Evolution of the phytochrome gene family in land plants and its utility for phylogenetic analyses of flowering plants

by Sarah L Mathews

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biological Sciences

Montana State University

Abstract:
The phytochrome nuclear gene family encodes photoreceptor proteins that mediate diverse developmental responses to red and far red light throughout the life of a plant. From studies of the angiosperm Arabidopsis thaliana, the family has been modeled as comprising five loci, PHYA-PHYE. In gymnosperms, two loci have been detected, while in other nonangiosperm groups just one locus has been observed. A polymerase chain reaction (PCR) based strategy to sample plant DNAs was developed in order to test the gene family model inferred from Arabidopsis in other plant species and' to evaluate the utility of phytochrome DNA sequence data for phylogenetic studies. Results presented here indicate that the Arabidopsis model is not completely appropriate for all angiosperm groups because additional PHY loci related to PHYA and PