay.....	26
III.	RESULTS	27
	Identification of CDPKs.....	27
	Kinase Assay.....	45
	<i>CPK6</i> Promoter- <i>GUS</i> Expression	51
	CDPK RNAi	62
IV.	DISCUSSION	70
	Identification of CDPKs.....	70
	Kinase Assay.....	77
	<i>CPK6</i> Promoter- <i>GUS</i> Expression	78
	CDPK RNAi	82
	LIST OF REFERENCES	86

LIST OF TABLES

Table 1.	Percent identity and percent divergence between the variable domains of <i>Physcomitrella</i> CDPKs as calculated by MegAlign from the DNASTar package (Lasergene).....	30
Table 2.	Percent identity and percent divergence between the kinase domains of <i>Physcomitrella</i> CDPKs as calculated by MegAlign from the DNASTar package (Lasergene).....	32
Table 3.	Percent identity and percent divergence between the autoinhibitory domains of <i>Physcomitrella</i> CDPKs as calculated by MegAlign from the DNASTar package (Lasergene).....	34
Table 4.	Percent identity and percent divergence between the calmodulin-like domains of <i>Physcomitrella</i> CDPKs as calculated by MegAlign from the DNASTar package (Lasergene).....	36
Table 5.	Percent identity and percent divergence between full-length amino acid sequences of <i>Physcomitrella</i> CDPKs as calculated by MegAlign from the DNASTar package (Lasergene).....	37
Table 6.	Characteristics of predicted CDPK proteins from <i>Physcomitrella</i>	43
Table 7.	Variable domain characteristics of predicted CDPK proteins from <i>Physcomitrella</i>	44
Table 8.	Scintillation counts of <i>in vitro</i> kinase assays	49
Table 9.	Phylogenetic groupings of <i>Physcomitrella</i> and <i>Arabidopsis</i> CDPKs.....	74

LIST OF FIGURES

Figure 1.	Life Cycle of <i>Physcomitrella</i>	2
Figure 2.	Domain structure of calcium-dependent protein kinases	5
Figure 3.	Regulatory region of <i>CPK6</i> showing primer placement for overlap extension amplification	23
Figure 4.	Intron-exon structure of <i>Physcomitrella</i> CDPKs for which cDNA sequences were amplified	28
Figure 5.	Amino acid alignment of the variable domain of <i>Physcomitrella</i> CDPKs	29
Figure 6.	Amino acid alignment of the kinase domain of <i>Physcomitrella</i> CDPKs	31
Figure 7.	Amino acid alignment of the autoinhibitory domain of <i>Physcomitrella</i> CDPKs	33
Figure 8.	Amino acid alignment of the calmodulin-like domain of <i>Physcomitrella</i> CDPKs	35
Figure 9.	Dendrogram of <i>Physcomitrella</i> CDPKs.....	39
Figure 10.	Dendrogram of CDPKs <i>Physcomitrella</i> , <i>Arabidopsis</i> , <i>Oryza</i> , additional bryophytes, green algae and protist.....	41
Figure 11.	Schematic of CDPK fusion protein expressed in <i>E. coli</i>	46
Figure 12.	Detection of affinity purified CPK4 and CPK6	47
Figure 13.	Autoradiography after kinase assays.....	50
Figure 14.	Diagram of <i>CPK6</i> promoter- <i>GUS</i> cassette in pT-GUS-6a and pT-GUS-6b.....	52
Figure 15.	Diagram of <i>CPK6</i> promoter- <i>GUS</i> cassette in pT-GUS-6a (-) and pT-GUS-6b (-).....	53
Figure 16.	Anatomy of a ~4 week old <i>Physcomitrella</i> plant.....	54

Figure 17.	Light micrographs of <i>Physcomitrella</i> stably transformed with pT-GUS-6a; transformant #1	55
Figure 18.	Light micrographs of <i>Physcomitrella</i> stably transformed with pT-GUS-6a; transformant #2	56
Figure 19.	Light micrographs of <i>Physcomitrella</i> stably transformed with pT-GUS-6b; transformant #1	57
Figure 20.	Light micrographs of <i>Physcomitrella</i> stably transformed with pT-GUS-6a(-); transformant #1	59
Figure 21.	Light micrographs of <i>Physcomitrella</i> stably transformed with pT-GUS-6b(-); transformant #1	60
Figure 22.	Light micrographs of <i>Physcomitrella</i> stably transformed with pT-GUS-6b(-); transformant #2	61
Figure 23.	Segment of DNA from a conserved region of the kinase domain used for CPK6 RNAi aligned to corresponding segments of closely-related CDPKs	64
Figure 24.	Diagram of pTUGGi-RNAi6 construct for CDPK silencing.....	65
Figure 25.	Diagram of pTUGi construct used as control for RNAi	66
Figure 26.	<i>Physcomitrella</i> NLS4 line	67
Figure 27.	<i>Physcomitrella</i> NLS4 line transiently transformed with pTUGi	68
Figure 28.	<i>Physcomitrella</i> NLS4 line transiently transformed with pTUGGi-RNAi6.....	69

ABSTRACT

CALCIUM-DEPENDENT PROTEIN KINASES IN

PHYSCOMITRELLA PATENS

by

Kathryn Chervincky

University of New Hampshire, December, 2006

Calcium-dependent protein kinases (CDPKs) are a group of plant serine-threonine kinases that are stimulated in response to increases in intracellular calcium concentrations. *Physcomitrella patens* is a moss and a member of the bryophyte family. *Physcomitrella* has gained popularity as a model research organism due to the simple morphology of the plant as well as many other desirable characteristics. My research utilized *Physcomitrella* to accomplish four goals. The first goal was to identify and obtain sequence for as many CDPKs as possible in *Physcomitrella*. The second goal was to confirm the calcium-activated protein kinase activity of the identified CDPKs. The third goal consisted of two parts: first, to determine the expression pattern of a specific CDPK (*CPK6*) and second, to determine the role of a 5' UTR intron in *CPK6* expression. The fourth goal was to silence expression of *CPK6*, and possibly other closely related CDPKs, by RNA interference (RNAi) in order to screen for a phenotype.

My results showed that there are at least 16 CDPKs in *Physcomitrella patens*. Two of those CDPKs were tested and displayed calcium-activated protein kinase activity. Expression analysis of *CPK6* showed strongest expression in protonemal tip cells, immature phyllids, base of the gametophore at the site of protonemal attachment, and

occasionally in axillary hair cells. Weaker expression was seen in non-tip protonemal cells and mature phyllids. Removal of the 5' UTR intron resulted in overall decreased expression levels, but did not change the expression pattern. RNAi results indicated that the RNAi response pathway was activated, however no phenotype was observed under normal growth conditions.

CHAPTER I

INTRODUCTION

Physcomitrella patens

Physcomitrella patens is a moss and a member of the Bryophyte family, the oldest group of land plants that is presently part of Earth's flora. Bryophyta includes all mosses, hornworts and liverworts. Organisms in this group lack the complexity of true leaves, roots, and a vascular system, all of which are present in other groups of land plants.

Physcomitrella has been recently gaining popularity as a research organism due to many characteristics that make it a convenient plant to study. First, unlike most vascular plants, *Physcomitrella* spends most of its life cycle in the haploid gametophytic phase (Figure 1), and has only a brief, diploid sporophytic phase (Schaefer and Zryd, 2001). This is an attractive attribute for a research organism since there is only one copy of each gene that needs to be inactivated when creating a knockout mutant. Second, *Physcomitrella* more efficiently undergoes homologous recombination with introduced DNA than other land plants do (Schaefer and Zryd, 1997). This characteristic allows for site-directed mutagenesis of genes and provides increased reproducibility in expression analyses using reporter genes since the site of gene insertion can be controlled. Third, *Physcomitrella* has successfully been used in RNA interference (RNAi) experiments where gene expression was downregulated through mRNA degradation (Bezanilla et al., 2005). Fourth, as confirmed by my own experience, *Physcomitrella* is a very compact plant and is easily cultured on simple growth medium in a laboratory environment. The plant can also remain dormant for years when refrigerated, allowing for easy storage (E. Hrabak,

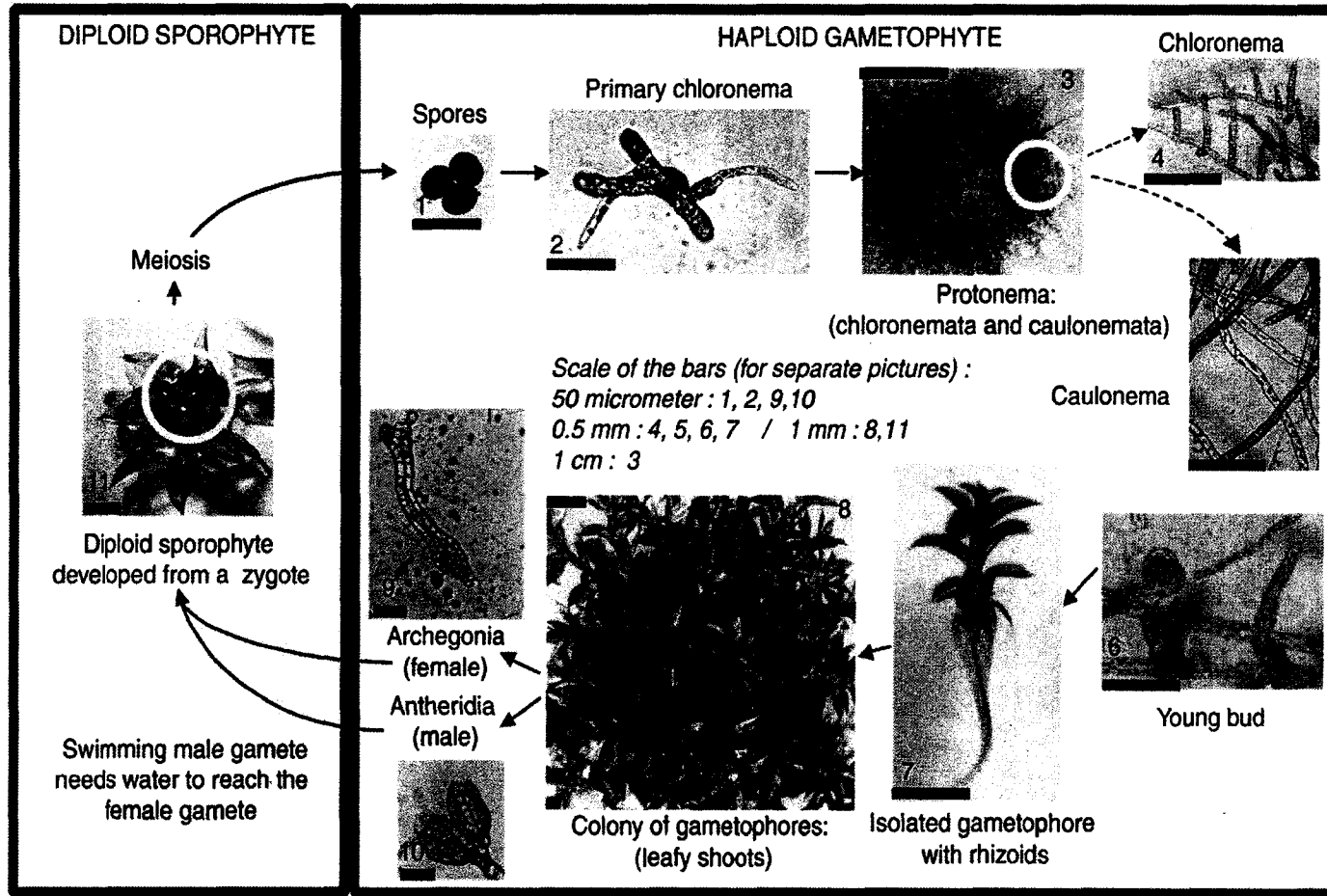


Figure 1. Life cycle of *Physcomitrella* (from Schaefer and Zryd, 2001). Stages are listed in order from 1 to 11, starting with spores and ending with sporophyte (in the diploid phase, circled yellow).

personal communication). Fifth, the sequence of the relatively small genome of 460 Mb (three times the size of the popular model vascular plant *Arabidopsis thaliana*) has been completed recently and is currently being annotated (<http://www.jgi.doe.gov>). My research utilizes *Physcomitrella* to investigate calcium-dependent protein kinases (CDPKs).

Calcium-Dependent Protein Kinases

Ca^{2+} is an important second messenger in many cellular signal transduction pathways. Fluctuations in calcium levels occur within cells in response to external stimuli such as salt and drought stress (Knight et al., 1997), hypo-osmotic stress (Takahashi et al., 1997), cold stress (Knight et al., 1996) and oxidative stress (Price et al., 1994). Each stimulus is predicted to cause differences in the amplitude, duration and frequency, as well as the location, of the Ca^{2+} signal within the cell (Sanders et al., 1999). Differences in Ca^{2+} spiking lead to different biochemical responses within the cell. For these biochemical responses to occur, different types of calcium signals must be perceived correctly and then converted into responses. Examples of enzymes that respond to calcium signals include protein kinase C (PKC), two types of calcium/calmodulin-dependent protein kinases (CaMKs and CCaMKs) and calcium-dependent protein kinases (CDPKs). Each of these calcium-sensing kinases is activated by Ca^{2+} in a different way. Protein kinase C is fully activated by a two-step process. First Ca^{2+} binds to cytosolic PKC, which results in the exposure of the PKC phospholipid-binding site through a conformational change of the protein. After binding Ca^{2+} , PKC translocates to the membrane and binds diacylglycerol, resulting in the fully

active form of the enzyme (Oancea and Meyer, 1998). CaMKs are activated by the binding of a calcium-calmodulin complex, where calmodulin is a protein that binds Ca^{2+} (Colbran et al., 1988). CCaMKs are slightly activated by the binding of Ca^{2+} , but require binding of a calcium-calmodulin complex for full activation (Takezawa et al., 1996). CDPKs are activated directly by the binding of Ca^{2+} and do not require calmodulin (Harper et al., 1991). CaMKs and PKCs are the two major types of kinases that respond to Ca^{2+} signaling in animal cells, but do not seem to be present in plants (Hrabak et al., 2003). CCaMKs and CDPKs are two of the major kinases that respond to Ca^{2+} in plant cells and do not seem to be present in animal cells (Hrabak et al., 2003). Only CDPKs will be discussed in detail here since they are the focus of my research.

CDPKs form a major family of serine/threonine kinases that are found only in plants and organisms of the Apicomplexan protist subgroup (Hrabak, 2000) such as *Plasmodium falciparum*. These kinases phosphorylate their corresponding substrate proteins in a Ca^{2+} -dependent manner, making them the only group of kinases consistently found in all plants that are directly activated by Ca^{2+} . The calcium-dependent protein kinase activity of CDPKs makes them prime candidates for interpreting Ca^{2+} fluctuations within plant cells and converting the calcium signal into a biochemical response by activating or inactivating substrate proteins through phosphorylation. Certain CDPKs are involved in responses to stresses such as drought, salt and cold (Sheen, 1996; Saijo et al., 2000) and pathogen attack (Romeis et al., 2001).

Each CDPK enzyme consists of four domains: the variable domain, the catalytic domain, the autoregulatory domain, and the calmodulin-like domain (Figure 2). The variable domain differs in length and amino acid content between individual CDPKs. In

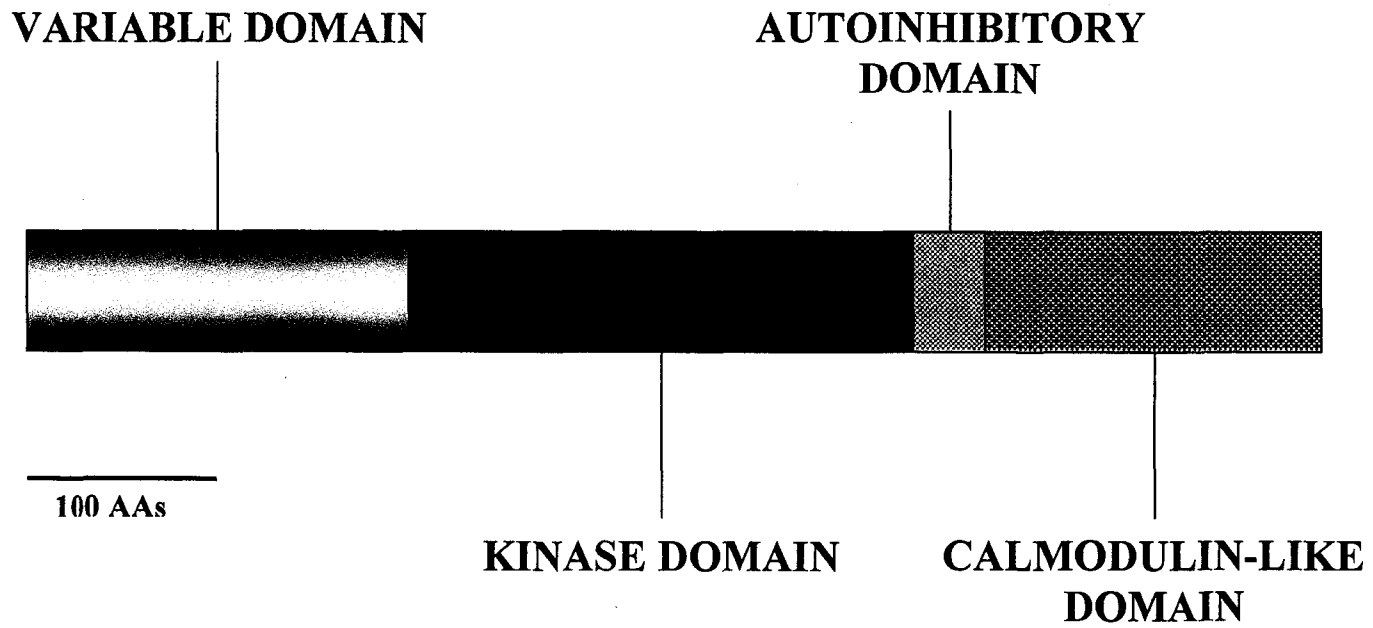


Figure 2. Domain structure of calcium-dependent protein kinases. Variable domain is not to scale.

some CDPKs, this domain contains sites for the addition of the fatty acids myristate and palmitate that enable localization of CDPKs to membranes (Resh, 1999; Martin and Busconi, 2000; Lu and Hrabak, 2002; A. Argyros and E. Hrabak, unpublished data). The catalytic domain transfers the γ -phosphate of ATP to a serine or threonine residue of the corresponding substrate protein (Hrabak et al., 2003). The autoinhibitory domain acts as a pseudosubstrate of the kinase domain, inhibiting phosphorylation (Harmon et al., 1994; Harper et al., 1994). This inhibition is released when calcium binds to the four EF-hands of the calmodulin-like domain causing a conformational change in the protein that results in the removal of the autoinhibitory domain from the kinase domain (Harmon et al., 1987).

No previous research has been done on CDPKs and CDPK function in *Physcomitrella*. My research includes identifying and obtaining sequence for as many CDPKs in *Physcomitrella* as possible and confirming calcium-dependent protein kinase activity of *CPK4* and *CPK6*. Additional experiments focus on *CPK6* with the goals of defining the expression pattern of *CPK6*, determining whether an intron found in the 5' UTR of *CPK6* affects gene expression, and attempting to silence expression of *CPK6*, as well as three closely related CDPKs in the same *Physcomitrella* transformant, through the use of RNAi.

The *GUS* Reporter Gene

Understanding the regulation of gene expression is instrumental in discerning the function of any given gene. Reporter genes are a class of naturally-occurring genes that have been isolated and manipulated in such a way that they become useful in indicating

the expression pattern of a gene of interest. In order to utilize a reporter gene, it is fused downstream of the promoter of a gene of interest. There are two general types of reporter gene fusions: the reporter can be fused immediately downstream of the promoter (transcriptional fusion) or the reporter gene can be fused to the open reading frame (translational fusion). If all regulatory elements are contained in the DNA segment that has been fused to the reporter gene, then either approach should result in the reporter gene being transcribed in the same fashion as the gene of interest. The protein product of the reporter gene can then be detected in order to characterize the expression profile of the gene of interest. The use of reporter genes has become common in research. Due to the popularity of reporter genes, many different types have been developed and include *GUS* (β -glucuronidase; Jefferson, 1987), *GFP* (green fluorescent protein; Chalfie et al., 1994) and *LUC* (luciferase; Palomares et al., 1989). Only *GUS* will be described here in detail since this is the reporter gene I utilized in my research.

The *GUS* reporter gene was isolated from *Escherichia coli* and encodes β -glucuronidase. This enzyme functions as a tetramer whose monomers have a molecular weight of 68.2 kDa (Jefferson, 1987). During colorimetric detection, β -glucuronidase cleaves the colorless substrate X-Gluc (5-bromo-4-chloro-3-indolyl β -D-glucuronide) into glucuronic acid and compound X, which dimerizes to form an insoluble blue product called dichloro-dibromo-indigo (Lojda, 1970). When the coding region of β -glucuronidase is fused to the promoter or coding region of a gene of interest, expression of β -glucuronidase should occur in the same pattern throughout the organism as the gene in question. The organism can then be incubated with the X-gluc substrate, which will result in a blue color wherever the given gene is normally expressed. However, some

drawbacks to using the *GUS* reporter gene to determine expression profiles include the fact that the half-life of the GUS protein may be different than the half-life of the protein of the gene of interest, and it is always possible that not 100% of the gene regulatory information was included in the promoter-*GUS* construct. Aside from these drawbacks, the *GUS* reporter gene method is likely to give a close to accurate expression profile of the gene of interest.

Intron Effects on Gene Expression

Introns are sequences that are spliced out of pre-mRNA transcripts in eukaryotes during processing in the nucleus. Two functions that have been described for introns include the ability to produce multiple proteins from a single gene through alternative intron splicing and allowing the evolution of new protein encoding genes through exon shuffling (Le Hir et al., 2003). More recently, introns have been shown to positively influence gene expression (Furger et al., 2002; Nott et al., 2004; Chung et al., 2006; Morello et al., 2006). Introns are thought to be involved in many stages of gene expression, including transcription, processing of the initial transcript and nuclear export, as well as translation and decay of the mature mRNA (Le Hir et al., 2003). In addition, introns found in the 5' UTR have a greater positive effect on gene expression than introns found in the coding region or the 3' UTR. Any positive effect of an intron in the coding region is reduced the farther downstream the intron is placed from the start codon (Rose, 2004; Chung et al., 2006). Genes that do not have introns in the 5' UTR often show enhancement of expression if one is inserted (Chaubet-Gigot et al., 2001).

Some introns affect transcription of genes since a reduction in transcription has been observed following removal of introns from the coding region, inactivation of the 5' splice site, or moving the position of the splice site farther downstream of the transcription start site (McKenzie and Brennan, 1996; Furger et al., 2002). Some transgenes have shown 10 to 100 times greater transcription efficiency than genes that are otherwise identical but lack introns (Brinster et al., 1988). It is also interesting to note that although only ~3.8% of genes in the *Saccharomyces cerevisiae* genome contain introns (Spingola et al., 1999), nearly 10,000 of the almost 38,000 mRNA transcripts produced per hour in yeast cells are from this small percentage of intron-containing genes (Holstege et al., 1998). Two different mechanisms can regulate gene expression at the transcriptional level. First, the presence of introns can result in an ordered nucleosome arrangement that becomes disrupted upon removal of the introns. It is hypothesized that a signal present in the genomic sequence results in a characteristic nucleosome arrangement that optimizes RNA polymerase binding (Liu et al., 1995). Second, transcription can be upregulated due to the presence of enhancer sequences contained within an intron immediately upstream of the first exon (Sleckman et al., 1996).

Introns can affect additional aspects of pre-mRNA processing which may impact gene expression. An increase in polyadenylation efficiency has been seen in response to intron splicing (Vagner et al., 2000). Since the poly A tail is involved in a variety of processes such as mRNA stability, export from the nucleus and translation in the cytoplasm, increased efficiency of polyadenylation could lead to an increase in expression (Furger et al., 2002).

Three processes subsequent to pre-mRNA processing are affected based on whether or not an mRNA has been spliced: nuclear export of mRNA, translation efficiency of the mRNA once in the cytoplasm, and stability of the mRNA as regulated by nonsense-mediated decay (NMD; (Le Hir et al., 2003)). The act of splicing a pre-mRNA results in the binding of the exon junction complex (EJC) 20-24 bp upstream of exon-exon junctions (Le Hir et al., 2000). At least seven proteins have been identified in the EJC of *Drosophila melanogaster*: REF/Aly, RNPS1, SRm160, Y14, Magoh, UAP56 and Upf3. Each of these proteins, with the exception of Upf3, is involved in at least one of the following processes: RNA splicing, nuclear export, or shuttling between the nucleus and cytoplasm (Blencowe et al., 1998; Mayeda et al., 1999; Conti and Izaurralde, 2001; Hachet and Ephrussi, 2001; Mohr et al., 2001; Reed and Hurt, 2002). Upf3 is a nonsense-mediated decay (NMD) factor, which will be explained in more detail below. After export from the nucleus to the cytoplasm, Y14 and Magoh remain bound to the mRNA and are not removed until the first ribosomal passage (Le Hir et al., 2003). All other known nuclear EJC proteins are no longer associated with the mRNA in the cytoplasm (Le Hir et al., 2001a). In addition, the cytoplasmic NMD factor Upf2 associates with the mRNA once the transcript reaches the cytoplasm.

There is strong evidence that the EJC is involved in mRNA nuclear export. Since EJCs bind 20-24 nt upstream of exon-exon junctions, a truncated mRNA that would be too short (<19 bases in the 5' exon) to allow EJC binding was created by the Le Hir lab at Brandeis University. The Le Hir lab also created a second mRNA with a 5' exon longer than 19 bases, but otherwise identical to the truncated mRNA, that would allow EJC binding. The longer mRNA was shown to be associated with all aforementioned EJC

proteins and was efficiently exported whereas the shorter mRNA was not efficiently exported and was not associated with the EJC proteins (Le Hir et al., 2001b). Since intron splicing is needed for the binding of the EJC to mRNAs and the EJC allows for efficient transport of mRNAs out of the nucleus, these results show that introns are indirectly involved in efficient export of mRNA into the cytoplasm.

An increase in translational yield and higher ribosomal affinity has been observed in spliced mRNA when compared to an identical version of the mRNA that was not produced by splicing (Nott et al., 2004). Furthermore, translational yield and ribosomal affinity were restored when individual EJC proteins or NMD factors were tethered to an mRNA that was not produced by splicing (Nott et al., 2004). These results suggest that EJC and NMD proteins, deposited either by tethering or by post-splicing events, increase translational efficiency of the mRNA. Other studies have also shown that more protein molecules are produced per molecule of spliced mRNA than from identical mRNAs that were not produced by splicing (Braddock et al., 1994; Matsumoto et al., 1998; Lu and Cullen, 2003; Wiegand et al., 2003) and this increase in translation is independent of the rate of nuclear export.

Nonsense-mediated decay (NMD) is the process whereby mRNAs containing a premature stop codon, due to processes such as mutation, are degraded. In mammalian cells, the stop codon is defined as premature when it is located 50 nucleotides or more upstream of an exon-exon junction (Maquat and Carmichael, 2001). It is not surprising then to find that few mammalian genes contain introns in their 3' UTR (Le Hir et al., 2003). Since an NMD factor is always associated with the EJC which is located 20-24 nucleotides upstream of all exon-exon junctions, a mechanism is provided that explains

how exon-exon junctions that occur after stop codons in mammalian cells are recognized for NMD (Le Hir et al., 2003). Since EJC proteins remain bound to mRNA until the first ribosomal passage, if the EJC occurs after the stop codon, EJC proteins will not dissociate as a function of ribosomal passage and the NMD factor will most likely remain bound to the mRNA. This is the only case in which the presence of an intron results in downregulation of gene expression.

It is well documented that introns in the 5' UTR usually enhance gene expression in vascular plants. Examples of genes that are upregulated due to an intron in the 5' UTR include the triosephosphate isomerase gene (*tpi*) in *Oryza sativa*, the sucrose synthase gene (*Sus3*) in potato and the polyubiquitin gene (*Ubi.U4*) in *Nicotiana tabacum* (Fu et al., 1995; Snowden et al., 1996; Plesse et al., 2001). A 291 bp intron in the 5' UTR of a CDPK gene has also been shown to influence expression in *Oryza* (Morello et al., 2006). This CDPK study was particularly interesting because, not only did the removal of the 5' UTR intron almost completely abolish *OsCDPK2* expression in both stable and transient transformants, but the removal of 83 bp at the 3' end of the intron, which resulted in the inactivation of the 3' splice site of the intron, also showed almost complete absence of expression in transient and stable transformants as detected through the use of the *GUS* reporter gene (Morello et al., 2006). Since the majority of the intron was left intact, these results show that the act of splicing is more likely responsible for enhanced expression of the CDPK than sequences within the intron itself, indicating that the EJC is involved in increasing expression. Little has been published on intron-mediated enhancement of gene expression in moss and no results have been published on intron-mediated enhancement of CDPKs in non-vascular plants.

RNA Interference (RNAi)

RNA interference or RNAi is the general term for targeted degradation of mRNA through the RNA-induced silencing complex (RISC) in eukaryotes. There are two types of RNAi triggers: microRNA (miRNA) and small interfering RNA (siRNA). miRNAs are created from stem loops in endogenous RNAs and commonly results in gene down-regulation (Sontheimer and Carthew, 2005). siRNAs are created from long double-stranded RNA from a variety of sources, such as pathogens and transposons, and result in the targeted degradation of RNA (Sontheimer and Carthew, 2005). The RNAi pathway can be taken advantage of by creating artificial triggers to downregulate or silence expression of a gene of interest. Here I will focus on siRNA since this is what I utilized in my research.

Long dsRNA is recognized by the Dicer enzyme, a ~200 kDa protein that generally contains an ATPase/helicase domain, a PAZ domain, two catalytic RNase III domains, and a C-terminal dsRNA binding domain (Filipowicz, 2005). Dicer cleaves dsRNA precursors into segments called siRNA duplexes that are 21-25 nucleotides in length. The RISC binds the siRNA duplexes produced by Dicer. Once an siRNA duplex is incorporated into the RISC, the double-stranded siRNA duplex is unwound through an unknown mechanism, although ATP-dependent helicases embedded within the complex are suspected to be involved (Preall and Sontheimer, 2005). Only one of the siRNA strands is retained in the RISC. It is believed that which strand is retained is based on its thermodynamic properties, typically low internal stability especially at the 5' end (Khvorova et al., 2003). Once the siRNA duplex becomes a single-stranded siRNA, it is

used as a guide to target the RISC to RNA molecules containing complementary sequence. Once bound, the RISC will degrade the RNA molecule to which the siRNA has annealed. This results in either the silencing or downregulation of any gene that produces an mRNA with sequence similarity above the required threshold of identity to the siRNA.

The RISC has been shown to have multiple proteins associated with it; however the only group of proteins consistently found in all RISC complexes are the Argonautes (Parker et al., 2004). All Argonaute proteins are composed of an N-terminal PAZ domain of ~20 kDa and a C-terminal PIWI domain of ~40 kDa (Parker et al., 2004). Recently, the PAZ domain has been shown to interact with RNA and may serve as the binding site for the 21-25 nucleotide siRNA (Ma et al., 2004). The PIWI domain has been shown to contain an RNase H-like fold, suggesting that this domain of eukaryotic Argonautes could possess RNase activity and may be responsible for the cleavage of target mRNAs (Parker et al., 2004). Phylogenetic analysis shows that the Argonaute group can be divided into two subgroups, Ago and Piwi; however no functional distinction between the two subgroups has yet been detected (Carmell et al., 2002).

RNAi has been utilized in *Physcomitrella* and in some cases has produced a phenotype. The *FtsZ2-1* gene, whose loss of function is known to produce altered chloroplast morphology in vascular plants (Strepp et al., 1998), was successfully silenced to produce the altered chloroplast phenotype in *Physcomitrella* through the use of RNAi (Bezanilla et al., 2005). RNAi has also been used to silence a CDPK in *Medicago truncatula*, which resulted in a significant reduction in root hair and root cell length, as well as significantly lower levels of rhizobial and mycorrhizal symbiotic association

(Ivashuta et al., 2005). RNAi of the same CDPK also resulted in altered expression of cell wall and defense-related genes (Ivashuta et al., 2005). RNAi is capable of silencing multiple members of a closely-related gene family when using an siRNA trigger sequence that has ~80% sequence identity to the target genes (Miki et al., 2005). Taken together, these facts support the hypothesis that RNAi can be utilized in *Physcomitrella* to silence multiple closely-related CDPKs. Since closely-related CDPKs are thought to have functional redundancy (Cheng et al., 2002), silencing of multiple CDPKs may be necessary to observe a phenotype.

CHAPTER II

MATERIALS AND METHODS

Identification of CDPKs

CDPKs were identified using two methods. The first method consisted of work done by Dr. Estelle Hrabak using *Arabidopsis thaliana* CDPKs to execute nucleotide-nucleotide BLAST searches (blastn) for *Physcomitrella* expressed sequence tags (ESTs) on PHYSCObase (<http://moss.nibb.ac.jp>). For ESTs that showed sequence identity to the *Arabidopsis* CDPKs, primers were designed using the OLIGO program (Molecular Biology Insights) to bind within both the 5' and 3' ESTs. Each open reading frame along with any available untranslated region (UTR) sequences was amplified by from *Physcomitrella* total cDNA template by either myself or Dr Estelle Hrabak. In addition, the entire transcribed region between each primer set was amplified from genomic DNA. PCR products were cloned into the pCR[®]2.1-TOPO[®] vector (Invitrogen) for sequencing.

The second method of identifying CDPKs utilized the *Physcomitrella* CDPK sequences obtained from the previous method. I used these sequences to execute nucleotide-nucleotide BLAST searches (blastn) against the *Physcomitrella* genome project in the trace archives at NCBI (<http://ncbi.nih.gov>) to identify individual sequences that showed identity or similarity to the previously identified CDPKs. At the time that this research was performed, the genome had not been assembled and only individual sequence fragments of about 600-1000 bp were available. These sequences were sorted and assembled into individual CDPK gene coding regions using the SeqMan

program from the DNASTar package (Lasergene). Sequences upstream and downstream of the coding regions ranging from 300 bp to 2 kb were also assembled.

Expression of CDPK Proteins in *Escherichia coli*

The coding sequences, including the stop codons, of CDPK genes were amplified from genomic DNA. PCR products were cloned into the pET200/D-TOPO[®] (Invitrogen) expression vector. These expression plasmids were provided by Estelle Hrabak. Upstream of the cloning site, pET200/D-TOPO contains sequence which adds a 6xHis tag and an Xpress epitope to the N-terminus of the expressed CDPK protein. Clones were sequenced by the University of New Hampshire sequencing facility to determine whether or not clones were error-free. An error-free clone of each CDPK was electroporated into *Escherichia coli* BL21(DE3). Cells containing the expression vector were grown to an OD₆₀₀ of ~0.5 in 50 ml LB broth [1% tryptone, 0.5% yeast extract, 1% sodium chloride], then induced with 1 mM IPTG for 2.5 hours. Cells were pelleted and resuspended in 1.25 ml TBS [140 mM NaCl, 20 mM Tris-Cl, pH 7.6] containing 0.1% Tween 20, 2 mM EDTA, 1 mM DTT, and 1 mM PMSF. Cells were lysed with a 100 µg/ml lysozyme treatment for 0.5 hours on ice followed by five sonications of 5 sec each with chilling on ice for 1 min between each sonication. The supernatant was added to 0.5 ml HIS-Select Nickel Affinity Gel (Sigma) that had been washed with 10 ml of equilibration buffer [0.3 M NaCl, 50 mM sodium phosphate, pH 8.0]. The *E. coli* protein extract and affinity gel were incubated at 4°C for 30 min with gentle agitation. The nickel gel complexed with 6His-tagged protein was pelleted at 500 x g for 2 min, resuspended in wash buffer [50 mM sodium phosphate, pH 8.0, 0.3 M sodium chloride,

10 mM imidazole], and re-pelleted. The wash was repeated a total of three times. The protein was eluted by resuspending the nickel gel in elution buffer [50 mM sodium phosphate, pH 8.0, 0.3 M sodium chloride, 250 mM imidazole] and incubating with agitation (~175 rpm) for five min at room temperature. The nickel gel was pelleted at 500 x g for 2 min and supernatant containing protein was collected. The elution process was repeated a total of three times. Eluted proteins were stored at -80°C.

CDPK Immunodetection

Samples containing either 1 µg or 5 µg of protein were combined with an equal volume of sample loading buffer [2% (w/v) SDS, 62.5 mM Tris-HCl, pH 6.8, 10% (v/v) glycerol, 5% (v/v) β-mercaptoethanol, 0.05% (w/v) bromophenol blue] on an 8% SDS polyacrylamide gel (Harlow and Lane, 1988). The protein was transferred to a PVDF membrane (Millipore) by electroblotting in transfer buffer [25 mM Tris-HCl, 192 mM glycine, 0.025% (w/v) SDS] for one hour following manufacturer's instructions.

The PVDF membrane was incubated in blocking buffer (TBS-T [20 mM Tris-Cl, pH 7.6, 137 mM NaCl, 0.1% Tween-20] containing 5% non-fat dry milk) at 4°C overnight. The membrane was then incubated at room temperature in blocking buffer containing anti-Xpress antibody conjugated to horseradish peroxidase (Invitrogen) at a 1:5000 dilution for one hour. The membrane was washed in TBS-T containing 0.5% non-fat dry milk for 10 minutes at room temperature for a total of three times. Antibody was detected by incubating the membrane in SuperSignal West Pico Chemiluminescent Substrate (Pierce) for ~5 minutes, followed by exposure of X-ray film (ISC BioExpress). Size of the proteins was determined using the BenchMark protein ladder (Invitrogen).

Kinase Assay

Three separate 50 μ l reactions were set up for each CDPK: one with substrate and calcium, one lacking substrate, and one lacking calcium. Reactions consisted of 1X kinase buffer [20 mM Tris-Cl, pH 7.3, 6 mM MgCl₂], 0.5 mg/ml Histone III-S (Sigma) (not present in reactions lacking substrate), 0.1 mM CaCl₂ (replaced with 0.2 mM EGTA in reactions lacking calcium), 50 μ M adenosine triphosphate, 5 μ Ci adenosine 5'-[γ -³²P]triphosphate (Amersham, catalog # AA0068), and 1 μ g of CDPK (expressed in and isolated from *E. coli* as described above). Reactions were incubated at room temperature for 20 min.

After incubation, 25 μ l of each reaction was spotted on blotting paper (VWR), which was washed three times at 5 min per wash in ~250 mls 10% trichloroacetic acid, 10 mM disodium pyrophosphate. Filter papers were counted for radioactivity using a scintillation counter.

The remaining 25 μ l of each reaction was added to 25 μ l of sample loading buffer. Twenty μ l of each sample was electrophoresed on an 8% SDS-polyacrylamide gel and incubated with Coomassie stain [0.1% Coomassie blue R-250, 40% methanol, 10% acetic acid] for fifteen min with gentle agitation. The gel was incubated in destaining solution [40% methanol, 10% acetic acid] with gentle agitation overnight. Gel drying was done between two sheets of cellophane (Research Products International Corp.). The dried gel was exposed to X-ray film.

CPK6 Promoter-GUS Constructs

Using *Pfu* Ultra DNA polymerase (Stratagene), the *GUS* gene-*NOS* terminator cassette was amplified from pBI101.2 (Clontech) using a forward primer that added both *Sal*I and *Hind*III sites and a reverse primer that added an *Xho*I site (restriction sites are bold and underlined):

Forward:

5'-TTT**GTCGACTTCTCGAGTTTAAGCTT**GGTAGGTCAGTCCCTTATG-3'

Reverse:

5'-**ACTCGAGCCAGTGAATTC**CCGATCTAG-3'

The PCR product was co-digested with *Sal*I and *Xho*I (Promega) for 4 hours following manufacturer's instructions. The pT vector (obtained from the Bezanilla lab at the University of Massachusetts, Amherst) that contains a unique *Sal*I site was also digested with *Sal*I for 4 hrs. Ligation of the digested PCR product into the pT vector was done using T4 DNA ligase (Promega) to create the pT-GUS construct. The ligation reaction was electroporated into *E. coli* TOP10 cells. Bacteria were selected by plating on LB plates containing 100 µg/ml carbenicillin. Resistant colonies were checked for presence of the *GUS* terminator-*NOS* insert in the pT vector by PCR amplification and by restriction digestion with *Sal*I and *Xho*I. One clone containing the *GUS* terminator-*NOS* cassette was saved and named pT-GUS.

A 3 kb region of *CPK6* genomic DNA consisting of 1.8 kb of sequence upstream of the 5' UTR, the 1.1 kb 5' UTR containing a ~0.3 kb intron, and 110 bp of coding sequence was amplified in two independent reactions using *Ex Taq* DNA polymerase

(Takara) and primers that added a *SalI* site at the 5' end and a *HindIII* site at the 3' end (restriction sites are bold and underlined):

Forward: 5'-TTT**GTCGAC**CTTTTTCCAACCTTTGATGTC-3'

Reverse: 5'-GGGA**AGCTT**ATTGGTCTTTCGCTCTCATTC-3'

Each PCR product was digested with *SalI* and *HindIII* in sequential reactions for 4 hrs per digestion. The pT-GUS construct was digested with the same enzymes. The digested products from each PCR were ligated into the pT-GUS construct at the *SalI* and *HindIII* sites using T4 DNA ligase to create the pT-GUS-6a and pT-GUS-6b constructs for *CPK6* promoter-*GUS* assays. Constructs were electroporated into the *E. coli* TOP10 strain and plated on LB plates containing 100 µg/ml carbenicillin. Colonies were screened for plasmids containing the correct insert by checking with PCR amplification and by restriction digestion with *SalI* and *HindIII*.

CPK6 Promoter-GUS Constructs Lacking the 5' UTR Intron

Overlap extension amplification (Horton et al., 1993) was used to create *CPK6* promoter-*GUS* constructs in which the intron in the 5' UTR was removed. The forward and reverse primers from the previous section were used, as well as a second nested set of forward and reverse primers designed to flank the intron:

Nested Fwd. 5'-**GGCACTGGAGTTGGTTTTG**-3'

Nested Rev. 5'-**CAAACCAACTCCAGTGCCGTAGTTGGCAAGCGCTTCAG**-3'

The nested reverse primer contained 19 bp of sequence (bold and underlined) that overlapped with the 5' sequence of the nested forward primer (Figure 3). DNA sequences upstream and downstream of the intron were amplified with *Pfu* Ultra DNA polymerase in separate reactions pairing the forward primer with the nested reverse primer and the nested forward primer with the reverse primer, respectively. Once amplified, the upstream and downstream PCR products were combined in a new PCR with the un-nested forward and reverse primers to amplify the *CPK6* regulatory region lacking the 5' UTR intron. Two overlap extension reactions were performed independently using pT-GUS-6a and pT-GUS-6b as templates. The final PCR products and the pT-GUS vector were digested with *SalI* and *HindIII* and combined in a ligation reaction with T4 DNA ligase to create pT-GUS-6a(-) and pT-GUS-6b(-). Constructs were electroporated into *E. coli* TOP10 cells. Transformants were selected on LB plates containing 100 ug/ml carbenicillin. Colonies were screened for plasmids containing the correct insert by PCR amplification and by restriction digestion with *SalI* and *HindIII*.

CDPK RNAi Construct

An RNAi construct was made using the pTUGGi Gateway vector (Bezanilla et al., 2005). A 255 bp segment of the *CPK6* kinase region which had ~80-90% sequence identity to the analogous region in the *CPK1*, *CPK3*, *CPK4* and *CPK8* genes was amplified using a forward primer with a 5' *attB1* sequence and a reverse primer with a 5' *attB2* sequence for Gateway cloning (*attB1* and *attB2* sites are bold and underlined):

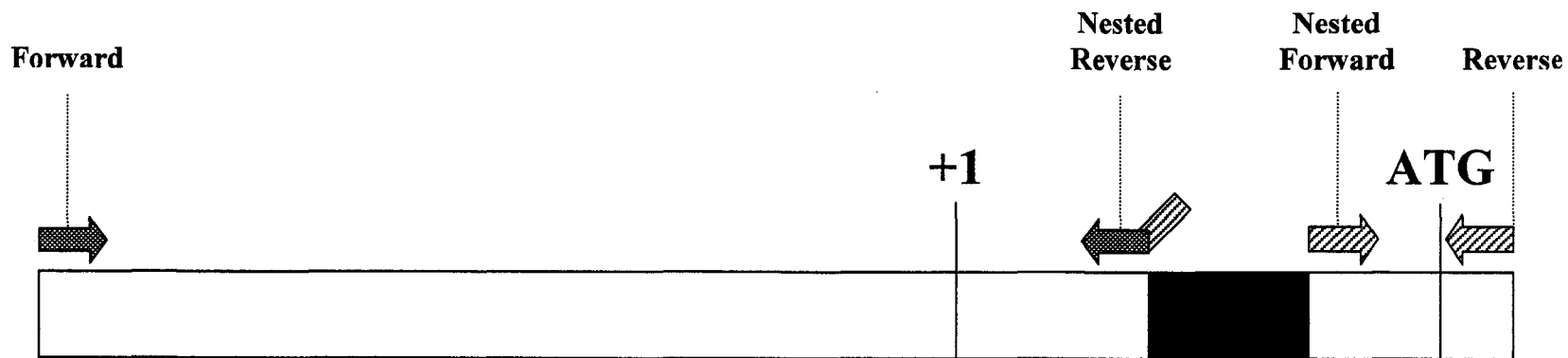


Figure 3. Regulatory region of *CPK6* showing primer placement for overlap extension amplification. Forward primer and nested reverse primer are shaded gray. Reverse primer, forward nested primer and the tail on the nested reverse primer that matches the sequence of the nested reverse primer are striped. The 5' UTR intron is black. +1 is the transcription start site and ATG is the translation start site. Not to scale.

Forward:

5'-ACAAGTTTGTACAAAAAAGCAGGCTGAGGACAGGCATTCCGTGCA-3'

Reverse:

5'-ACCACTTTGTACAAGAAAGCTGGGTGCTTGAAGAAAGTCGAAAGACC-3'

Gateway cloning was done using a Gateway cloning kit (Invitrogen) according to the manufacturer's standard protocol. Clones were checked for correctness by PCR and sequencing.

Physcomitrella Growth Conditions

The *Physcomitrella* Gransden isolate obtained from Dr. David Cove at University of Leeds was grown on PpNH₄ growth medium (Marella et al., 2006) overlaid with cellophane discs (A.A. Packaging Limited). Plates were kept at 25°C with a 12-hr photoperiod. Moss was subcultured by homogenizing tissue in sterile water with a PowerGen 125 grinder equipped with Omni disposable generator probes (Fisher Scientific) and plating the homogenate onto fresh PpNH₄ medium overlaid with cellophane.

Physcomitrella Transformation

The most commonly used method for moss transformation was biolistics, also known as particle bombardment. Target tissue was prepared by homogenizing and plating moss onto PpNH₄ medium lacking di-ammonium tartrate and overlaid with cellophane discs. These plates were incubated for 4-5 days at 25°C with a 12-hr photoperiod.

Three mg of tungsten particles (Sylvania GTE) were coated with 12 µg of plasmid and loaded onto macrocarriers (Bio-Rad) for bombardment according to Bio-Rad's standard protocol. Bombardment was done using a Biolistics Particle Delivery System (Bio-Rad Model # PDS-1000/He) and a 1350 psi rupture disc (Bio-Rad).

Bombarded moss tissue was grown on the same plates for 3 days, then homogenized and plated on PpNH₄ plates containing 15 µg/ml hygromycin B (Roche). Transformed plant cells produced visible protonema after ten day's growth at 25°C in a 12-hr photoperiod.

A second method used for moss transformation was protoplast transformation. Moss tissue was homogenized and plated on PpNH₄ medium overlaid with cellophane discs and grown for 4-5 days at 25°C. Moss tissue was then used for protoplast transformation (Schaefer et al., 1991). Protoplasts were plated and allowed to regenerate cell walls on PpNH₄ for 4-5 days at 25°C. Regenerated plant cells were moved to PpNH₄ medium containing 15µg/ml hygromycin B. Resistant plantlets became visible to the eye after ~15 days.

Selecting for Stable Transformants

Initial transformants obtained from either biolistics or protoplast transformation were removed from PpNH₄ medium containing 15 µg/ml hygromycin and transferred to PpNH₄ medium without selection for 12-14 days at 25°C with a 12-hr photoperiod. Transformants were transferred back to PpNH₄ medium containing 15 µg/ml hygromycin. Plants that were only transiently transformed either died or grew at a slower

rate. Stably transformed plants grew at the same rate in the presence or absence of hygromycin.

Histochemical GUS Assay

Stable transformants were grown to the leafy gametophore stage on PpNH₄ growth medium (~3 weeks). Plant tissue was incubated in GUS substrate solution [0.5 mg/ml X-Gluc, 0.1 M sodium phosphate, pH 7.0, 1 mM potassium ferricyanide, 1 mM potassium ferrocyanide, 10 mM EDTA, 4% Triton X-100] at 37°C. Plants were destained in 70% ethanol for 30 min. Digital images were obtained using a Q-color 3 camera attached to an SZX9 dissecting zoom microscope (Olympus).

CHAPTER III

RESULTS

Identification of CDPKs

I have identified a total of 16 CDPKs in *Physcomitrella*. Initially, eight CDPKs were identified from paired 5' and 3' EST sequences available on the PHYSCObase website (<http://moss.nibb.ac.jp>). Based on the EST data, cDNA sequences for these eight CDPKs were amplified by PCR, cloned, and sequenced. In order to identify introns, genomic copies of each gene were also amplified, cloned and sequenced (Figure 4). Eight additional CDPKs were identified by compiling raw sequence data that showed sequence similarity to the previously identified CDPKs. The raw sequence data was produced by the *Physcomitrella* sequencing project conducted by the DOE Joint Genome Institute and available on the NCBI website (<http://www.ncbi.nlm.nih.gov/>). Sequences from the NCBI website were also compared to the original eight CDPK sequences to confirm that they were correct since they were generated by PCR. This analysis revealed only one nucleotide difference among all eight genes.

The amino acid sequences for each CDPK were predicted using the EditSeq program of the DNASTar package (Lasergene). The amino acid sequences for each CDPK domain were aligned using MegAlign of the DNASTar package (Lasergene; Figures 5-8). Distance matrices were calculated using MegAlign based on the alignments of each domain (Tables 1-4). A distance matrix was also calculated based on the amino acid sequence alignment of the full-length CDPK proteins (Table 5). The distance matrices revealed that, on average, the kinase domain is the most conserved of the four



Figure 4. Intron-exon structure of *Physcomitrella* CDPKs for which cDNA sequences were amplified. Untranslated regions are shaded yellow, coding regions are shaded blue, and introns are represented by black lines. CDPK gene names are listed to the left of each structure.

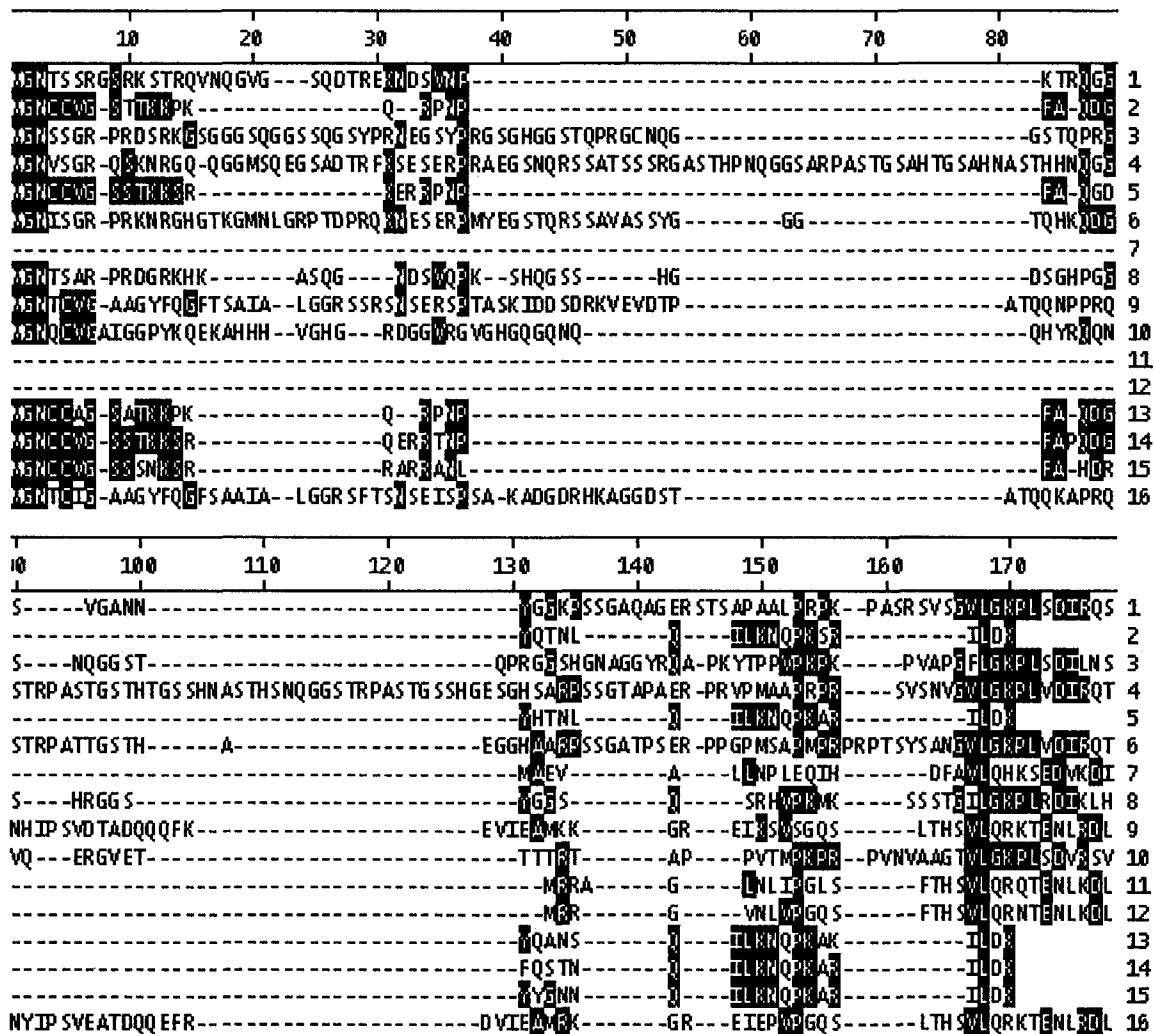


Figure 5. Amino acid alignment of the variable domain of *Physcomitrella* CDPKs. Alignment was done using the ClustalW method in MegAlign from the DNASTar package (Lasergene). Shaded areas represent amino acids identical to the consensus sequence at that position. Dashes indicate sequence gaps. CDPK gene number is listed to the right of each line.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
1		27.9	35.6	50.6	28.9	46.7	13.8	49.2	15.6	26.9	10.3	10.7	30.2	23.9	22.2	14.3	1
2	174.2		23.3	23.3	81.4	25.6	11.1	30.0	20.9	31.0	18.8	12.5	86.0	79.1	67.4	18.6	2
3	129.7	210.0		26.4	20.0	27.9	10.3	54.5	13.1	26.7	6.9	10.7	25.6	21.7	22.2	15.3	3
4	78.3	210.0	186.3		26.7	68.1	10.3	32.9	11.0	27.6	13.8	14.3	23.3	23.9	20.0	12.0	4
5	167.3	21.4	246.0	183.8		22.2	11.1	28.6	20.0	29.5	18.8	12.5	74.4	80.0	73.3	17.8	5
6	89.2	193.1	174.1	41.5	220.0		13.8	35.1	15.4	25.6	13.8	14.3	25.6	23.9	20.0	13.6	6
7	360.0	451.0	493.0	493.0	451.0	360.0		10.7	24.1	17.9	30.8	26.9	11.1	11.1	11.1	27.6	7
8	81.9	159.9	68.5	143.2	169.5	132.3	471.0		13.2	21.3	11.1	14.8	35.0	25.6	26.2	14.5	8
9	317.0	234.0	377.0	456.0	246.0	321.0	202.0	377.0		14.0	60.7	67.9	20.9	21.7	17.8	72.5	9
10	181.8	154.1	183.2	176.6	162.9	193.3	276.0	230.0	354.0		14.3	14.3	23.8	26.7	27.3	16.5	10
11	493.0	262.0	1000.0	360.0	262.0	360.0	155.2	451.0	55.1	347.0		85.7	18.8	18.8	18.8	67.9	11
12	471.0	399.0	471.0	347.0	399.0	347.0	181.8	332.0	41.9	347.0	15.9		12.5	12.5	12.5	75.0	12
13	158.5	15.5	193.1	210.0	31.3	193.1	451.0	132.6	234.0	205.0	262.0	399.0		76.7	65.1	20.9	13
14	204.0	24.6	226.0	204.0	23.3	204.0	451.0	193.1	226.0	183.8	262.0	399.0	27.9		73.3	17.4	14
15	220.0	42.6	220.0	246.0	33.0	246.0	451.0	187.8	277.0	179.0	262.0	399.0	46.7	33.0		15.6	15
16	347.0	264.0	321.0	414.0	277.0	365.0	176.6	342.0	34.3	299.0	41.9	30.4	234.0	284.0	317.0		16
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	

Table 1. Percent identity (upper right) and divergence (lower left) between the variable domains of *Physcomitrella* CDPKs as calculated by MegAlign from the DNASTar package (Lasergene). Percent identity is the percentage of identical amino acids between two sequences. Percent divergence is derived from the phylogeny reconstructed by the amino acid alignment. Note that percent identity and divergence do not usually equal 100%. CDPK gene numbers are listed around the perimeter of table. The lowest percent identity is shaded red, while the highest percent identity is shaded red and bolded. Sequences that are the most diverged from each other have their percent divergence shaded blue, while the two genes that are the least diverged have their percent divergence shaded blue and bolded.

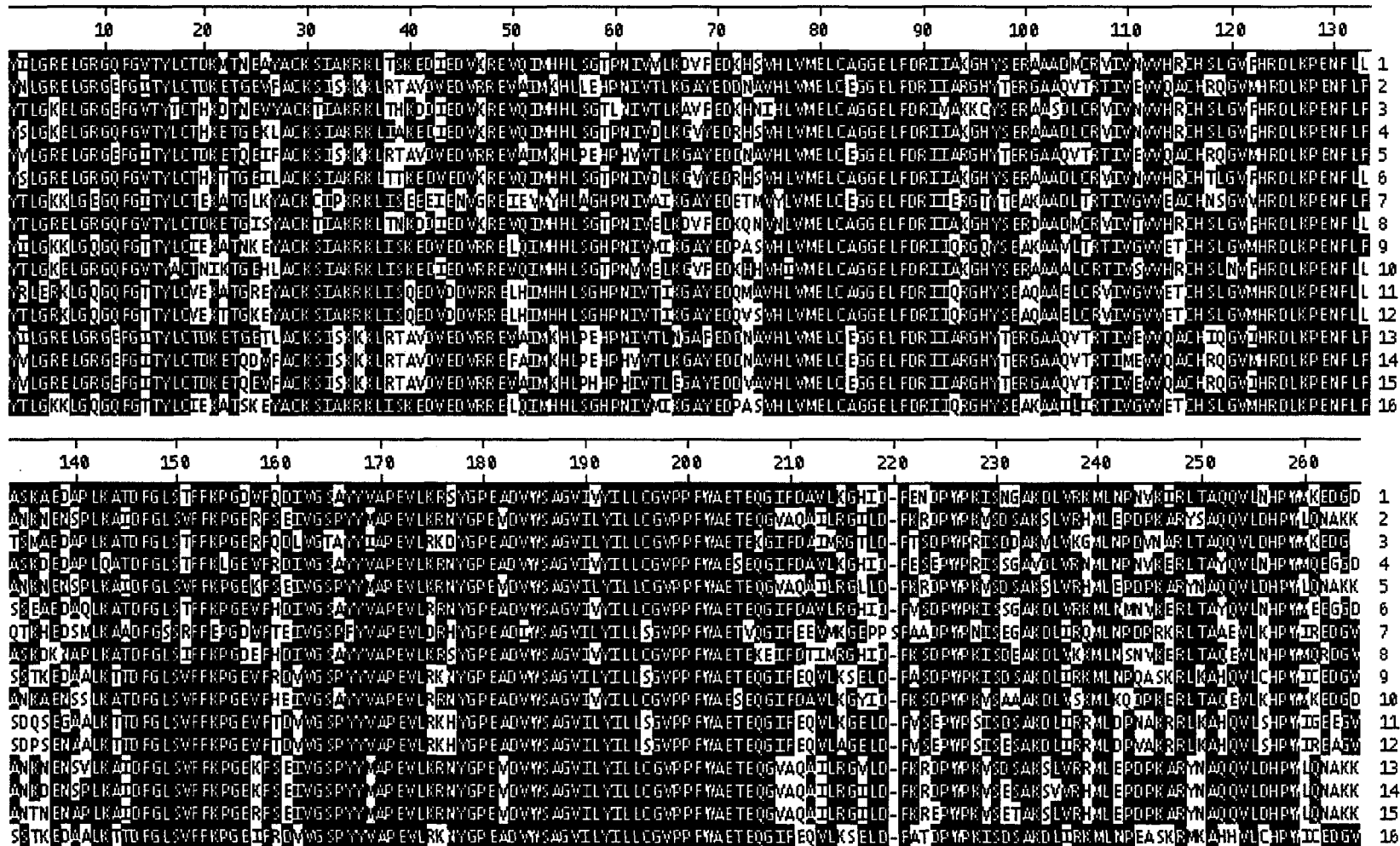


Figure 6. Amino acid alignment of the kinase domain of *Physcomitrella* CDPKs. Alignment was done using the ClustalW method in MegAlign from the DNASTar package (Lasergene). Shaded areas represent amino acids identical to the consensus sequence at that position. CDPK gene number is listed to the right of each line.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
1		68.2	81.0	87.9	67.4	88.6	65.9	86.0	73.5	82.6	71.6	71.2	68.2	65.9	67.0	72.3	1
2	41.3		65.8	67.8	96.6	68.6	63.3	67.0	66.3	68.9	66.3	66.3	95.1	95.1	94.7	66.3	2
3	22.0	45.5		77.9	64.3	78.7	60.5	78.7	68.8	74.5	66.2	66.2	64.6	62.7	64.6	68.1	3
4	13.3	42.0	26.2		66.7	90.5	65.5	80.7	72.7	81.4	71.6	71.2	67.0	65.5	67.0	72.0	4
5	42.6	3.5	48.3	43.9		67.8	62.1	65.9	65.5	68.6	65.2	65.2	95.1	97.0	95.1	65.5	5
6	12.4	40.7	25.1	10.1	42.0		64.4	81.8	73.9	82.2	73.5	73.9	68.2	65.9	67.4	73.1	6
7	45.3	50.2	55.6	46.0	52.3	48.0		62.9	72.7	65.2	69.7	70.1	62.5	61.7	61.4	71.6	7
8	15.6	43.3	25.1	22.4	45.3	20.9	50.9		67.4	78.8	67.4	67.4	66.7	64.8	65.2	67.8	8
9	32.7	44.6	40.2	33.9	46.0	32.2	33.9	42.6		70.5	86.4	86.0	65.5	64.4	64.8	95.8	9
10	19.9	40.0	31.2	21.4	40.7	20.4	46.6	25.0	37.5		68.9	69.3	69.3	67.0	66.7	70.5	10
11	35.7	44.6	44.8	35.7	46.6	32.7	38.8	42.6	15.1	40.0		95.5	65.5	64.0	64.8	84.8	11
12	36.3	44.6	44.8	36.3	46.6	32.2	38.1	42.6	15.6	39.4	4.7		65.5	64.8	65.9	84.8	12
13	41.3	5.1	47.6	43.3	5.1	41.3	51.6	43.9	46.0	39.4	46.0	46.0		92.8	92.8	65.2	13
14	45.3	5.1	51.1	46.0	3.1	45.3	53.1	47.3	48.0	43.3	48.7	47.3	7.6		94.3	64.4	14
15	43.3	5.5	47.6	43.3	5.1	42.6	53.8	46.6	47.3	43.9	47.3	45.3	7.6	5.9		64.8	15
16	34.5	44.6	41.5	35.1	46.0	33.3	35.7	42.0	4.3	37.5	17.0	17.0	46.6	48.0	47.3		16
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	

Table 2. Percent identity (upper right) and divergence (lower left) between the kinase domains of *Physcomitrella* CDPKs as calculated by MegAlign from the DNASTar package (Lasergene). Percent identity is the percentage of identical amino acids between two sequences. Percent divergence is derived from the phylogeny reconstructed by the amino acid alignment. Note that percent identity and divergence do not usually equal 100%. CDPK gene numbers are listed around the perimeter of table. The lowest percent identity is shaded red, while the highest percent identity is shaded red and bolded. Sequences that are the most diverged from each other have their percent divergence shaded blue, while the two genes that are the least diverged have their percent divergence shaded blue and bolded.

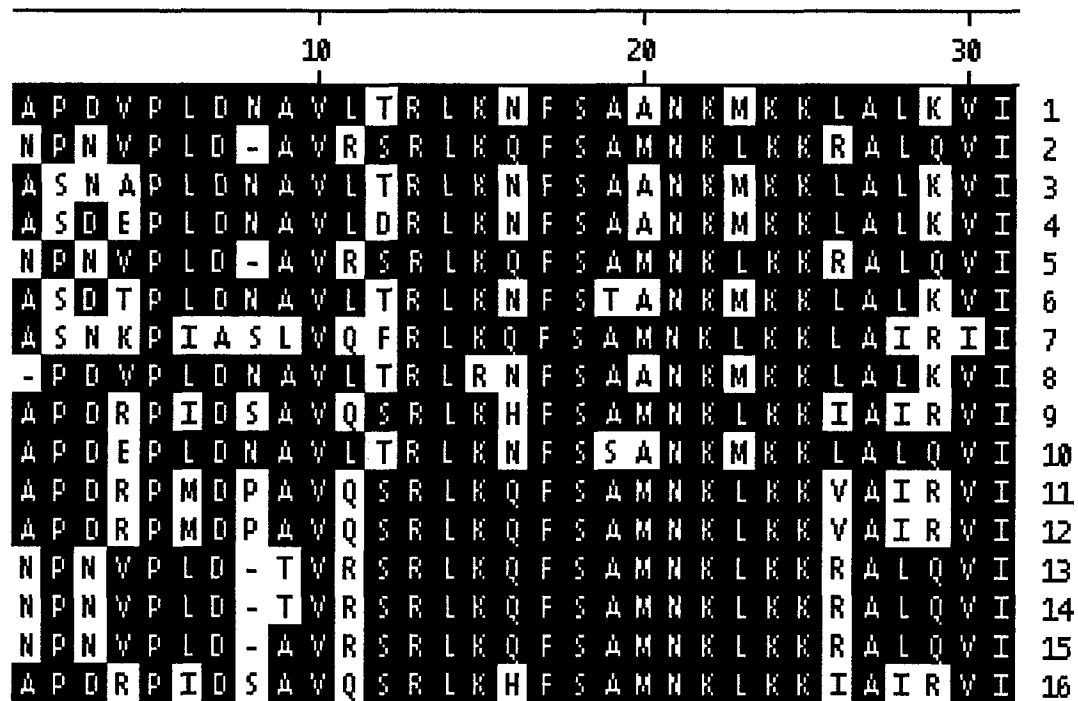


Figure 7. Amino acid alignment of the autoinhibitory domain of *Physcomitrella* CDPKs. Alignment was done using the ClustalW method in MegAlign of the DNASTar package (Lasergene). Shaded areas represent amino acids identical to the consensus sequence at that position. CDPK gene number is listed to the right of each line.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
1		67.7	90.3	90.3	67.7	90.3	51.6	93.5	64.5	90.3	64.5	64.5	64.5	64.5	67.7	64.5	1
2	38.3		66.7	63.3	100.0	60.0	60.0	66.7	70.0	66.7	73.3	73.3	96.7	96.7	100.0	70.0	2
3	10.4	43.9		90.3	64.5	90.3	58.1	83.9	58.1	83.9	58.1	58.1	61.3	61.3	64.5	58.1	3
4	10.4	50.0	10.4		61.3	90.3	54.8	83.9	61.3	87.1	61.3	61.3	58.1	58.1	61.3	61.3	4
5	38.3	0.0	43.9	50.0		60.0	60.0	66.7	70.0	66.7	73.3	73.3	96.7	96.7	100.0	70.0	5
6	10.4	56.6	10.4	10.4	56.6		51.6	83.9	58.1	87.1	58.1	58.1	54.8	54.8	58.1	58.1	6
7	75.7	56.6	60.6	67.8	56.6	75.7		45.2	71.0	48.4	67.7	67.7	58.1	58.1	58.1	71.0	7
8	3.4	40.0	14.7	14.7	40.0	14.7	89.2		60.0	86.7	60.0	60.0	63.3	63.3	66.7	60.0	8
9	47.8	38.3	60.6	54.0	38.3	60.6	36.7	56.6		61.3	87.1	87.1	64.5	64.5	67.7	100.0	9
10	10.4	43.9	18.2	14.2	43.9	14.2	84.3	14.7	54.0		61.3	61.3	61.3	61.3	64.5	61.3	10
11	47.8	33.0	60.6	54.0	33.0	60.6	42.1	56.6	14.2	54.0		100.0	67.7	67.7	71.0	87.1	11
12	47.8	33.0	60.6	54.0	33.0	60.6	42.1	56.6	14.2	54.0	0.0		67.7	67.7	71.0	87.1	12
13	43.9	3.4	50.0	56.6	3.4	63.7	56.6	46.0	43.9	50.0	38.3	38.3		100.0	96.7	66.7	13
14	43.9	3.4	50.0	56.6	3.4	63.7	56.6	46.0	43.9	50.0	38.3	38.3	0.0		96.7	66.7	14
15	38.3	0.0	43.9	50.0	0.0	56.6	56.6	40.0	38.3	43.9	33.0	33.0	3.4	3.4		70.0	15
16	47.8	38.3	60.6	54.0	38.3	60.6	36.7	56.6	0.0	54.0	14.2	14.2	43.9	43.9	38.3		16
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	

Table 3. Percent identity (upper right) and divergence (lower left) between the autoinhibitory domains of *Physcomitrella* CDPKs as calculated by MegAlign from the DNASTar package (Lasergene). Percent identity is the percentage of identical amino acids between two sequences. Percent divergence is derived from the phylogeny reconstructed by the amino acid alignment. Note that percent identity and divergence do not usually equal 100%. CDPK gene numbers are listed around the perimeter of table. The lowest percent identity is shaded red, while the highest percent identity is shaded red and bolded. Sequences that are the most diverged from each other have their percent divergence shaded blue, while the two genes that are the least diverged have their percent divergence shaded blue and bolded.

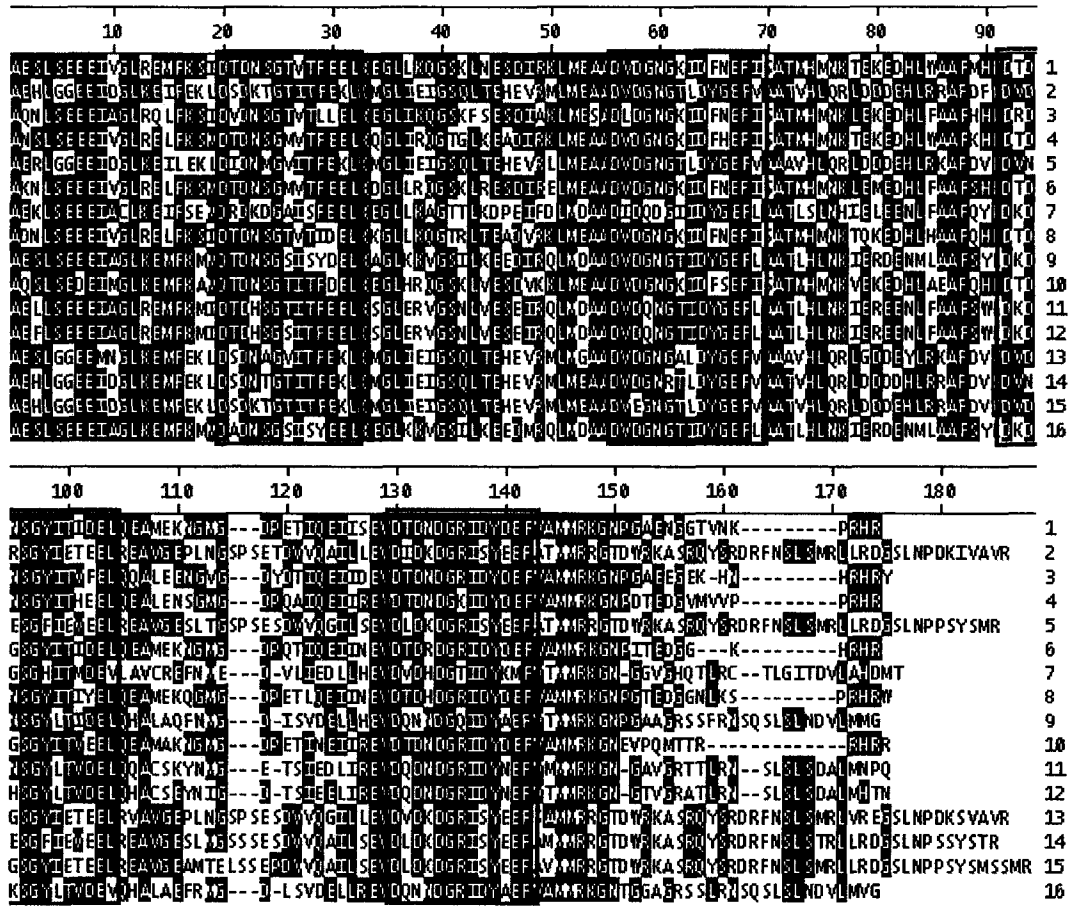


Figure 8. Amino acid alignment of the calmodulin-like domain of *Physcomitrella* CDPKs. Alignment was done using the ClustalW method in MegAlign from the DNAStar package (Lasergene). Shaded areas represent amino acids that match the consensus sequence at that position. CDPK gene number is listed to the right of figure. Ca²⁺ binding domains are boxed in pink.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
1		46.9	79.5	82.1	45.1	86.8	49.4	84.6	59.0	79.7	63.1	61.9	45.7	48.1	47.5	57.1	1
2	88.5		43.8	45.7	85.9	47.2	41.4	45.4	45.3	47.2	48.2	48.2	87.1	87.6	88.7	45.9	2
3	24.0	98.0		73.3	42.0	77.4	45.0	74.1	53.8	71.1	57.5	57.5	42.0	43.8	42.6	53.8	3
4	20.5	92.2	33.0		44.4	82.4	46.9	79.6	55.9	74.7	56.9	58.1	44.4	46.3	45.1	54.7	4
5	94.1	15.6	104.3	96.0		45.3	39.1	42.9	43.5	45.3	45.8	46.4	83.2	90.3	87.0	44.7	5
6	14.6	87.7	27.0	20.1	93.4		49.0	83.0	56.3	77.1	60.5	61.8	45.9	47.2	47.2	55.1	6
7	81.5	106.3	94.3	88.6	115.1	82.4		46.0	58.1	47.1	59.5	60.7	39.1	38.5	40.8	59.3	7
8	17.3	93.0	31.8	23.8	101.0	19.3	91.3		54.0	74.2	55.3	55.3	42.9	45.4	44.8	52.2	8
9	58.6	93.4	70.4	65.4	99.0	64.4	60.6	69.7		56.1	73.1	74.3	44.7	44.1	44.1	90.0	9
10	23.7	87.7	36.5	30.9	93.4	27.4	87.8	31.6	65.0		59.9	59.2	46.5	48.4	47.8	57.3	10
11	50.4	84.7	61.8	63.2	91.7	55.5	57.5	66.8	33.4	56.8		91.7	46.4	48.2	47.6	72.5	11
12	52.8	84.7	61.8	60.5	89.9	53.0	55.1	66.8	31.6	58.1	8.9		46.4	46.4	47.0	74.9	12
13	92.2	14.2	104.3	96.0	19.0	91.5	115.1	101.0	95.2	89.6	89.9	89.9		81.1	81.2	46.5	13
14	84.9	13.6	98.0	90.3	10.4	87.7	117.5	93.0	97.1	84.1	84.7	89.9	21.9		89.7	45.9	14
15	86.7	12.3	102.2	94.1	14.3	87.7	108.4	95.0	97.1	85.9	86.4	88.2	21.7	11.1		45.3	15
16	62.6	91.6	70.4	68.2	95.2	67.3	58.0	74.2	10.8	62.2	34.3	30.7	89.8	91.6	93.4		16
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	

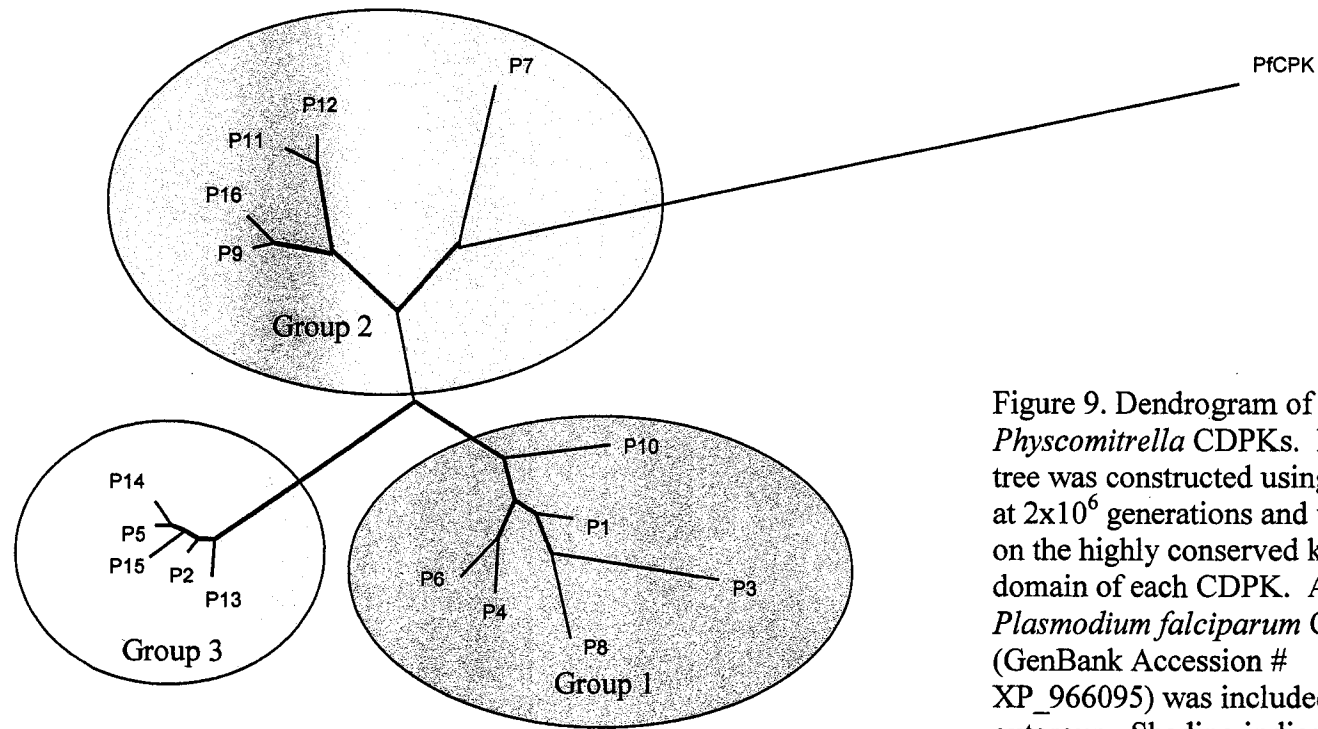
Table 4. Percent identity (upper right) and divergence (lower left) between the calmodulin-like domains of *Physcomitrella* CDPKs as calculated by MegAlign from the DNASTar package (Lasergene). Percent identity is the percentage of identical amino acids between two sequences. Percent divergence is derived from the phylogeny reconstructed by the amino acid alignment. Note that percent identity and divergence do not usually equal 100%. CDPK gene numbers are listed around the perimeter of table. The lowest percent identity is shaded red, while the highest percent identity is shaded red and bolded. Sequences that are the most diverged from each other have their percent divergence shaded blue, while the two genes that are the least diverged have their percent divergence shaded blue and bolded.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
1		57.5	74.4	78.7	56.5	80.0	57.1	79.9	58.3	73.4	64.9	64.4	56.7	55.8	56.1	58.1	1
2	61.7		55.3	56.2	91.8	56.8	52.4	56.1	55.7	58.3	56.9	57.6	91.6	90.6	90.6	55.9	2
3	31.3	66.8		69.3	53.4	70.2	52.5	73.7	53.0	66.5	59.2	59.5	53.6	53.3	53.6	52.9	3
4	25.1	64.6	39.4		55.2	82.9	55.7	75.4	56.1	70.1	62.7	63.0	55.2	55.0	55.8	55.2	4
5	64.0	8.7	71.2	66.9		55.5	50.8	54.4	54.1	56.9	55.4	56.1	89.3	93.1	90.6	54.5	5
6	23.4	63.4	37.9	19.5	66.2		55.5	76.6	56.5	71.0	64.4	65.2	56.0	54.6	55.9	55.5	6
7	62.8	73.5	73.4	65.8	77.7	66.3		53.4	65.0	55.7	63.7	64.7	51.2	50.5	51.0	65.0	7
8	23.4	65.0	32.4	29.8	68.7	28.2	71.1		55.7	70.3	59.7	60.0	55.1	54.1	54.2	55.3	8
9	60.0	65.8	72.3	64.8	69.5	64.1	47.0	65.7		57.5	79.8	81.0	54.9	54.0	53.9	89.9	9
10	32.8	60.1	44.3	38.1	63.2	36.7	65.7	37.7	61.9		61.8	62.3	57.7	56.5	56.2	57.7	10
11	47.0	63.2	58.3	51.2	66.5	47.9	49.2	57.2	23.5	53.0		93.5	55.6	55.5	55.8	79.2	11
12	47.9	61.6	57.6	50.6	64.8	46.5	47.5	56.5	21.9	51.9	6.8		56.4	56.0	56.8	81.0	12
13	63.5	8.9	70.6	66.9	11.6	65.2	76.6	67.3	67.6	61.4	66.0	64.4		87.2	86.6	55.5	13
14	65.6	10.1	71.4	67.4	7.2	68.3	78.5	69.5	69.7	63.9	66.3	65.1	14.1		90.8	54.6	14
15	64.9	10.0	70.7	65.5	10.0	65.3	77.2	69.2	69.9	64.6	65.5	63.4	14.8	9.8		54.3	15
16	60.6	65.3	72.5	67.0	68.5	66.2	47.0	66.8	10.9	61.3	24.4	21.9	66.2	68.3	69.0		16
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	

Table 5. Percent identity (upper right) and divergence (lower left) between the full-length amino acid sequences of *Physcomitrella* CDPKs as calculated by MegAlign from the DNASTar package (Lasergene). Percent identity is the percentage of identical amino acids between two sequences. Percent divergence is derived from the phylogeny reconstructed by the amino acid alignment. Note that percent identity and divergence do not usually equal 100%. CDPK gene numbers are listed around the perimeter of table. The lowest percent identity is shaded red, while the highest percent identity is shaded red and bolded. Sequences that are the most diverged from each other have their percent divergence shaded blue, while the two genes that are the least diverged have their percent divergence shaded blue and bolded.

CDPK domains while the variable domain, as expected, is the least conserved. Amino acid identity for CDPK domains range from 6.9% (*CPK3* and *CPK11*) to 86.0% (*CPK2* and *CPK13*) identity for the variable domain, 60.5% (*CPK3* and *CPK7*) to 97.0% (*CPK5* and *CPK14*) identity for the kinase domain, 45.2% (*CPK7* and *CPK8*) to 100% (*CPK2*, *CPK5* and *CPK15*; *CPK9* and *CPK16*; *CPK11* and *CPK12*; *CPK13* and *CPK14*) identity for the autoinhibitory domain, and 38.5% (*CPK7* and *CPK14*) to 91.7% (*CPK11* and *CPK12*) identity for the calmodulin-like domain. The broad range of sequence identities for each domain indicates that some CDPKs are more closely related to each other than they are to others. In order to better define relationships between the individual CDPKs, Phylogenetic analysis was done.

The well-conserved kinase domain was used to construct a *Physcomitrella* CDPK tree. The analysis showed that *Physcomitrella* CDPKs formed three major groups (Figure 9). This was not unexpected since a similar pattern has been observed in trees created from *Arabidopsis* CDPKs based on the kinase domain amino acid sequence (Hrabak et al., 2003) and *Oryza* CDPKs based on full-length amino acid sequence (Asano et al., 2005). CDPKs in Group 3 (*CPK2*, *CPK5*, *CPK13*, *CPK14* and *CPK15*) have the shortest branch lengths when compared with the other two groups, indicating that this group contains the CDPKs that are the most closely related to each other. The close relationship of CDPKs in Group 3 was supported by results from the distance matrices for the full-length amino acid sequences as well as the individual domains. The percent divergence from the full-length distance matrix between the CDPKs in Group 3 is on average smaller than what is seen in Group 1, ranging from 9.0% to 14.8%, indicating that these CDPKs have diverged more recently than Group 1 CDPKs. CDPKs in Group 1



Bootstrap Support Key

- | | |
|-------------|-------------|
| ■ 1.00 | ■ 0.70-0.79 |
| ■ 0.90-0.99 | ■ 0.60-0.69 |
| ■ 0.80-0.89 | ■ 0.50-0.59 |

Figure 9. Dendrogram of *Physcomitrella* CDPKs. Rooted tree was constructed using MrBayes at 2×10^6 generations and was based on the highly conserved kinase domain of each CDPK. A *Plasmodium falciparum* CDPK (GenBank Accession # XP_966095) was included as an outgroup. Shading indicates the three groups formed by *Physcomitrella* CDPKs. Bootstrap support is color coded.

(*CPK1*, *CPK3*, *CPK4*, *CPK6*, *CPK8* and *CPK10*) and Group 2 (*CPK7*, *CPK9*, *CPK11*, *CPK12* and *CPK16*) on average showed more divergence, ranging from 23.4% to 44.3% and 6.8% to 49.2% respectively. *CPK7* has the greatest branch length relative to all moss CDPKs, which is also reflected in the complete amino acid distance matrix where the percent divergence between *CPK7* and all other CDPKs ranges from 47.0% to 78.5%.

When compared to CDPK sequences from vascular plants, the three groups formed by the *Physcomitrella* CDPKs associated with the three groups formed by the *Arabidopsis* and *Oryza* CDPKs (Figure 10); however, *Arabidopsis* and *Oryza* CDPKs were interspersed amongst each other within each group whereas *Physcomitrella* CDPKs were always on a separate branch. CDPK sequences from two other moss species, *Tortula ruralis* and *Funaria hygrometrica*, associated with *Physcomitrella* CDPKs in Groups 2 and 3 respectively. A CDPK sequence from the liverwort *Marchantia polymorpha* segregated into Group 1, but did not associate with *Physcomitrella* CDPKs any more so than vascular plant CDPKs. A minor group of *Arabidopsis* and *Oryza* CDPKs (A16, A18, A24, O4 and O18) consisted of CDPKs that are more closely related to the CDPK-related kinases (CRKs) in amino acid sequence than they are to other CDPKs (Hrabak et al., 2003). The CRKs are a group of serine/threonine kinases that are similar in protein domain structure to CDPKs, except that they lack Ca²⁺-binding EF-hands (Lindzen and Choi, 1995). Interestingly, the intron/exon structure of this minor CDPK group is much more similar to the CRK group (data not shown), which supports the close phylogenetic relationship of these proteins (Hrabak et al., 2003). CDPK sequences from *Plasmodium falciparum* (protist), *Dunaliella tertiolecta* (green alga) and *Clamydomonas eugametos* (green alga) did not clearly associate with CDPKs from any

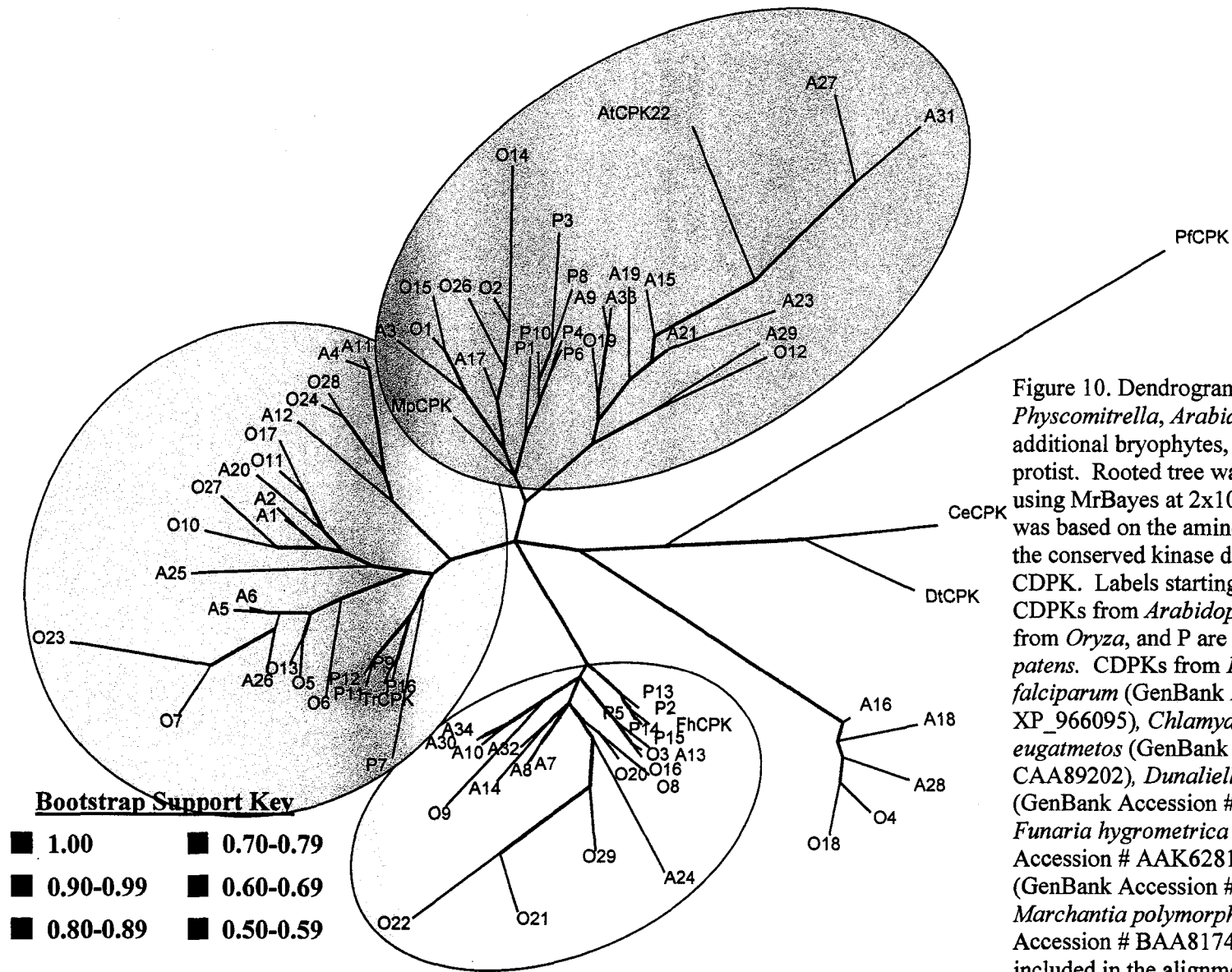


Figure 10. Dendrogram of CDPKs from *Physcomitrella*, *Arabidopsis*, *Oryza*, additional bryophytes, green algae and protist. Rooted tree was constructed using MrBayes at 2×10^6 generations and was based on the amino acid sequence of the conserved kinase domain of each CDPK. Labels starting with A are CDPKs from *Arabidopsis*, O are CDPKs from *Oryza*, and P are CDPKs from *P. patens*. CDPKs from *Plasmodium falciparum* (GenBank Accession # XP_966095), *Chlamydomonas eugatmetos* (GenBank Accession # CAA89202), *Dunaliella tertiolecta* (GenBank Accession # AAF21062), *Funaria hygrometrica* (GenBank Accession # AAK62812), *Tortula ruralis* (GenBank Accession # AAB70706) and *Marchantia polymorpha* (GenBank Accession # BAA81749) are also included in the alignment.

group and functioned as outgroups. Characteristics of *Physcomitrella* CDPK protein sequences are listed in Table 6. The number of amino acids for each CDPK ranges from 491 to 628 and the molecular weight of each CDPK ranges from 55.1 to 64.0 kDa. Both the number of amino acids and molecular weights are within the normal range for CDPKs from *Arabidopsis* and *Oryza* (Hrabak et al., 2003; Asano et al., 2005; Cheng et al., 2002). The isoelectric points (pI) for the CDPKs range from 4.74 to 6.20, where the most acidic CDPKs belong to Group 2. All CDPKs contain four EF hands, the most common number of EF hands in CDPKs (Hrabak, 2000), which happens to be an even number since EF hands bind Ca^{2+} as pairs (Zhang and Yuan, 1998). Characteristics of the CDPK variable domains are listed in Table 7. Variable domains range from 28 to 171 amino acids in length, which accounts for the majority of variability in CDPK length. A plant myristoylation prediction algorithm (available at <http://plantsp.genomics.purdue.edu/plantsp/html/myrist.html>) predicted 11 CDPKs to be myristoylated, indicating a slightly lower percentage of myristoylated CDPKs in *Physcomitrella* than what is observed in *Arabidopsis* (Argyros, 2005). Myristoylation is the attachment of a myristate (14-carbon fatty acid) to an N-terminal glycine of a protein (Towler et al., 1988). Two CDPKs additional CDPKs contain an N-terminal glycine, but were not predicted to be myristoylated. The hydrophobicity of the myristate enables the protein to loosely associate with a membrane which brings the protein in contact with palmitoyl transferases, most of which are integral membrane proteins (Resh, 1999). Palmitoylation is the attachment of a palmitate (16-carbon fatty acid) to a cysteine near a myristoylated N-terminal glycine. The hydrophobicity of the palmitate provides a stronger association of the protein to a membrane than does myristate. Of the 11 CDPKs

Table 6. Characteristics of predicted CDPK proteins from *Physcomitrella*. Phylogenetic group refers to sequence clusters shown in Figure 8. Molecular weight and pI (isoelectric point) were calculated using the ExPasy website (http://ca.expasy.org/tools/pi_tool.html). Number of EF hands was calculated using the SMART website (<http://smart.embl-heidelberg.de/>).

Gene Name	Group Number from Figures 9 and 10	Number of Amino Acids	Molecular Weight (kDa)	pI	Number of EF Hands
<i>CPK1</i>	1	549	60.6	5.86	4
<i>CPK2</i>	3	523	59.3	6.13	4
<i>CPK3</i>	1	567	62.5	6.05	4
<i>CPK4</i>	1	628	68.5	6.07	4
<i>CPK5</i>	3	524	59.3	6.18	4
<i>CPK6</i>	1	578	64.0	6.05	4
<i>CPK7</i>	2	494	55.4	4.74	4
<i>CPK8</i>	1	534	59.7	6.20	4
<i>CPK9</i>	2	575	63.8	5.58	4
<i>CPK10</i>	1	545	60.6	6.17	4
<i>CPK11</i>	2	492	55.3	5.23	4
<i>CPK12</i>	2	491	55.1	5.30	4
<i>CPK13</i>	3	523	58.5	6.10	4
<i>CPK14</i>	3	525	59.6	5.99	4
<i>CPK15</i>	3	527	59.8	6.07	4
<i>CPK16</i>	2	568	63.0	5.74	4

Table 7. Variable domain characteristics of predicted CDPK proteins from *Physcomitrella*. Potential myristoylation sites (bold) were determined by a plant-specific myristoylation algorithm available at <http://plantsp.genomics.purdue.edu/plantsp/html/myrist.html>. Potential palmitoylation sites (underlined) were determined by the presence of a cysteine near a predicted myristoylation site (Resh, 1999). The cysteines in the variable domains of CPK10 and CPK16 are not underlined since both proteins are not predicted to be myristoylated by the plant-specific myristoylation algorithm.

Gene Name	Amino Acids in Variable Domain	Potential Myristoylation Site	Potential Palmitoylation Sites	N-terminal Amino Acids
<i>CPK1</i>	92	YES	0	MGNTSSRG
<i>CPK2</i>	43	YES	2	MGNCCVGS
<i>CPK3</i>	111	YES	0	MGNSSGRP
<i>CPK4</i>	171	YES	0	MGNVSGRQ
<i>CPK5</i>	45	YES	2	MGNCCVGS
<i>CPK6</i>	124	YES	0	MGNISGRP
<i>CPK7</i>	29	NO	0	MAEVALLN
<i>CPK8</i>	78	YES	0	MGNTSARP
<i>CPK9</i>	111	YES	1	MGNTCVGA
<i>CPK10</i>	91	NO	0	MGNQCVGA
<i>CPK11</i>	29	NO	0	MRRAGLNL
<i>CPK12</i>	28	NO	0	MRRGVNLV
<i>CPK13</i>	43	YES	2	MGNCCAGS
<i>CPK14</i>	46	YES	2	MGNCCVGS
<i>CPK15</i>	45	YES	2	MGNCCVGS
<i>CPK16</i>	109	NO	0	MGNTCIGA

predicted to be myristoylated, 6 contained one or more potential palmitoylation sites as well. Further information on myristoylation and palmitoylation will be presented in the discussion.

Kinase Assay

A kinase assay was performed on recombinant CPK4 and CPK6 proteins in order to determine whether they were bona fide protein kinases, and if so, whether their activity was calcium-dependent. To accomplish this, CPK4 and CPK6 proteins were expressed in *E. coli*. Expression clones contained the *CPK4* or *CPK6* open reading frames downstream of both a 6xHis tag and an Xpress epitope. Thus, CPK4 and CPK6 were expressed as fusion proteins with a 36 amino acid sequence containing the 6xHis motif (6 amino acids) and the Xpress epitope (8 amino acids) at the N-terminus (Figure 11). The CDPKs were isolated by affinity purification on nickel-agarose beads using the 6xHis tag. Proteins were separated by SDS-PAGE and detected with a total protein stain (Figure 12A). The major protein products detected in the Coomassie-stained gel were of the expected size for CPK4 and CPK6, 72.6 kDa and 68.2 kDa respectively. Proteins on a duplicate gel were transferred to a PVDF membrane and detected by immuno blotting using anti-Xpress antibody conjugated to horseradish peroxidase. Horseradish peroxidase was detected following incubation with a chemiluminescent substrate by exposing X-ray film (Figure 12B). Western blotting confirmed that the affinity-purified proteins contained an Xpress epitope tag and represented the CDPK fusion proteins.

An *in vitro* kinase assay was performed using the purified CDPK, [γ -³²P] adenosine 5'-triphosphate and histone, which is a general substrate for serine/threonine kinases. Three reactions were set up for each CDPK: one reaction contained CDPK

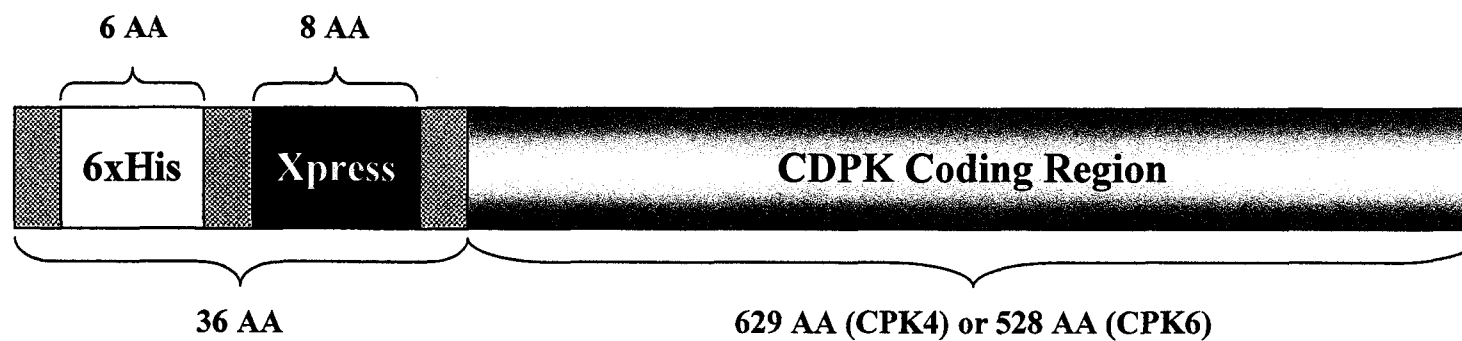


Figure 11. Schematic of the CDPK fusion protein expressed in *E. coli*. The 6xHis tag was used for affinity purification. The Xpress epitope was used for immunodetection of the fusion protein. Figure is not to scale. AA = amino acids.

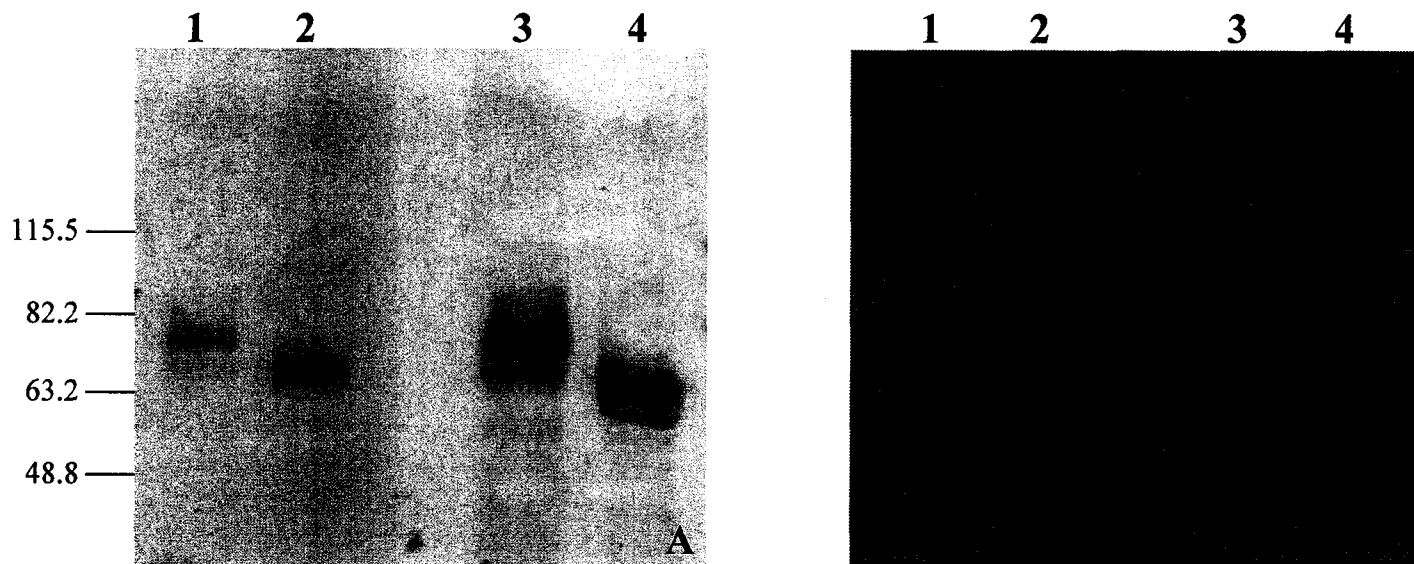


Figure 12. Detection of affinity-purified CPK4 and CPK6. A) Coomassie-stained polyacrylamide gel showing affinity-purified CPK4 fusion protein (72.6 kDa) and CPK6 fusion protein (68.2 kDa). The sizes of protein markers are given in kDa to the left of the figure. B) Chemiluminescent immunodetection of CPK4 and CPK6 fusion proteins. Lanes 1 and 3 contain CPK4 while lanes 2 and 4 contain CPK6. Lanes 1 and 2 contain 1 µg protein, lanes 3 and 4 contain 5 µg protein.

enzyme, histone substrate and Ca^{2+} , one reaction contained CDPK enzyme and histone substrate but lacked Ca^{2+} , and one reaction contained CDPK enzyme and Ca^{2+} but lacked histone substrate. Following the 20 min reaction, half of the sample was spotted on filter paper and washed to remove unincorporated ATP before scintillation counting. The other half was used for SDS-PAGE and subsequent detection using X-ray film. For both CDPKs, a ~15 fold increase of radioactivity was detected by scintillation counting in reactions which contained both histone substrate and calcium as compared to reactions containing histone but lacking Ca^{2+} (Table 8). These results confirm the calcium-stimulated phosphorylation activity of CPK4 and CPK6. Lanes 1 and 4 (Figure 13), which contain CDPK, histone and Ca^{2+} , show strong phosphorylation of the histone substrate protein, which appeared as multiple bands on the polyacrylamide gel as expected due the highly basic nature of the protein. Lanes 2 and 5, which contain only CDPK and histone but lack Ca^{2+} , showed less intense phosphorylation of the histone, indicating that less kinase activity occurred in the reactions that lack Ca^{2+} . Reactions in lanes 3 and 6, which contained Ca^{2+} but lacked histone, showed that CDPK autophosphorylation had occurred. This was not unexpected as CDPK autophosphorylation activity has previously been observed in CDPKs from *Arabidopsis* (Hrabak et al., 1996; Hegeman et al., 2006), *Oryza* (Abo-el-Saad and Wu, 1995), *Nicotiana tabacum* (Iwata et al., 1998) and *Glycine max* (Putnam-Evans et al., 1990).

Table 8. Scintillation counts after *in vitro* kinase assays. Results are shown as the average counts per minute over a course of three minutes. The blank reaction was a filter paper that was not spotted with any sample but was processed in the same manner as the filter papers spotted with the kinase reaction components in order to detect the level of radiation retained on the filter papers after washing.

Reaction Contents	Raw Scintillation Counts	Counts Corrected for Background	Fold Stimulation by Ca ²⁺
Blank	13,891	-	-
<i>CPK4</i> + Ca ²⁺ + histone	379,462	365,571	16.3
<i>CPK4</i> + histone	36,266	22,373	-
<i>CPK4</i> + Ca ²⁺	23,636	9745	-
<i>CPK6</i> + Ca ²⁺ + histone	476,888	462,996	14.3
<i>CPK6</i> + histone	46,273	32,382	-
<i>CPK6</i> + Ca ²⁺	48,474	34,584	-

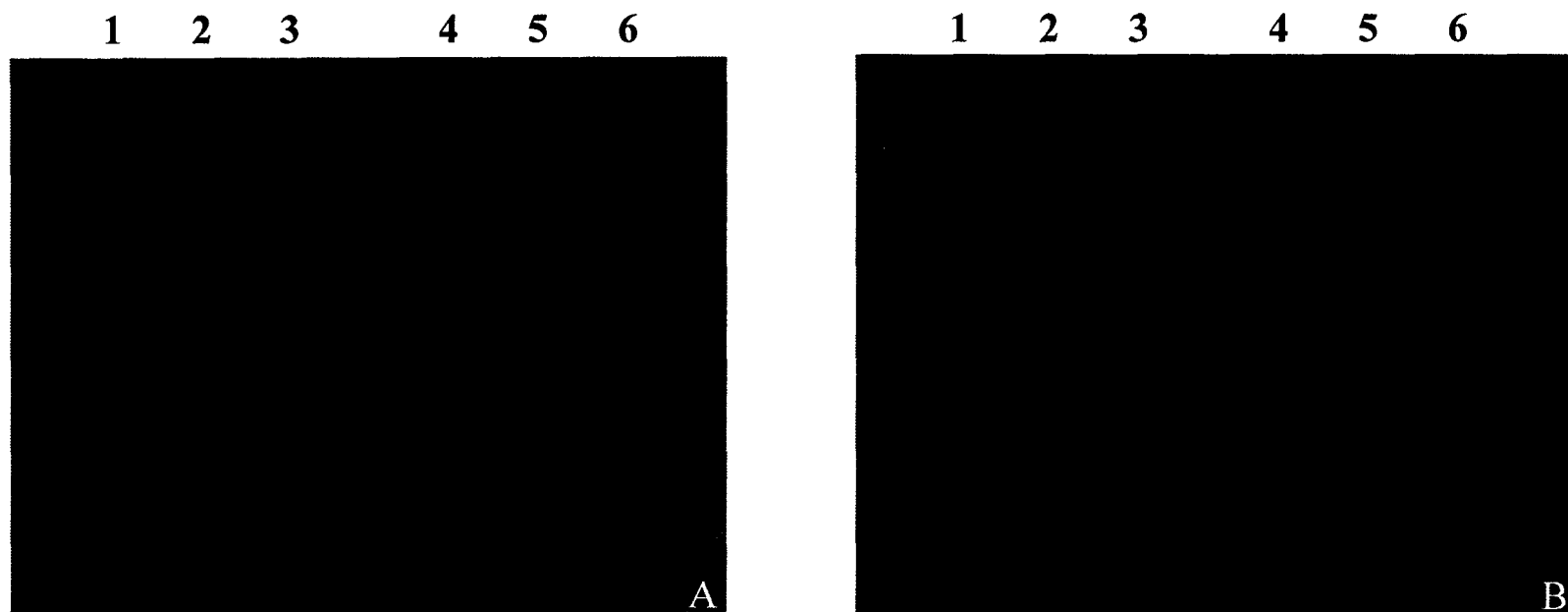


Figure 13. Autoradiography after kinase assays. A) X-ray film after 2 ½ hour exposure to polyacrylamide gel containing kinase reactions. B) X-ray film after overnight exposure to polyacrylamide gel containing kinase reactions. Lanes 1-3 contain CPK4 and lanes 4-6 contain CPK6. Lanes 1 and 4 contain histone substrate and Ca^{2+} , lanes 2 and 5 contain histone substrate without Ca^{2+} , lanes 3 and 6 contain Ca^{2+} without histone substrate.

CPK6 Promoter-GUS Expression

Duplicate *CPK6* promoter-*GUS* constructs named pT-GUS-6a and pT-GUS-6b (Figure 14) were made in order to define the expression pattern of *CPK6* in *Physcomitrella*. The constructs were made as independent duplicates in order to confirm the reproducibility of results. These constructs contained 1.8 kb of sequence upstream of the transcription start site, the 1.1 kb 5' UTR containing a ~0.3 kb intron, and 110 bp of coding sequence from *CPK6* fused to the open reading frame of the *GUS* gene. A second set of duplicate *CPK6* promoter-*GUS* constructs named pT-GUS-6a(-) and pT-GUS-6b(-) (Figure 15) were constructed that are identical to pT-GUS-6a and pT-GUS-6b except that they lacked the intron in the 5' UTR. These constructs were used to determine the role of the intron in *CPK6* expression.

One or two stable transformants were obtained for each of the four promoter-*GUS* constructs and are labeled either #1 or #2. Transformants were incubated with *GUS* substrate at different developmental stages. After incubation, chlorophyll was removed by ethanol. Transformants were examined under a microscope for expression in the protonema, and the developing and mature phyllids of the gametophore (Figure 16). The different stages of gametophore development examined included the gametophore bud, young gametophore and more mature gametophore. Plants that were stably transformed with pT-GUS-6a or pT-GUS-6b containing the 5' UTR intron showed strongest *GUS* expression in gametophore buds, immature phyllids, the base of the gametophore at the site of protonemal attachment and protonemal tip cells (Figures 17-19). Strong expression was also seen in the axillary hair cells, though this was not consistent.

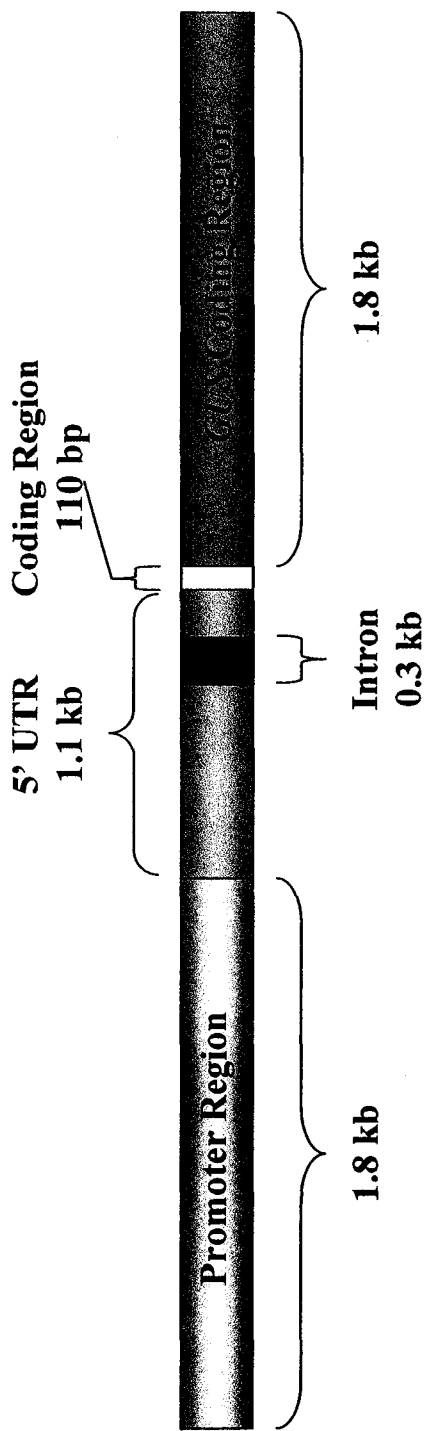


Figure 14. Diagram of *CPK6* promoter-*GUS* cassette in pT-*GUS*-6a and pT-*GUS*-6b.

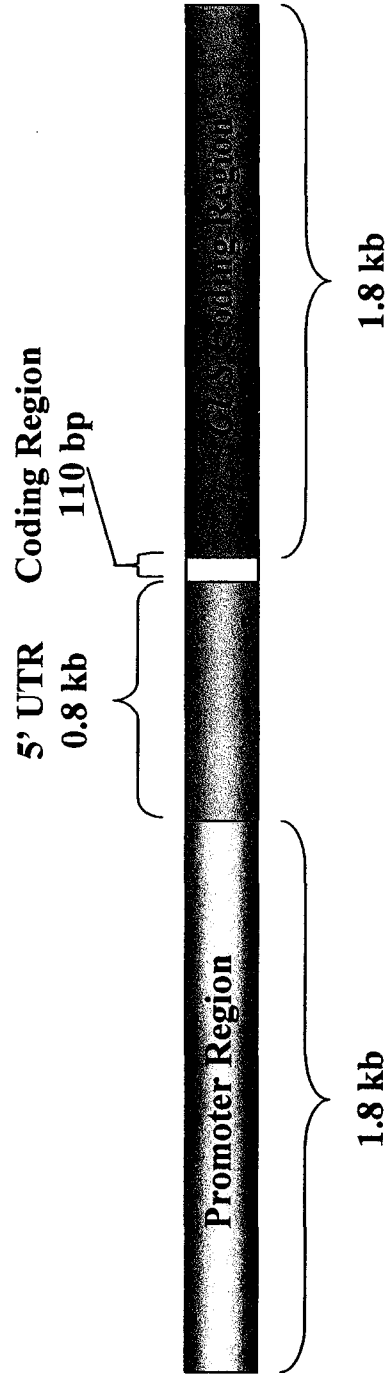


Figure 15. Diagram of *CPK6* promoter-*GUS* cassette in pT-*GUS*-6a(-) and pT-*GUS*-6b(-).

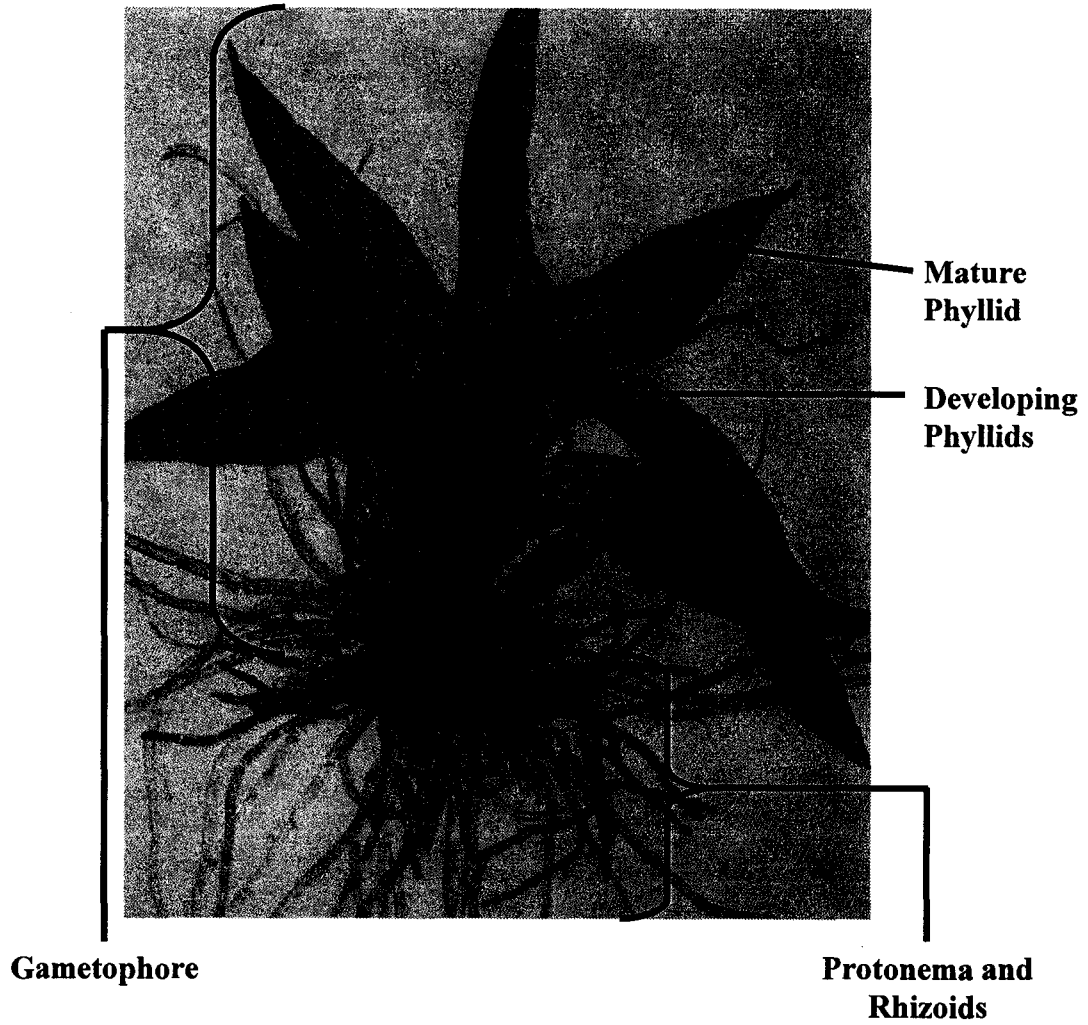


Figure 16. Anatomy of a ~4 week old wild-type *Physcomitrella* plant.

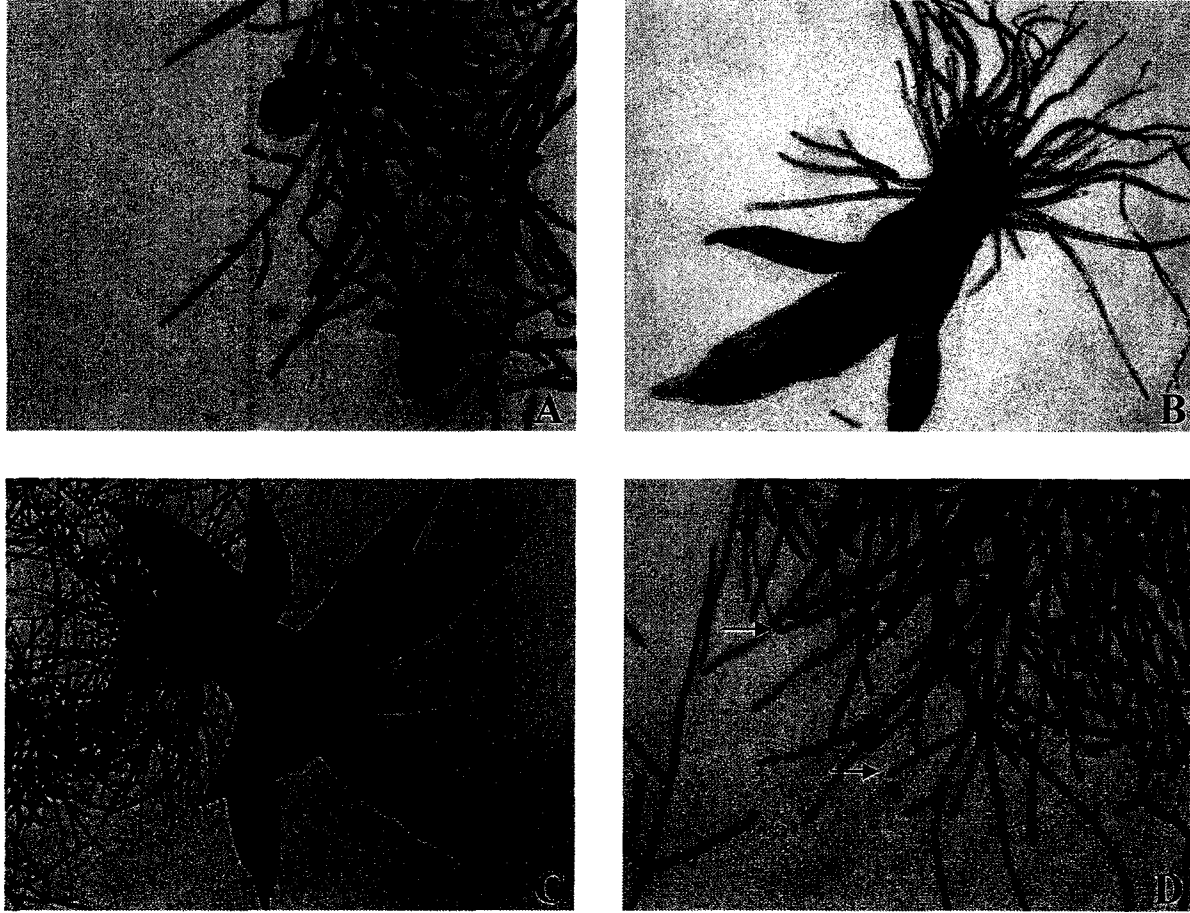


Figure 17. Light micrographs of *Physcomitrella* stably transformed with pT-GUS-6a; transformant #1. Plants were incubated with *GUS* substrate for 5 hours. A) Two gametophore buds showing stronger *GUS* expression than surrounding protonemal tissue. B) Young gametophore showing strongest *GUS* expression in the youngest phyllids and at the site of attachment to protonema. C) Gametophore showing strong overall *GUS* expression. D) Protonemal tissue showing strongest *GUS* expression in certain tip cells indicated by arrows.

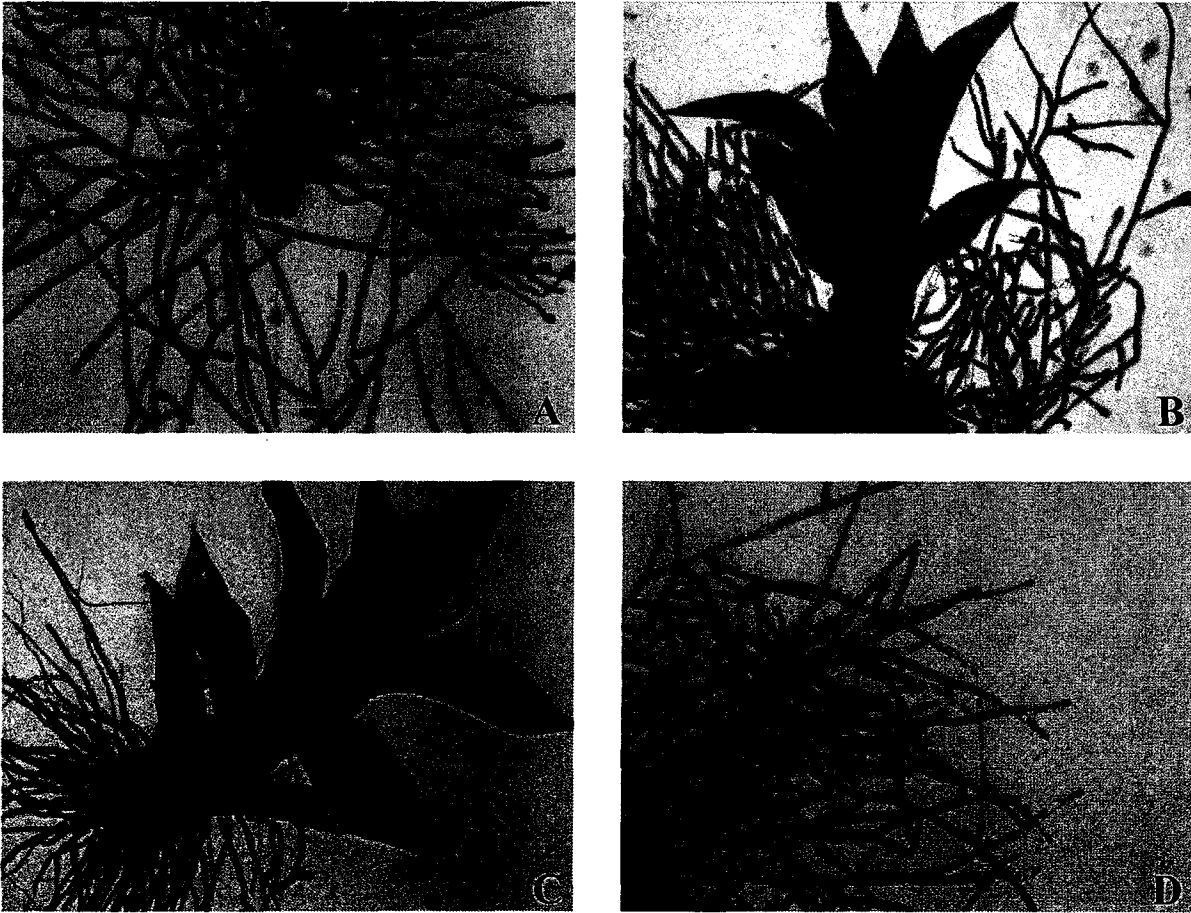


Figure 18. Light micrographs *Physcomitrella* stably transformed with pT-GUS-6a; transformant #2. Plants were incubated with *GUS* substrate for 5 hours. A) A gametophore bud showing stronger *GUS* expression than surrounding protonemal tissue. B) Young gametophore showing strongest expression in the immature phyllids and at the site of attachment to protonema. C) Gametophore showing strongest expression in the immature phyllids and at the site of attachment to the protonema. D) Protonemal tissue showing strongest *GUS* expression in the tip cells.

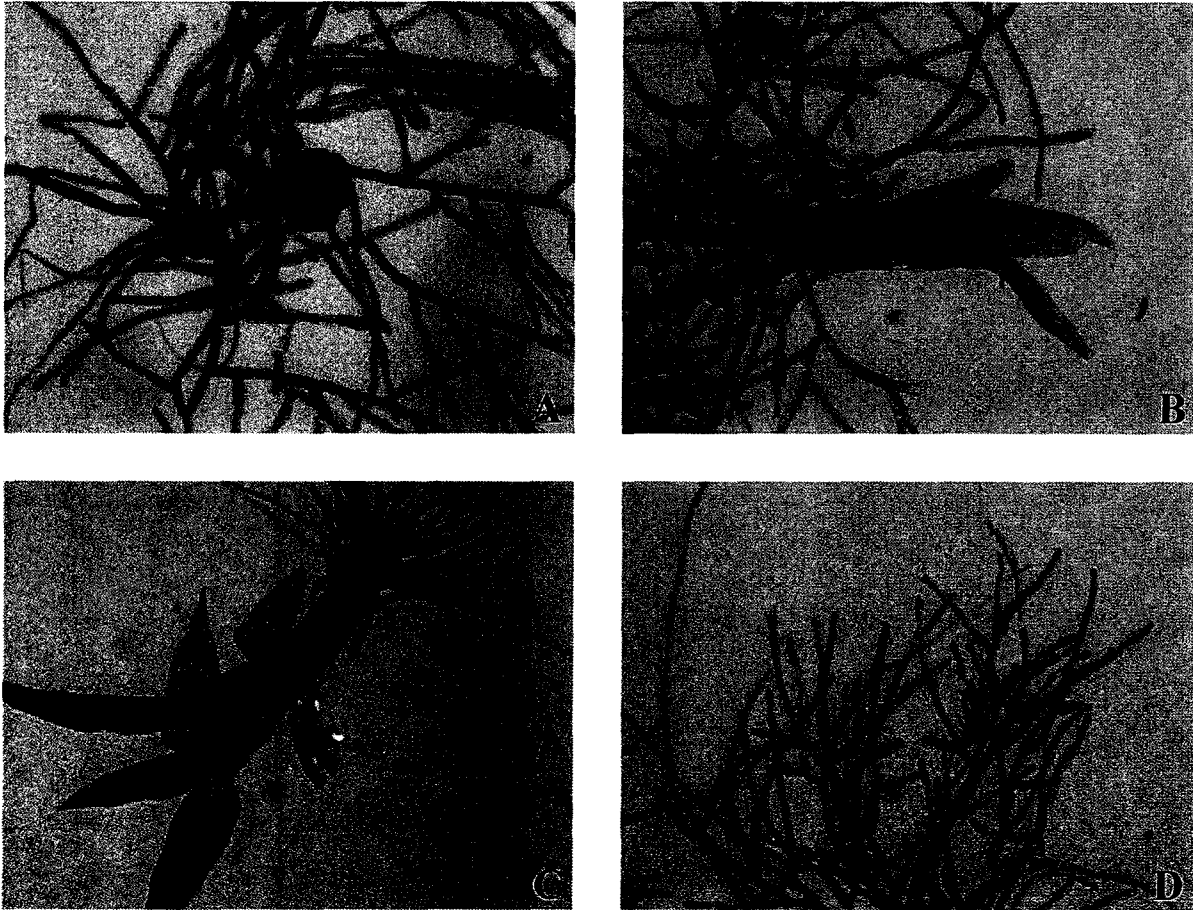


Figure 19. Light micrographs of *Physcomitrella* stably transformed with pT-GUS-6b; transformant #1. Plants were incubated with *GUS* substrate for 5 hours. A) A gametophore bud showing stronger *GUS* expression than surrounding protonemal tissue. B) Young gametophore showing strongest *GUS* expression in the youngest phyllids and at the site of attachment to the protonema. C) Gametophore showing strongest *GUS* expression in the youngest phyllids and at the site of attachment to the protonema. D) Protonemal tissue showing strongest *GUS* expression in the tip cells.

Weaker expression was seen in the mature phyllids and protonemal tissue with the exception of tip cells (Figures 17-19). Overall, these results indicated that *CPK6* is more strongly expressed in areas of the plant where cells are actively dividing, with the exception of the axillary hair cells and gametophore base, and less strongly in areas of the plant that are not actively dividing, i.e. the mature phyllids and protonemal cells that were not tip cells. It should be pointed out that there are some drawbacks to using the *GUS* reporter gene method for detecting expression patterns since the half-life of the *GUS* protein may be different from that of *CPK6*, and it is possible that not 100% of the regulatory elements were included in the *CPK6* promoter-*GUS* construct. However, aside from these drawbacks, it is likely that the *GUS* expression results are close to the actual expression profile of *CPK6*.

Plants that were stably transformed with pT-GUS-6a(-) or pT-GUS-6b(-) lacking the 5' UTR intron showed less *GUS* expression than plants transformed with pT-GUS-6a or pT-GUS-6b. However, the overall pattern of expression remained the same. Strongest *GUS* expression was seen in gametophore buds, immature phyllids, base of the gametophore at the site of protonemal attachment, axillary hair cells of mature gametophores and protonemal tip cells. Little or no expression was seen in mature phyllids and protonemal cells that weren't tip cells (Figures 20-22). These results indicate that the intron in the 5' UTR plays a role in upregulating the overall level of *CPK6* expression, but doesn't change when or where *CPK6* is expressed in the plant

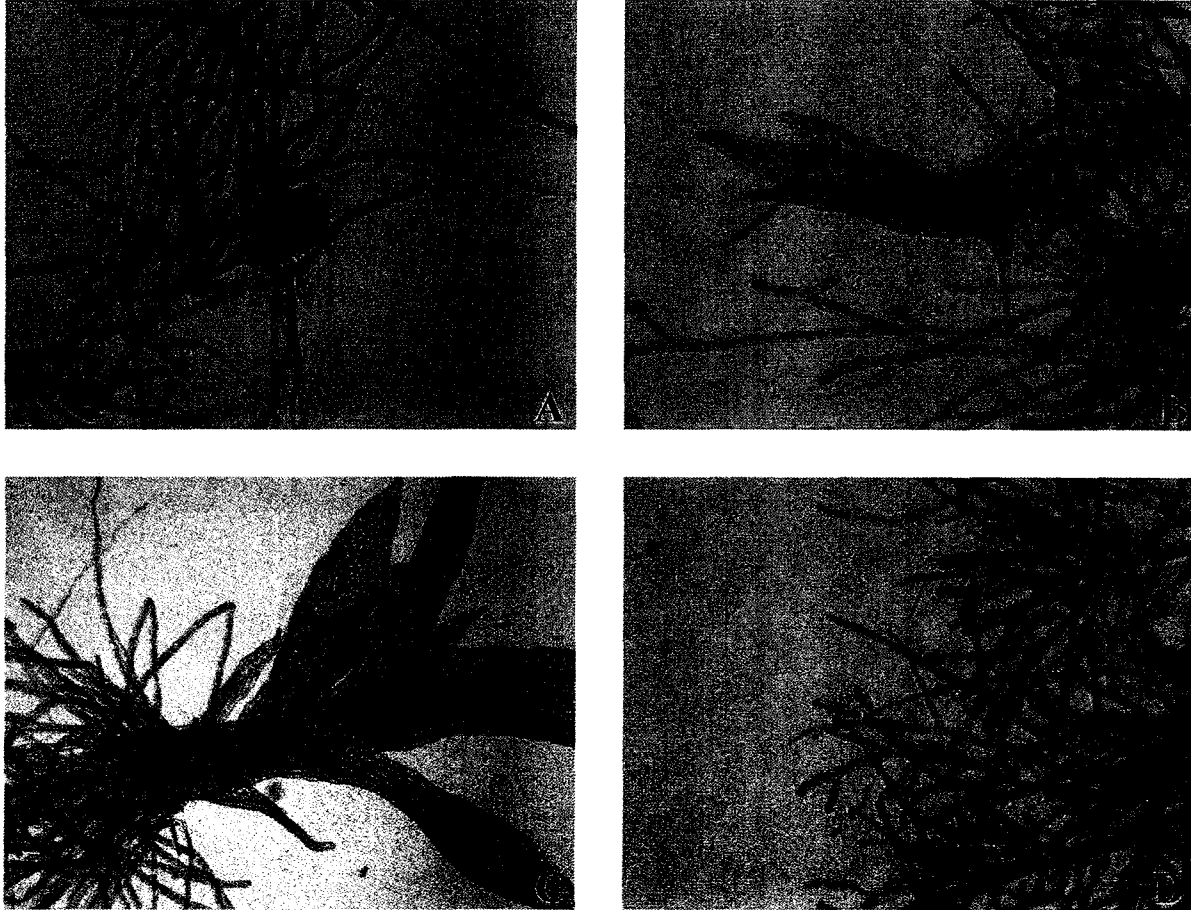


Figure 20. Light micrographs of *Physcomitrella* stably transformed with pT-GUS-6a(-); transformant #1. Plants were incubated with *GUS* substrate for 5 hours. A) A gametophore bud showing weak *GUS* expression. B) Young gametophore showing weak *GUS* expression in the youngest phyllids and at the site of attachment to the protonema. C) Gametophore showing strongest *GUS* expression in the youngest phyllids, axillary hair cells and at the site of attachment to the protonema. D) Protonemal tissue showing strongest *GUS* expression in the tip cells.

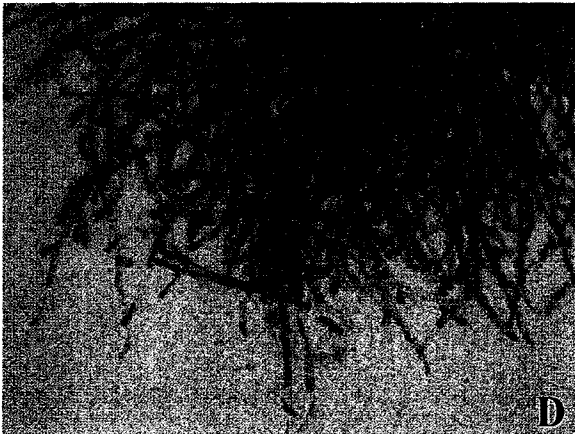


Figure 21. Light micrographs of *Physcomitrella* stably transformed with pT-GUS-6b(-); transformant #1. Plants were incubated with *GUS* substrate for 5 hours. A) A gametophore bud showing weak *GUS* expression. B) Young gametophore showing very slight *GUS* expression in youngest phyllids and at the site of attachment to protonema. C) Gametophore showing extremely weak *GUS* expression in the youngest phyllids. D) Protonemal tissue showing weak *GUS* expression in the tip cells.

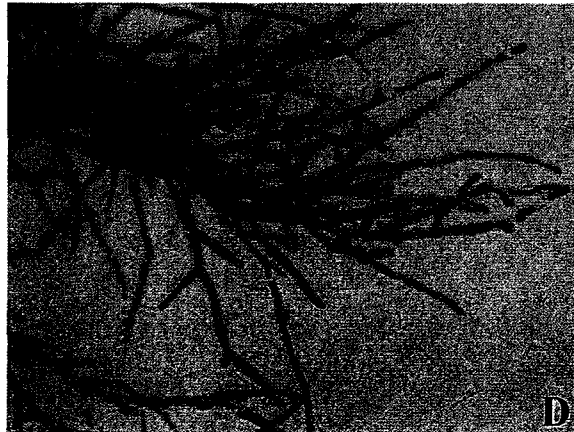
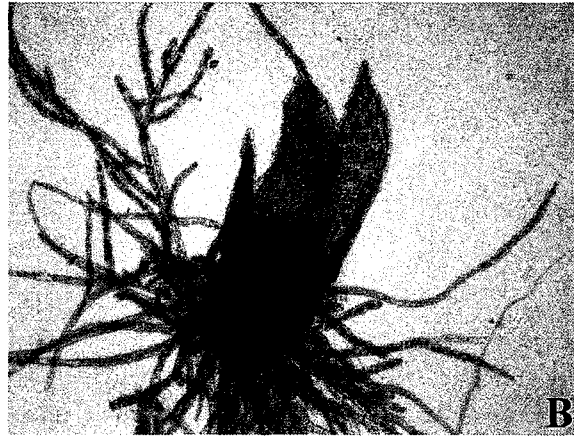
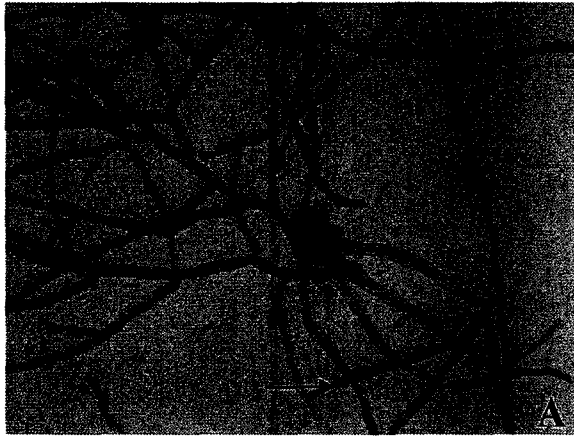


Figure 22. Light micrographs of *Physcomitrella* stably transformed with pT-GUS-6b(-); transformant #2. Plants were incubated with *GUS* substrate for 5 hours. A) A gametophore bud showing weak *GUS* expression and rhizoids (arrow) showing stronger *GUS* expression. B) Young gametophore showing strongest *GUS* in the base at the site of protonemal attachment. C) Gametophore showing *GUS* expression in immature phyllids, axillary hair cells and the base at the site of protonemal attachment. D) Protonemal tissue showing strongest *GUS* expression in the tip cells.

CDPK RNAi

An RNAi construct (pTUGGi-RNAi6) was made that would produce a 255 bp double-stranded segment of RNA containing part of the *CPK6* coding sequence (Figure 23) in *Physcomitrella* following transformation in order to silence *CPK6* expression. The segment of *CPK6* sequence used to produce the dsRNA was amplified from the kinase region and shows high sequence identity (~80-90%) to *CPK1*, *CPK3*, *CPK4* and *CPK8*, which may result in silencing of these genes also since identity of ~80% or greater is sufficient to downregulate gene expression (Miki et al., 2005). The pTUGGi-RNAi6 (Figure 24) vector and the pTUGi control vector (Figure 25), were transformed independently into the *Physcomitrella* NLS-4 line (Bezanilla et al., 2003), where each vector will produce a double-stranded RNA segment of *GUS* sequence. The NLS-4 line expresses a GFP-GUS fusion protein that is localized to the nucleus (Figure 26). The nuclear GFP can be detected easily using a fluorescence microscope. The double-stranded RNA segment of *GUS* sequence matches a segment of the *GFP-GUS* mRNA, and is sufficient to cause degradation of the transcript through RNAi (Bezanilla et al., 2003). When either vector is transformed into the NLS-4 line, RNAi should be activated and this can be confirmed since fluorescence will no longer be detectable. *GFP* silencing acts as an internal control to show that RNAi is being triggered by dsRNA produced from either pTUGi or pTUGGi-RNAi6. Transient transformants were obtained for both pTUGi and pTUGGi-RNAi6. Examination of these transformants with a fluorescence microscope revealed that expression of *GFP-GUS* was being silenced (Figures 27 & 28, D-F), indicating that RNAi was being triggered by dsRNA produced from both constructs. However, no obvious difference in phenotype was observed between

transformants containing the pTUGi control vector and transformants containing the pTUGGi-RNAi6 vector (Figures 27 & 28, A-C).

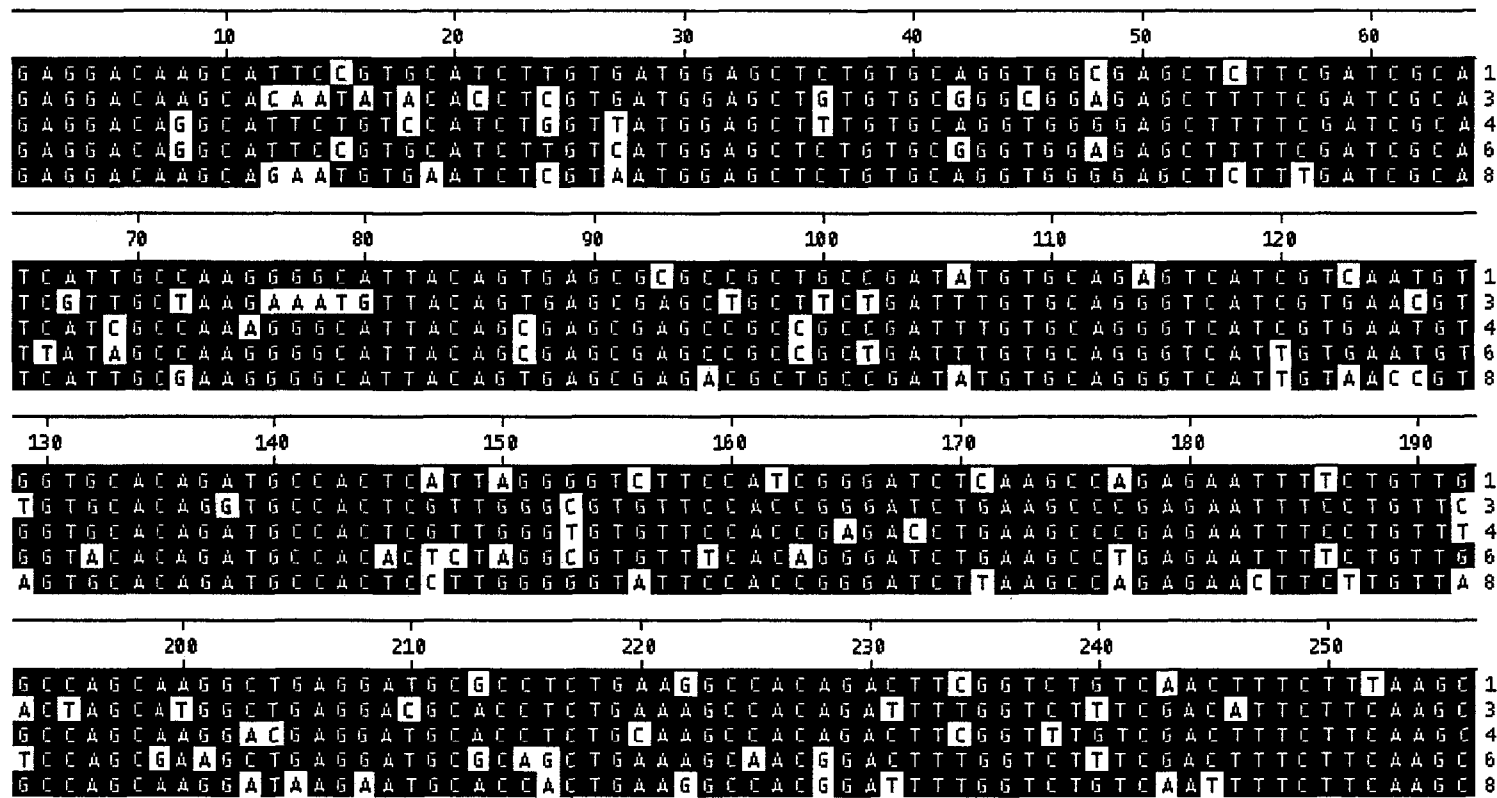
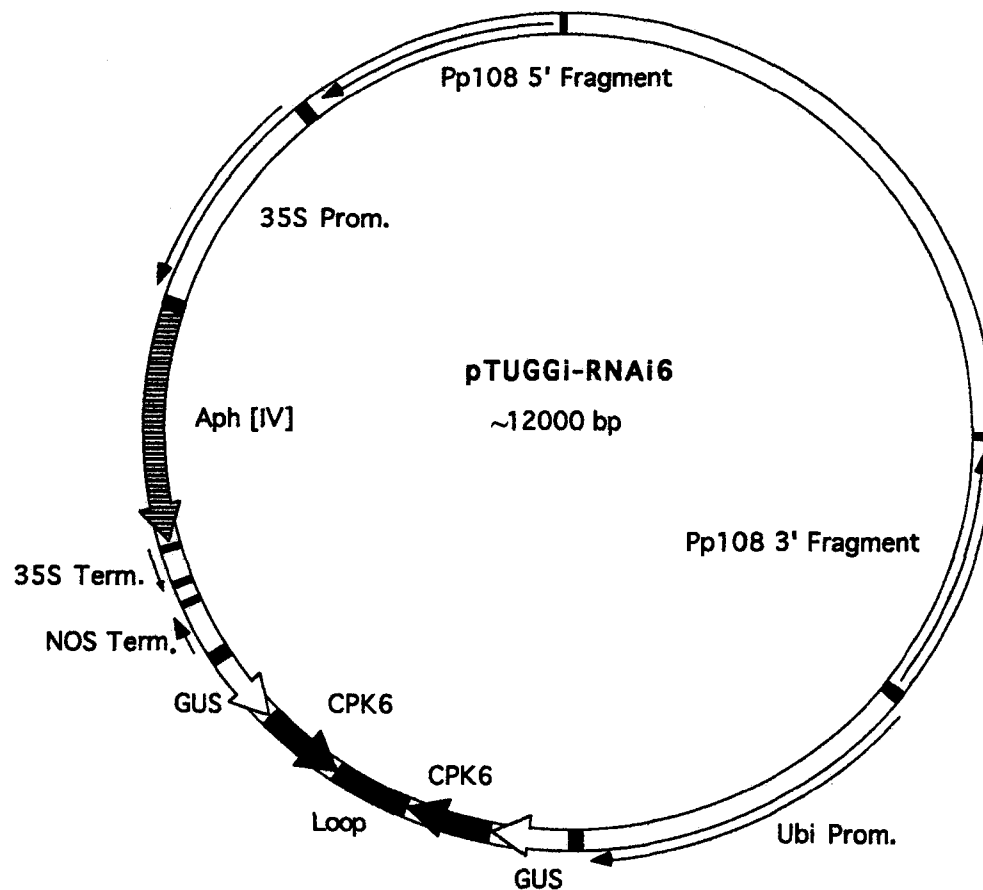


Figure 23. Segment of DNA from a conserved region of the kinase domain used for CPK6 RNAi aligned to corresponding segments of closely-related CDPKs. Alignment was done using the ClustalW method in MegAlign from the DNASTar package (Lasergene). Shaded areas represent nucleotides identical to consensus sequence at the position. CDPK gene number is listed to the right of each line.



65

Figure 24. Diagram of pTUGi-RNAi6 construct for CDPK silencing. Pp108 5' and 3' fragments are identical to sites in the *Physcomitrella* genome and can serve as sites for homologous recombination. The 35S promoter is from cauliflower mosaic virus (Odell et al., 1985) is a moderate-strength promoter in *Physcomitrella* (M. Bezanilla, personal communication). In this construct, the 35S promoter is used to express the *Aph [IV]* gene in concert with the 35S terminator. *Aph [IV]* confers hygromycin resistance. The maize ubiquitin promoter drives transcription of the *GUS* and *CPK6* inverted repeats in concert with the *NOS* terminator. Transcription yields *GUS* and *CPK6* double-stranded RNA.

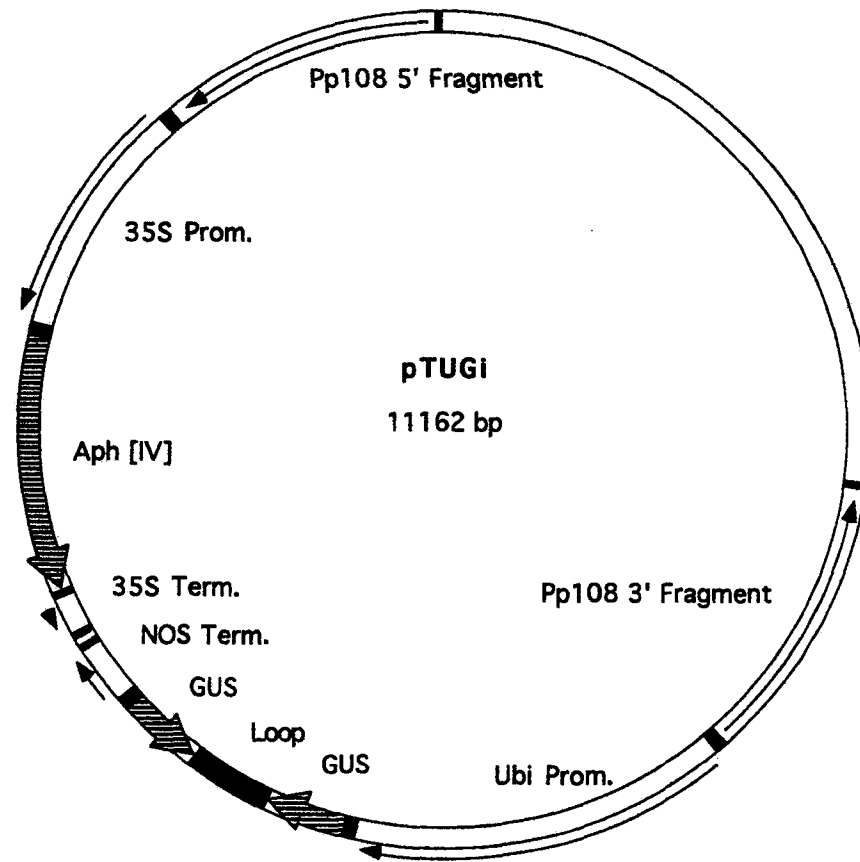


Figure 25. Diagram of pTUGi construct used as control for RNAi. Pp108 5' and 3' fragments are identical to sites in the *Physcomitrella* genome and can serve as sites for homologous recombination. The 35S promoter is from cauliflower mosaic virus (Odell et al., 1985) and is a moderate-strength promoter in *Physcomitrella* (M. Bezanilla, personal communication). In this construct, the 35S promoter is used to express the *Aph [IV]* gene in concert with the 35S terminator. *Aph [IV]* confers hygromycin resistance. The maize ubiquitin promoter drives transcription of the inverted GUS sequences in concert with the NOS terminator. Transcription of GUS inverted repeats produces GUS double-stranded RNA.

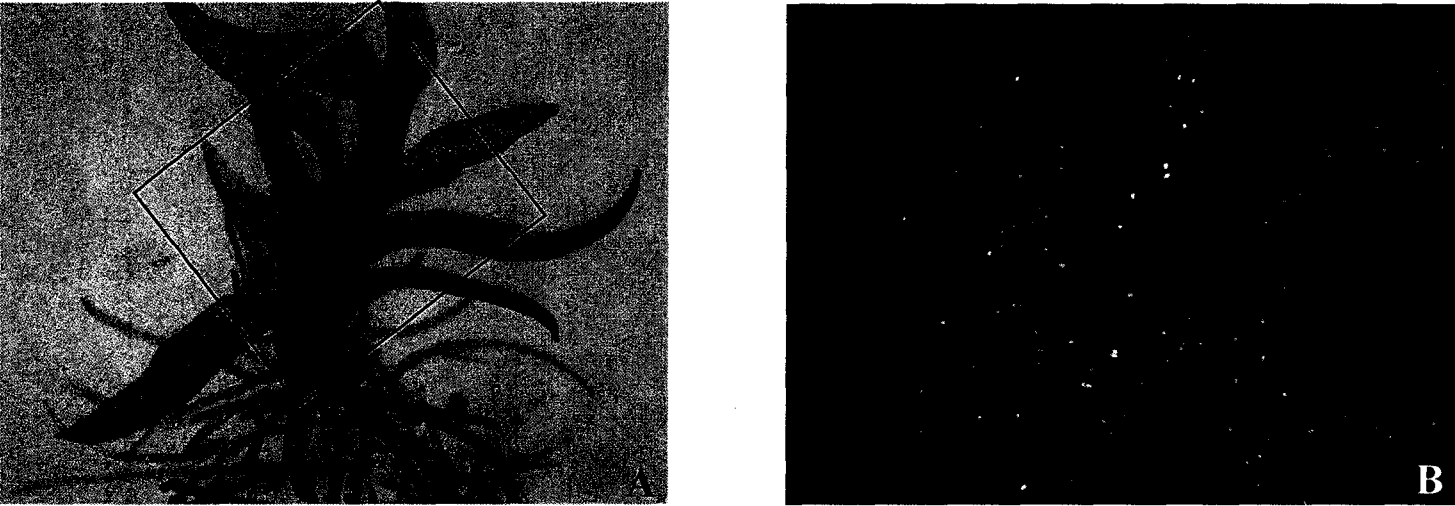


Figure 26. *Physcomitrella* NLS4 line. A) Light micrograph of NLS4 plant. B) NLS4 plant showing GFP localized to the nucleus under a fluorescence microscope. Section of plant in light micrograph that corresponds to image taken by fluorescence scope is outlined in black.

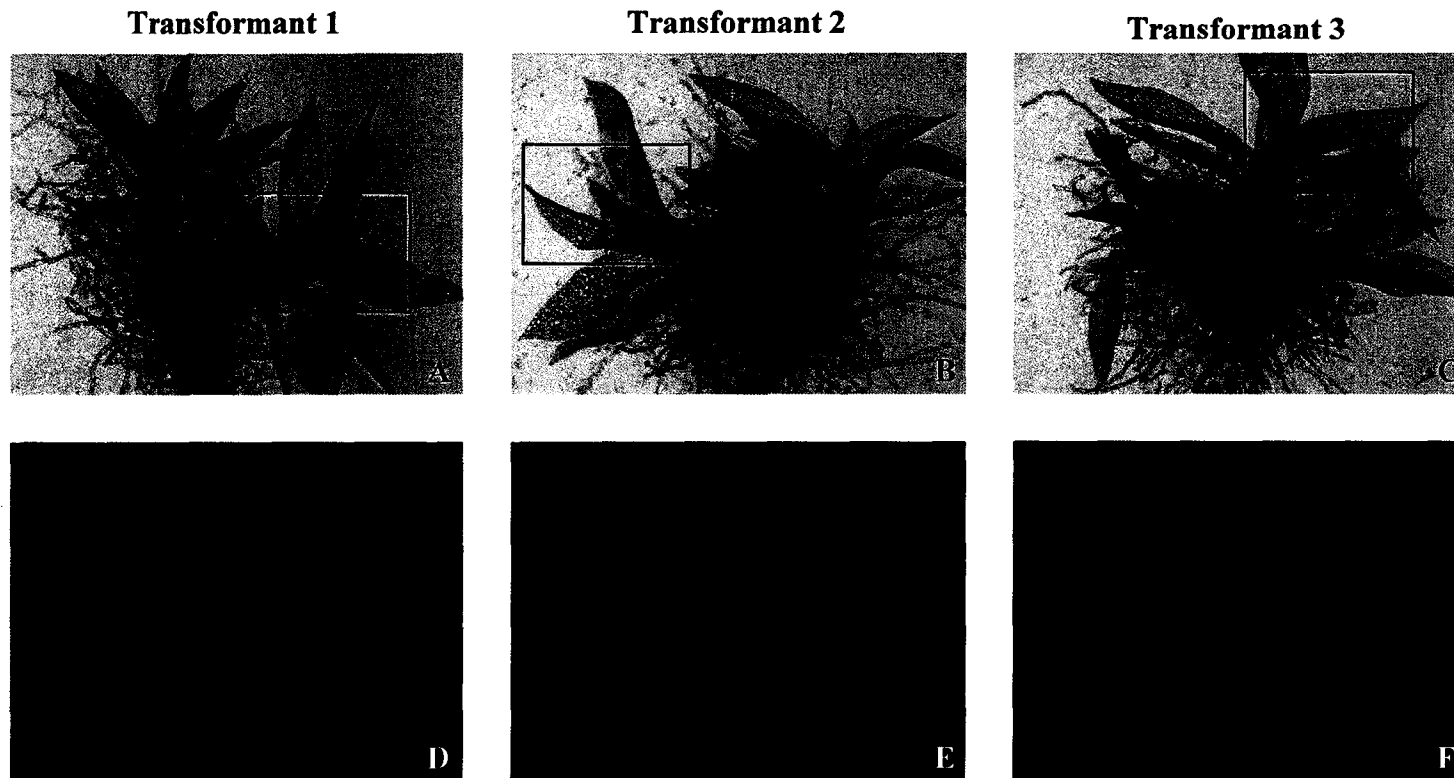


Figure 27. Independent *Physcomitrella* NLS4 plants transiently transformed with the *GUS*-RNAi control vector, pTUGi. A-C) Light micrographs of transformants. D-F) Confirmation of silencing of transformants by a lack of GFP localized to the nucleus under a fluorescence microscope. Section of plant in light micrograph that corresponds to image taken by fluorescence scope is outlined in black.

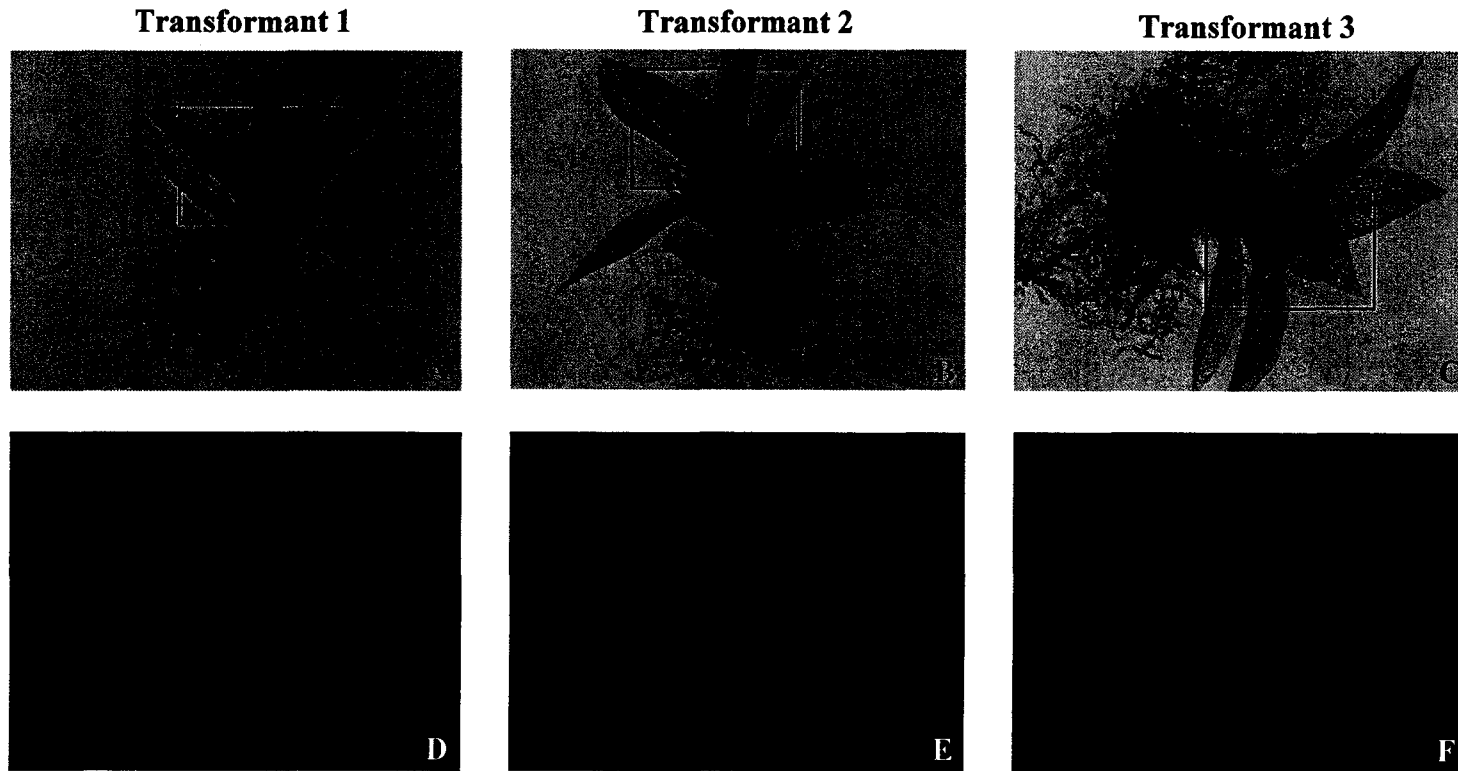


Figure 28. Independent *Physcomitrella* NLS4 plants transiently transformed with the *GUS-CPK6-RNAi* construct, pTUGGi-RNAi6. A-C) Light micrographs of transformants. D-F) Confirmation of silencing of transformants by a lack of GFP localized to the nucleus under a fluorescence microscope. Section of plant in light micrograph that corresponds to image taken by fluorescence scope is outlined in black.

CHAPTER IV

DISCUSSION

Identification of CDPKs

To date, I have identified sixteen CDPKs in *Physcomitrella* using the two previously described methods. This is fewer than the number of CDPKs that have been identified in vascular plants where the genome has been completely sequenced and annotated (34 CDPKs in *Arabidopsis* and 29 CDPKs in *Oryza*; Hrabak et al., 2003; Asano et al., 2005). Since there seems to be fewer CDPKs in *Physcomitrella* than in *Arabidopsis* and *Oryza*, there may be less functional overlap of individual CDPKs. Less functional redundancy would make it easier to produce a phenotype through gene inactivation. This makes *Physcomitrella* a good plant in which to study CDPKs.

Previous phylogenetic analysis showed that CDPKs in vascular plants are arranged in three major groups with one minor group (Hrabak et al., 2003; Asano et al., 2005). The minor group consists of CDPKs that resemble the CDPK-related kinases (CRKs). CRKs are kinases that have similar sequence to CDPKs but lack Ca²⁺ binding motifs (EF-hands) in the C-terminal domain. The CDPKs in the minor group still contain EF-hand, Ca²⁺-binding motifs, but are more similar in sequence and in intron-exon structure to CRKs than to CDPKs. Interestingly, *Physcomitrella* does contain at least five CRK genes (K. Chervincky, personal observation), but I have not identified any CRK-like CDPKs. Similar to vascular plants, the *Physcomitrella* CDPKs that I have identified also formed three major groups; however no minor group of CRK-like CDPKs was present. Either *Physcomitrella* CDPK members of the CRK-like group were not

identified by my search of the moss sequence database or this minor group is missing altogether. It would be interesting if this minor group of CDPKs was missing from all bryophytes since that would indicate that these CRK-like CDPKs arose in vascular plants after the split from bryophytes.

Phylogenetic analyses were done comparing *Physcomitrella* CDPKs to CDPKs from *Arabidopsis*, *Oryza*, additional bryophytes, green algae and *Plasmodium falciparum*, the best-studied protist from the CDPK-containing Apicomplexans. CDPKs from *Arabidopsis* and *Oryza* were found interspersed amongst each other within each major group, whereas *Physcomitrella* CDPKs remain on their own branch within each group. This result indicates that within each of the three groups, CDPKs from the vascular plants are more closely related to each other than they are to CDPKs from *Physcomitrella*. I found that CDPKs from the mosses *Tortula* and *Funaria* were located on the same branch as *Physcomitrella* CDPKs in groups 2 and 3 respectively, indicating that these moss CDPKs are more closely related to each other than they are to the CDPKs from vascular plants. The CDPK from the liverwort *Marchantia* was located in Group 1; however, it did not associate any more so with *Physcomitrella* CDPKs than with vascular plant CDPKs, indicating that the liverwort CDPK isn't more closely-related to either type of plant. CDPKs from *Plasmodium* and from the green algae *Dunaliella* and *Chlamydomonas* were not obviously associated with any of the major groups of CDPKs from the plants that I used to construct my tree, indicating the distant relationship of CDPKs from these organisms to CDPKs in land plants.

Of the three groups formed by *Physcomitrella* CDPKs, Group 1 had the least amount of overall amino acid sequence conservation, showing 66.5% to 82.9% overall

identity to each other. The variable domains in Group 1 were the least conserved when compared to each other, ranging from 14.0% to 47.8% identity. The variable domain for each CDPK in Group 1 is predicted to be myristoylated by the plant-based myristoylation algorithm (available at <http://plantsp.genomics.purdue.edu/plantsp/html/myrist.html>) with the exception of CPK10. Myristoylation is a type of acylation where a myristate (14-carbon fatty acid) is attached to an N-terminal glycine. The hydrophobicity of the myristate can cause the protein to loosely associate with a membrane. Myristoylation algorithms are created by examining consensus sequences of known myristoylated proteins. It should be pointed out that myristoylation prediction algorithms are not 100% accurate, especially the algorithm for plants which is still being developed. Since CPK10 contains an N-terminal glycine it may in fact be myristoylated, however this would have to be determined experimentally. All *Physcomitrella* CDPKs in Group 1 (with the possible exception of CPK10) contain predicted myristoylation sites. All *Arabidopsis* CDPKs in Group 1 also contain predicted myristoylation sites, and nine of the twelve CDPKs have been experimentally tested and shown to be myristoylated (A. Argyros and E. Hrabak, unpublished data).

Palmitoylation is another type of acylation where a palmitate (16-carbon fatty acid) is attached to a cysteine. A common acylation motif, the SH4 domain consists of one or more cysteines near a myristoylation site (Resh, 1999). CDPKs that contain an SH4 domain are usually palmitoylated (A. Argyros and E. Hrabak, unpublished data; Martin and Busconi, 2000). The palmitate is hydrophobic enough to provide a strong association between a palmitoylated protein and a membrane. If CPK10 is in fact myristoylated, it is likely to be palmitoylated as well since it contains a cysteine near the

N-terminus. The variable domains of each *Arabidopsis* CDPK in Group 1 (except for CPK3) also contain at least one potential palmitoylation site. The subcellular localizations of CDPKs in *Arabidopsis* and *Oryza* have been tested using both GFP fusions and subcellular fractionation experiments (Martin and Busconi, 2000; Dammann et al., 2003; Argyros, 2005). Previously known *Arabidopsis* CDPK localization results are listed in Table 9. All tested CDPKs that contain both myristoylation and palmitoylation sites are localized to membranes, but are found in the soluble fraction upon the disruption of the myristoylation site (Argyros, 2005), indicating that palmitoylation is dependent upon myristoylation. All tested *Arabidopsis* CDPKs in Group 1 (except CPK3) were plasma membrane-associated. CPK3 was localized to the nucleus (Dammann et al., 2003; Argyros, 2005; Table 9). These results suggest that *Physcomitrella* CPK10 is likely to be membrane-associated if it is in fact myristoylated and palmitoylated. It is difficult to predict the subcellular localization of the remaining *Physcomitrella* CDPKs in Group 1 since myristate can only loosely associate a protein with a membrane (Peitzsch and McLaughlin, 1993).

Physcomitrella CDPKs in Group 2 on average showed slightly more overall amino acid sequence identity to each other (63.7% to 93.5%) than did CDPKs in Group 1. The variable domain of CDPKs in Group 2 showed a wider range of sequence identity (15.5% to 82.8%) than did CDPKs in Group 1. The variable domain of CPK9 is predicted to be myristoylated by the plant-specific myristoylation prediction algorithm and contains one potential palmitoylation site. CPK16 was not predicted to be myristoylated, though it does contain an N-terminal glycine. As previously mentioned, the plant-specific prediction algorithm is not 100% accurate, so CPK16 may in fact be

Table 9. Phylogenetic groupings of *Physcomitrella* and *Arabidopsis* CDPKs (group number listed on left). N-terminal amino acids are shown to the right of each gene. Potential myristoylation sites are shown in bold and potential palmitoylation sites are underlined. Subcellular localization results of *Arabidopsis* CDPKs (Argyros, 2005; Damman et al., 2003) are shown in the far right column. Dashed lines indicate there is no subcellular localization data available. *Arabidopsis* CDPKs from the CRK-like group are not shown. PM=plasma membrane; ER=endoplasmic reticulum.

<i>Physcomitrella</i>		<i>Arabidopsis</i>			
Gene Name	N-terminal AAs	Gene Name	N-terminal AAs	Subcellular localization	
1	<i>CPK1</i>	MGNTSSR	<i>CPK3</i>	MGHRH SK	Nucleus
	<i>CPK3</i>	MGNSSGR	<i>CPK9</i>	MGNC FAK	PM
	<i>CPK4</i>	MGNVSGR	<i>CPK15</i>	MG C FSSK	PM
	<i>CPK6</i>	MGNISGR	<i>CPK17</i>	MGNCC SH	PM
	<i>CPK8</i>	MGNTSAR	<i>CPK19</i>	MGCLCIN	----
	<i>CPK10</i>	MGNQCVG	<i>CPK21</i>	MG C FSSK	PM
			<i>CPK22</i>	MGNCCGS	----
			<i>CPK23</i>	MG C FSSK	PM
			<i>CPK27</i>	MG C FSSK	PM
			<i>CPK29</i>	MG F CFSK	----
			<i>CPK31</i>	MGCYSSK	PM
		<i>CPK33</i>	MGNCLAK	PM	
2	<i>CPK7</i>	MAEVALL	<i>CPK1</i>	MGNTCVG	Peroxisome
	<i>CPK9</i>	MGNTCVG	<i>CPK2</i>	MGNAC CVG	ER, Soluble
	<i>CPK11</i>	MRRAGLN	<i>CPK4</i>	MEKPNPR	Soluble
	<i>CPK12</i>	MRRGVNL	<i>CPK5</i>	MGN S CRG	PM, Soluble
	<i>CPK16</i>	MGNTCIG	<i>CPK6</i>	MGN S CRG	PM, Soluble
			<i>CPK11</i>	METKPNP	Soluble
			<i>CPK12</i>	MANKPRT	Soluble
			<i>CPK20</i>	MGNTCVG	----
		<i>CPK25</i>	MGNVC VH	PM	
		<i>CPK26</i>	MKHS G GN	Soluble	
3	<i>CPK2</i>	MGNCC VG	<i>CPK7</i>	MGNCCGN	PM
	<i>CPK5</i>	MGNCC VG	<i>CPK8</i>	MGNCCAS	PM
	<i>CPK13</i>	MGNCC AG	<i>CPK10</i>	MGN C NAC	PM
	<i>CPK14</i>	MGNCC VG	<i>CPK13</i>	MGNCCRS	PM
	<i>CPK15</i>	MGNCC VG	<i>CPK14</i>	MGNCCGT	----
			<i>CPK24</i>	MG S CVSS	----
			<i>CPK30</i>	MGN C IA C	PM
			<i>CPK32</i>	MGNCCGT	PM
		<i>CPK34</i>	MGNCC SH	PM	

myristoylated. If CPK16 is myristoylated, it is probably palmitoylated as well since there is a nearby cysteine. The variable domains of the other three CDPKs in Group 2 (CPK7, CPK11 and CPK12) had neither potential acylation site. Six of the *Arabidopsis* CDPK variable domains from Group 2 had one potential myristoylation site and one potential palmitoylation site, while the other four CDPK variable domains had neither.

Arabidopsis CDPKs that were tested for subcellular localization in Group 2 were associated with a membrane when potential myristoylation and palmitoylation sites were present. CDPKs that did not possess any potential acylation sites were soluble (Dammann et al., 2003; Argyros 2005, Table 9). If the mechanisms that control membrane binding of acylated proteins are the same in *Arabidopsis* and *Physcomitrella*, then CPK9 and CPK16 of the *Physcomitrella* CDPKs in Group 2 may be membrane-associated while the other three are most likely soluble.

Physcomitrella CDPKs from Group 3 are the most closely related to each other, showing an overall amino acid identity ranging from 86.6% to 93.1%. The variable domain, which showed higher divergence in the other two *Physcomitrella* CDPK groups, had sequence identity ranging from 65.1% to 81.4%. All variable domains of CDPKs in Group 3 contained one potential myristoylation site and two potential palmitoylation sites. This pattern was reflected in *Arabidopsis* CDPKs in Group 3, which also contained one potential myristoylation site and, except for CPK24, two potential palmitoylation sites. CPK24 contained only one potential palmitoylation site (Table 9). Of the *Arabidopsis* CDPKs in Group 3 that were tested for subcellular localization, all were plasma membrane-associated (Dammann et al., 2003; Argyros, 2005; Table 9). This suggests that all

Physcomitrella CDPKs in Group 3 are membrane-associated, again depending on whether or not similar acylation enzymes exist in *Arabidopsis* and *Physcomitrella*.

The kinase, autoinhibitory and calmodulin-like domains were all more conserved within each group than the variable domain. On average, Group 1 showed the lowest amount of amino acid identity for each domain, ranging from 74.5% to 90.5% for the kinase domain, 83.9% to 93.5% for the autoinhibitory domain and 71.1% to 86.8% for the calmodulin-like domain. Group 2 on average showed more amino acid identity for each domain with the exception of *CPK7*, which consistently showed the least amount of identity. Amino acid identity for Group 2 ranged from 70.1% to 95.8% for the kinase domain, 67.7% to 100% for the autoinhibitory domain and 58.1% to 91.7% for the calmodulin-like domain. Group 3 showed the highest amount of sequence identity for each domain, ranging from 92.8% to 97% for the kinase domain, 96.7% to 100% for the autoinhibitory domain and 81.1% to 90.0% for the calmodulin-like domain. The fact that all three domains show conservation indicates that these domains must be necessary for typical CDPK function. Since the variable domain is much less conserved, this region may be the most responsible for varying protein function amongst individual CDPKs.

Kinase Assay

The kinase assay was done to confirm that two of the genes that I identified encode enzymes that phosphorylate substrates in a Ca^{2+} -dependent manner. The kinase activity of CPK4 and CPK6 was stimulated ~15-fold in the presence of Ca^{2+} . This result showed that these two enzymes phosphorylate a model substrate protein in a Ca^{2+} concentration-dependent manner, confirming that these enzymes are indeed calcium-dependent protein kinases. Though only two CDPKs were actually assayed, the other fourteen CDPKs probably function in a similar manner also since the kinase, autoinhibitory and calmodulin-like domains are similar in amino acid sequence to those of CPK4 and CPK6. In addition, the calmodulin-like domain of every *Physcomitrella* CDPK contains four predicted EF-hands for binding Ca^{2+} .

CDPKs from vascular plants such as *Arabidopsis*, *Oryza*, *Nicotiana tabacum* and *Glycine max* have been shown to autophosphorylate *in vitro* (Putnam-Evans et al., 1990; Abo-el-Saad and Wu, 1995; Yoon et al., 1999; Hegeman et al., 2006). CPK4 and CPK6 enzymes from *Physcomitrella* showed ^{32}P incorporation indicating that *Physcomitrella* CDPKs also autophosphorylate *in vitro*, behaving in a similar fashion to CDPKs from vascular plants. Little is known of the function of CDPK autophosphorylation *in vivo*, though one study claims that autophosphorylation leads to CDPK activation (Romeis et al., 2001). However more research needs to be done on CDPK autophosphorylation before any conclusions can be drawn regarding its effect on CDPK activity.

CPK6 Promoter-GUS Expression

GUS expression in *Physcomitrella* stably transformed with the *CPK6* promoter-*GUS* constructs containing the 5' UTR intron (pT-GUS-6a or pT-GUS-6b) was strong in gametophore buds, young phyllids, protonemal tip cells and the base of the gametophore at the site of protonemal attachment. Expression was also strong in axillary hair cells, but this was not consistently observed. Perhaps this variation in expression occurs because *CPK6* is expressed in the axillary hair cells only at certain stages of development, and there is no easy way to ensure plant samples are at exactly comparable stages of development when they are collected for GUS assays. Less expression was seen in mature phyllids and protonemal cells with the exception of tip cells. In general these results indicate that *CPK6* seems to be more strongly expressed in areas of the plant that are actively growing, with the exception of the axillary hair cells and the gametophore base. *CPK6* expression seems to be weaker in areas of the plant that in which cells are not actively dividing, i.e. mature phyllids and non-tip protonemal cells. These results suggest that *CPK6* may play a larger role in actively dividing cells as opposed to non-dividing cells, however *CPK6* probably plays a role in non-dividing cells as well since was still observed in these parts of the plant. It is interesting that *CPK6* is expressed in virtually all tissues of *Physcomitrella* to some degree, whereas the expression of *CPK2* in *Arabidopsis* is only in specific areas of the plant (E. Hrabak, unpublished results). It should be pointed out that there are some drawbacks to using the *GUS* reporter gene, some of which include the fact that the GUS protein probably doesn't have the same half-life as the *CPK6* protein, and it is possible that not all of the *CPK6* regulatory information was contained in the *CPK6* promoter-*GUS* construct. The only way to obtain 100%

accurate *CPK6* expression results is through techniques such as an *in situ* hybridization with RNA probes or antibodies, both of which are difficult techniques to successfully execute (E. Hrabak, personal communication).

Stable transformants containing pT-GUS-6a or pT-GUS-6b showed overall stronger expression levels than stable transformants that lacked the 5' UTR intron, pT-GUS-6a(-) or pT-GUS-6b(-). Since these constructs are identical except for the presence of the 5' UTR intron in pT-GUS-6a and pT-GUS-6b, the change in expression can be attributed to the intron sequence. The overall expression pattern between the two sets of constructs is similar, but the total expression is greatly reduced in constructs without the intron. These results suggest that the intron in the 5'UTR enhances expression of *CPK6*, but doesn't have a major influence on when and where *CPK6* is expressed.

The reduced amount of *GUS* expression seen in *Physcomitrella* transformants that contain constructs lacking the intron in the 5' UTR is consistent with results from studies in vascular plants. For example, in *Oryza*, a *rubi3* promoter-*GUS* fusion construct containing a 1140 bp 5' UTR intron immediately before the translation start site was used for transient transformation and compared with transformants that lacked the intron. Removal of the intron resulted in a 20-fold decrease in *GUS* expression in transient transformants (Sivamani and Qu, 2006). Also in *Oryza*, removal of a 291 bp intron from the 5' UTR of a *CPK2* promoter-*GUS* construct almost completely abolished *GUS* expression in transformants. When an 83 bp 3' sequence of the same intron was removed, leaving the majority of the intron intact, but disrupting the 3' splice site, little to no *GUS* expression was observed in *Oryza* transformants. The results from the disruption of the splice site suggests that it is not the intron sequence, but rather the act of intron

splicing that is necessary for enhanced *CPK2* expression (Morello et al., 2006). In *Arabidopsis*, the absence of a 602 bp intron in the 5' UTR of the *EF1 α -A3* gene in promoter fusions to the firefly luciferase gene resulted in significantly less expression in both transient and stable transformants than in plants transformed with an identical construct containing the 602 bp intron in the 5' UTR (Chung et al., 2006). The effect of intron length was also examined by partial deletion of the intron internal to the splice sites. This experiment revealed that partial deletion of the intron can result in downregulation of expression as well suggesting that other factors, possibly enhancers, in addition to the act of splicing affect expression of the *EF1 α -A3* gene (Chung et al., 2006).

In other studies, removal of the first intron resulted in a different expression profile as compared to the expression of genes where the intron was left intact. For example, a subtle difference was seen in the sucrose synthase gene (*Sus3*) of potato where the deletion of an intron from the 5' UTR resulted in a 5-fold reduction of expression in roots, but no change in levels of expression in leaves, stems or tubers (Fu et al., 1995). A more drastic change in expression profile was observed in *Oryza* where two *OsTubA1* promoter-*GUS* constructs were made. One contained the first three exons and two introns, and the other contained only part of the first exon and no introns.

Transformants that contained the former construct showed strongest expression in root cells that were actively dividing and in young shoots. Transformants containing the latter construct showed strongest expression in the elongation zones of roots and equal expression in young and mature shoots (Jeon et al., 2000). This is one of the few reports where an intron in the 5' UTR impacts the expression profile of a gene. A change in the expression profile was not seen in my *CPK6* expression results.

Many mechanisms have been proposed for intron-mediated enhancement and are outlined in the introduction. However the mechanisms that seem to have received the most experimental attention are associated with the binding of the exon-junction complex (EJC) to an mRNA after intron splicing has occurred. The EJC enhances expression by increasing mRNA nuclear export and translation rates in eukaryotes (Le Hir et al., 2003). Proteins that make up the EJC in the nucleus include Y14, Magoh, SRm160, RNPS1, REF, UAP56 and Upf3. However, once in the cytoplasm most of these proteins are no longer associated with the mRNA with the exception of Y14 and Magoh. Upf2 becomes associated with the mRNA in the cytoplasm, although it was not associated with it in the nucleus (Le Hir et al., 2000; Gatfield et al., 2001; Kataoka et al., 2001; Kim et al., 2001; Le Hir et al., 2001a). The EJC proteins that are involved in nuclear export include REF and UAP56. REF upregulates mRNA export by directly contacting the nuclear export factor TAP. TAP, with its partner p15, directly binds to the nuclear pore complex (Conti and Izaurralde, 2001). UAP56 is a spliceosome factor, but is also crucial for recruitment of REF to the EJC (Reed and Hurt, 2002). Once in the cytoplasm, it is unclear how the EJC upregulates translation and increases ribosomal affinity to the mRNA. Y14 and Magoh may play a role since they are the only EJC proteins left associated with the mRNA, however there may be undiscovered components of the EJC that are also involved.

I suspect that the decrease of *GUS* expression in the construct containing the intronless *CPK6* regulatory region is at least partially due to the lack of an EJC binding to the pre-mRNA transcript. However, it is impossible to tell from my results if the decrease in *GUS* expression is due to processes involved in mRNA export, translation or

both. Other mechanisms may also be involved as mentioned in the introduction, including the upregulation of polyadenylation, which would increase the stability of the mRNA, and an increase in transcription due to sequences in the intron. It is also probable that not all processes that result in intron-mediated enhancement have been described yet. In all likelihood a combination of mechanisms, those mentioned and those not yet discovered, cooperate to enhance *CPK6* expression in *Physcomitrella*. Nevertheless, my results are among the first to demonstrate intron-mediated enhancement in *Physcomitrella*, and are the first to show intron-mediated enhancement of CDPKs in *Physcomitrella*.

CDPK RNAi

RNAi experiments were performed using the NLS4 line of *Physcomitrella* that contains a GFP-GUS fusion protein that is localized to the nucleus. *Physcomitrella* transiently transformed with vectors that produce dsRNA matching *GUS* RNA sequence (pTUGi and pTUGGi-RNAi6) displayed silencing of the GFP-GUS fusion protein as shown by a lack of GFP fluorescence in the nucleus. The fact that no GFP fluorescence was seen indicated that RNAi was occurring in plants containing these constructs. In addition to producing a segment of *GUS* dsRNA, pTUGGi-RNAi6 should be producing a segment of *CPK6* dsRNA that had ~80-90% sequence identity to *CPK1*, *CPK3*, *CPK4* and *CPK8*. Since expression of the GFP-GUS fusion protein was silenced, expression of at least *CPK6* and possibly *CPK1*, *CPK3*, *CPK4* and *CPK8* may have been silenced as well, however this would have to be proven experimentally through an approach such as reverse transcriptase-PCR. However, no obvious phenotype was observed in these

plants. There are a few possible explanations for this result. First, the lack of phenotype may be due to the fact that some CDPKs in vascular plants are involved in stress responses to drought, salt, cold (Saijo et al., 2000; Sheen, 1996) and pathogen attack (Romeis et al., 2001). If the CDPKs being silenced are in fact involved in stress responses, it is quite possible that a phenotype would not be observed since the transformants were not exposed to any stress. Since RNAi no longer silences *GFP* after a period of less than a month in transient transformants, it is difficult to expose plants to stress and effectively screen for a phenotype. Stable transformants are needed that consistently silence *CPK6* expression so that single stable plant lines can be exposed to numerous types of stress.

A second possibility is that even though the pTUGGi-RNAi6 construct was designed to silence multiple CDPKs, it is possible that only *CPK6* was silenced. If the *CPK6* protein function overlaps with the function of the other closely-related CDPK proteins that weren't silenced, a phenotype would not be observed due to CDPK functional redundancy.

A third reason that may explain why a phenotype was not observed is that, although silencing of *GFP* was observed, silencing of *CPK6* or any of the other CDPKs may not have occurred. However I believe that this is a less likely scenario since a study has shown that using the same vector to silence *FtsZ2-1* resulted in 100% of transient transformants displaying the altered chloroplast phenotype when *GFP* expression was silenced (Bezanilla et al., 2005). However, this is the only experiment that has been published using the pTUGGi vector in *Physcomitrella*, so it is difficult to make any assumptions. The only way

to tell whether *CPK6* or any of the other CDPKs are being silenced is by detecting the level of mRNA transcript.

It will be difficult to directly test whether expression of any CDPKs was being silenced since transient transformants do not produce enough tissue for RNA extraction before the RNAi pathway is no longer acting to silence *GFP*. In order to directly test whether or not any of the CDPKs are being silenced, an NLS4 plant must be stably transformed with the pTUGGi-RNAi6 construct and must show constitutive *GFP* silencing to indicate that the RNAi pathway is operating using dsRNA from the pTUGGi-RNAi6 construct as a trigger. It would also be desirable to obtain an NLS4 plant stably transformed with the pTUGi construct that shows consistent *GFP* silencing as a control. Unfortunately, no stable transformants have been obtained to date that showed consistent *GFP* silencing, although it has been shown that acquiring such a stable in the NLS4 line is possible (Bezanilla et al., 2003). Although RNAi of *GFP-GUS* was successfully triggered by the pTUGGi-RNAi6 and pTUGi constructs, no conclusion can be drawn about the gene silencing and function of *CPK6* at this point, except for the fact that a phenotype for a *CPK6* knockout is somewhat expected since results from the *GUS* experiments indicate that *CPK6* is expressed throughout the plant.

In conclusion, I have identified and sequenced sixteen CDPK genes in *Physcomitrella*. Calcium-dependent protein kinase activity of CPK4 and CPK6 was confirmed, showing that these genes produce enzymes that have similar calcium-stimulated kinase activity as CDPKs from vascular plants. I have defined the expression profile of *CPK6*, as well as demonstrated apparent intron-mediated enhancement of *CPK6* expression. I have also produced *Physcomitrella* transformants that demonstrated

an RNAi response, however no phenotype was observed in preliminary experiments under normal growth conditions.

LIST OF REFERENCES

- Abo-el-Saad M, Wu R (1995)** A rice membrane calcium-dependent protein kinase is induced by gibberellin. *Plant Physiol* **108**: 787-793
- Argyros A (2005)** Myristoylation of calcium-dependent protein kinases from *Arabidopsis thaliana*. University of New Hampshire, Durham
- Asano T, Tanaka N, Yang G, Hayashi N, Komatsu S (2005)** Genome-wide identification of the rice calcium-dependent protein kinase and its closely related kinase gene families: comprehensive analysis of the CDPKs gene family in rice. *Plant Cell Physiol* **46**: 356-366
- Bezanilla M, Pan A, Quatrano RS (2003)** RNA interference in the moss *Physcomitrella patens*. *Plant Physiol* **133**: 470-474
- Bezanilla M, Perroud PF, Pan A, Klueh P, Quatrano RS (2005)** An RNAi system in *Physcomitrella patens* with an internal marker for silencing allows for rapid identification of loss of function phenotypes. *Plant Biol (Stuttg)* **7**: 251-257
- Blencowe BJ, Issner R, Nickerson JA, Sharp PA (1998)** A coactivator of pre-mRNA splicing. *Genes Dev* **12**: 996-1009
- Braddock M, Muckenthaler M, White MR, Thorburn AM, Sommerville J, Kingsman AJ, Kingsman SM (1994)** Intron-less RNA injected into the nucleus of *Xenopus* oocytes accesses a regulated translation control pathway. *Nucleic Acids Res* **22**: 5255-5264
- Brinster RL, Allen JM, Behringer RR, Gelinas RE, Palmiter RD (1988)** Introns increase transcriptional efficiency in transgenic mice. *Proc Natl Acad Sci U S A* **85**: 836-840
- Carmell MA, Xuan Z, Zhang MQ, Hannon GJ (2002)** The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes Dev* **16**: 2733-2742
- Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC (1994)** Green fluorescent protein as a marker for gene expression. *Science* **263**: 802-805
- Chaubet-Gigot N, Kapros T, Flenet M, Kahn K, Gigot C, Waterborg JH (2001)** Tissue-dependent enhancement of transgene expression by introns of replacement histone H3 genes of *Arabidopsis*. *Plant Mol Biol* **45**: 17-30

- Cheng SH, Willmann MR, Chen HC, Sheen J** (2002) Calcium signaling through protein kinases. The Arabidopsis calcium-dependent protein kinase gene family. *Plant Physiol* **129**: 469-485
- Chung BY, Simons C, Firth AE, Brown CM, Hellens RP** (2006) Effect of 5'UTR introns on gene expression in Arabidopsis thaliana. *BMC Genomics* **7**: 120
- Colbran RJ, Fong YL, Schworer CM, Soderling TR** (1988) Regulatory interactions of the calmodulin-binding, inhibitory, and autophosphorylation domains of Ca²⁺/calmodulin-dependent protein kinase II. *J Biol Chem* **263**: 18145-18151
- Conti E, Izaurralde E** (2001) Nucleocytoplasmic transport enters the atomic age. *Curr Opin Cell Biol* **13**: 310-319
- Dammann C, Ichida A, Hong B, Romanowsky SM, Hrabak EM, Harmon AC, Pickard BG, Harper JF** (2003) Subcellular targeting of nine calcium-dependent protein kinase isoforms from Arabidopsis. *Plant Physiol* **132**: 1840-1848
- Filipowicz W** (2005) RNAi: the nuts and bolts of the RISC machine. *Cell* **122**: 17-20
- Fu H, Kim SY, Park WD** (1995) A potato Sus3 sucrose synthase gene contains a context-dependent 3' element and a leader intron with both positive and negative tissue-specific effects. *Plant Cell* **7**: 1395-1403
- Furger A, O'Sullivan JM, Binnie A, Lee BA, Proudfoot NJ** (2002) Promoter proximal splice sites enhance transcription. *Genes Dev* **16**: 2792-2799
- Gatfield D, Le Hir H, Schmitt C, Braun IC, Kocher T, Wilm M, Izaurralde E** (2001) The DExH/D box protein HEL/UAP56 is essential for mRNA nuclear export in Drosophila. *Curr Biol* **11**: 1716-1721
- Hachet O, Ephrussi A** (2001) Drosophila Y14 shuttles to the posterior of the oocyte and is required for oskar mRNA transport. *Curr Biol* **11**: 1666-1674
- Harlow E, Lane D** (1988) *Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Harmon AC, Putnam-Evans C, Cormier MJ** (1987) A calcium-dependent but calmodulin-independent protein kinase from soybean. *Plant Physiol* **83**: 830-837
- Harmon AC, Yoo BC, McCaffery C** (1994) Pseudosubstrate inhibition of CDPK, a protein kinase with a calmodulin-like domain. *Biochemistry* **33**: 7278-7287
- Harper JF, Huang JF, Lloyd SJ** (1994) Genetic identification of an autoinhibitor in CDPK, a protein kinase with a calmodulin-like domain. *Biochemistry* **33**: 7267-7277

- Harper JF, Sussman MR, Schaller GE, Putnam-Evans C, Charbonneau H, Harmon AC (1991)** A calcium-dependent protein kinase with a regulatory domain similar to calmodulin. *Science* **252**: 951-954
- Hegeman AD, Rodriguez M, Han BW, Uno Y, Phillips GN, Jr., Hrabak EM, Cushman JC, Harper JF, Harmon AC, Sussman MR (2006)** A phyloproteomic characterization of in vitro autophosphorylation in calcium-dependent protein kinases. *Proteomics* **6**: 3649-3664
- Holstege FC, Jennings EG, Wyrick JJ, Lee TI, Hengartner CJ, Green MR, Golub TR, Lander ES, Young RA (1998)** Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* **95**: 717-728
- Horton RM, Ho SN, Pullen JK, Hunt HD, Cai Z, Pease LR (1993)** Gene splicing by overlap extension. *Methods Enzymol* **217**: 270-279
- Hrabak EM (2000)** Calcium-dependent protein kinases and their relatives. *Adv Bot Res* **32**: 185-223
- Hrabak EM, Chan CW, Gribskov M, Harper JF, Choi JH, Halford N, Kudla J, Luan S, Nimmo HG, Sussman MR, Thomas M, Walker-Simmons K, Zhu JK, Harmon AC (2003)** The Arabidopsis CDPK-SnRK superfamily of protein kinases. *Plant Physiol* **132**: 666-680
- Hrabak EM, Dickmann LJ, Satterlee JS, Sussman MR (1996)** Characterization of eight new members of the calmodulin-like domain protein kinase gene family from *Arabidopsis thaliana*. *Plant Mol Biol* **31**: 405-412
- Ivashuta S, Liu J, Liu J, Lohar DP, Haridas S, Bucciarelli B, VandenBosch KA, Vance CP, Harrison MJ, Gantt JS (2005)** RNA interference identifies a calcium-dependent protein kinase involved in *Medicago truncatula* root development. *Plant Cell* **17**: 2911-2921
- Iwata Y, Kuriyama M, Nakakita M, Kojima H, Ohto M, Nakamura K (1998)** Characterization of a calcium-dependent protein kinase of tobacco leaves that is associated with the plasma membrane and is inducible by sucrose. *Plant Cell Physiol* **39**: 1176-1183
- Jefferson RA (1987)** Assaying chimeric genes in plants: the GUS gene fusion system. *Plant Mol Biol Rep* **5**: 387-405
- Jeon JS, Lee S, Jung KH, Jun SH, Kim C, An G (2000)** Tissue-preferential expression of a rice alpha-tubulin gene, OsTubA1, mediated by the first intron. *Plant Physiol* **123**: 1005-1014

- Kataoka N, Diem MD, Kim VN, Yong J, Dreyfuss G (2001)** Magoh, a human homolog of *Drosophila mago nashi* protein, is a component of the splicing-dependent exon-exon junction complex. *EMBO J* **20**: 6424-6433
- Khvorova A, Reynolds A, Jayasena SD (2003)** Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115**: 209-216
- Kim VN, Yong J, Kataoka N, Abel L, Diem MD, Dreyfuss G (2001)** The Y14 protein communicates to the cytoplasm the position of exon-exon junctions. *EMBO J* **20**: 2062-2068
- Knight H, Trewavas AJ, Knight MR (1996)** Cold calcium signaling in *Arabidopsis* involves two cellular pools and a change in calcium signature after acclimation. *Plant Cell* **8**: 489-503
- Knight H, Trewavas AJ, Knight MR (1997)** Calcium signalling in *Arabidopsis thaliana* responding to drought and salinity. *Plant J* **12**: 1067-1078
- Le Hir H, Gatfield D, Braun IC, Forler D, Izaurralde E (2001a)** The protein Mago provides a link between splicing and mRNA localization. *EMBO Rep* **2**: 1119-1124
- Le Hir H, Gatfield D, Izaurralde E, Moore MJ (2001b)** The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO J* **20**: 4987-4997
- Le Hir H, Izaurralde E, Maquat LE, Moore MJ (2000)** The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. *EMBO J* **19**: 6860-6869
- Le Hir H, Nott A, Moore MJ (2003)** How introns influence and enhance eukaryotic gene expression. *Trends Biochem Sci* **28**: 215-220
- Lindzen E, Choi JH (1995)** A carrot cDNA encoding an atypical protein kinase homologous to plant calcium-dependent protein kinases. *Plant Mol Biol* **28**: 785-797
- Liu K, Sandgren EP, Palmiter RD, Stein A (1995)** Rat growth hormone gene introns stimulate nucleosome alignment in vitro and in transgenic mice. *Proc Natl Acad Sci U S A* **92**: 7724-7728
- Lojda Z (1970)** Indigogenic methods for glycosidases. An improved method for beta-D-glucosidase and its application to localization studies on intestinal and renal enzymes. *Histochemie* **22**: 347-361

- Lu S, Cullen BR** (2003) Analysis of the stimulatory effect of splicing on mRNA production and utilization in mammalian cells. *RNA* **9**: 618-630
- Ma JB, Ye K, Patel DJ** (2004) Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* **429**: 318-322
- Maquat LE, Carmichael GG** (2001) Quality control of mRNA function. *Cell* **104**: 173-176
- Marella HH, Sakata Y, Quatrano RS** (2006) Characterization and functional analysis of ABCISIC ACID INSENSITIVE3-like genes from *Physcomitrella patens*. *Plant J* **46**: 1032-1044
- Martin ML, Busconi L** (2000) Membrane localization of a rice calcium-dependent protein kinase (CDPK) is mediated by myristoylation and palmitoylation. *Plant J* **24**: 429-435
- Matsumoto K, Wassarman KM, Wolffe AP** (1998) Nuclear history of a pre-mRNA determines the translational activity of cytoplasmic mRNA. *EMBO J* **17**: 2107-2121
- Mayeda A, Badolato J, Kobayashi R, Zhang MQ, Gardiner EM, Krainer AR** (1999) Purification and characterization of human RNPS1: a general activator of pre-mRNA splicing. *EMBO J* **18**: 4560-4570
- McKenzie RW, Brennan MD** (1996) The two small introns of the *Drosophila* *affinidisjuncta* Adh gene are required for normal transcription. *Nucleic Acids Res* **24**: 3635-3642
- Miki D, Itoh R, Shimamoto K** (2005) RNA silencing of single and multiple members in a gene family of rice. *Plant Physiol* **138**: 1903-1913
- Mohr SE, Dillon ST, Boswell RE** (2001) The RNA-binding protein Tsunagi interacts with Mago Nashi to establish polarity and localize oskar mRNA during *Drosophila* oogenesis. *Genes Dev* **15**: 2886-2899
- Morello L, Bardini M, Cricri M, Sala F, Breviario D** (2006) Functional analysis of DNA sequences controlling the expression of the rice *OsCDPK2* gene. *Planta* **223**: 479-491
- Nott A, Le Hir H, Moore MJ** (2004) Splicing enhances translation in mammalian cells: an additional function of the exon junction complex. *Genes Dev* **18**: 210-222
- Oancea E, Meyer T** (1998) Protein kinase C as a molecular machine for decoding calcium and diacylglycerol signals. *Cell* **95**: 307-318

- Odell JT, Nagy F, Chua NH** (1985) Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. *Nature* **313**: 810-812
- Palomares AJ, DeLuca MA, Helinski DR** (1989) Firefly luciferase as a reporter enzyme for measuring gene expression in vegetative and symbiotic *Rhizobium meliloti* and other gram-negative bacteria. *Gene* **81**: 55-64
- Parker JS, Roe SM, Barford D** (2004) Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity. *EMBO J* **23**: 4727-4737
- Peitzsch RM, McLaughlin S** (1993) Binding of acylated peptides and fatty acids to phospholipid vesicles: pertinence to myristoylated proteins. *Biochemistry* **32**: 10436-10443
- Plesse B, Criqui MC, Durr A, Parmentier Y, Fleck J, Genschik P** (2001) Effects of the polyubiquitin gene Ubi. U4 leader intron and first ubiquitin monomer on reporter gene expression in *Nicotiana tabacum*. *Plant Mol Biol* **45**: 655-667
- Preall JB, Sontheimer EJ** (2005) RNAi: RISC gets loaded. *Cell* **123**: 543-545
- Price AH, Taylor A, Ripley SJ, Griffiths A, Trewavas AJ, Knight MR** (1994) Oxidative signals in tobacco increase cytosolic calcium. *Plant Cell* **6**: 1301-1310
- Putnam-Evans CL, Harmon AC, Cormier MJ** (1990) Purification and characterization of a novel calcium-dependent protein kinase from soybean. *Biochemistry* **29**: 2488-2495
- Reed R, Hurt E** (2002) A conserved mRNA export machinery coupled to pre-mRNA splicing. *Cell* **108**: 523-531
- Resh MD** (1999) Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins. *Biochim Biophys Acta* **1451**: 1-16
- Romeis T, Ludwig AA, Martin R, Jones JD** (2001) Calcium-dependent protein kinases play an essential role in a plant defence response. *EMBO J* **20**: 5556-5567
- Rose AB** (2004) The effect of intron location on intron-mediated enhancement of gene expression in *Arabidopsis*. *Plant J* **40**: 744-751
- Saijo Y, Hata S, Kyojuka J, Shimamoto K, Izui K** (2000) Over-expression of a single Ca²⁺-dependent protein kinase confers both cold and salt/drought tolerance on rice plants. *Plant J* **23**: 319-327
- Sanders D, Brownlee C, Harper JF** (1999) Communicating with calcium. *Plant Cell* **11**: 691-706

- Schaefer D, Zryd JP, Knight CD, Cove DJ** (1991) Stable transformation of the moss *Physcomitrella patens*. *Mol Gen Genet* **226**: 418-424
- Schaefer DG, Zryd JP** (1997) Efficient gene targeting in the moss *Physcomitrella patens*. *Plant J* **11**: 1195-1206
- Schaefer DG, Zryd JP** (2001) The moss *Physcomitrella patens*, now and then. *Plant Physiol* **127**: 1430-1438
- Sheen J** (1996) Ca²⁺-dependent protein kinases and stress signal transduction in plants. *Science* **274**: 1900-1902
- Sivamani E, Qu R** (2006) Expression enhancement of a rice polyubiquitin gene promoter. *Plant Mol Biol* **60**: 225-239
- Sleckman BP, Gorman JR, Alt FW** (1996) Accessibility control of antigen-receptor variable-region gene assembly: role of cis-acting elements. *Annu Rev Immunol* **14**: 459-481
- Snowden KC, Buchholz WG, Hall TC** (1996) Intron position affects expression from the *tpi* promoter in rice. *Plant Mol Biol* **31**: 689-692
- Sontheimer EJ, Carthew RW** (2005) Silence from within: endogenous siRNAs and miRNAs. *Cell* **122**: 9-12
- Spingola M, Grate L, Haussler D, Ares M, Jr.** (1999) Genome-wide bioinformatic and molecular analysis of introns in *Saccharomyces cerevisiae*. *RNA* **5**: 221-234
- Strepp R, Scholz S, Kruse S, Speth V, 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 U S A* **95**: 4368-4373
- Takahashi K, Isobe M, Knight MR, Trewavas AJ, Muto S** (1997) Hypoosmotic shock induces increases in cytosolic Ca²⁺ in tobacco suspension-culture cells. *Plant Physiol* **113**: 587-594
- Takezawa D, Ramachandiran S, Paranjape V, Poovaiah BW** (1996) Dual regulation of a chimeric plant serine/threonine kinase by calcium and calcium/calmodulin. *J Biol Chem* **271**: 8126-8132
- Towler DA, Adams SP, Eubanks SR, Towery DS, Jackson-Machelski E, Glaser L, Gordon JI** (1988) Myristoyl CoA:protein N-myristoyltransferase activities from rat liver and yeast possess overlapping yet distinct peptide substrate specificities. *J Biol Chem* **263**: 1784-1790
- Vagner S, Vagner C, Mattaj IW** (2000) The carboxyl terminus of vertebrate poly(A)

polymerase interacts with U2AF 65 to couple 3'-end processing and splicing.
Genes Dev **14**: 403-413

Wiegand HL, Lu S, Cullen BR (2003) Exon junction complexes mediate the enhancing effect of splicing on mRNA expression. *Proc Natl Acad Sci U S A* **100**: 11327-11332

Yoon GM, Cho HS, Ha HJ, Liu JR, Lee HS (1999) Characterization of NtCDPK1, a calcium-dependent protein kinase gene in *Nicotiana tabacum*, and the activity of its encoded protein. *Plant Mol Biol* **39**: 991-1001

Zhang M, Yuan T (1998) Molecular mechanisms of calmodulin's functional versatility. *Biochem Cell Biol* **76**: 313-323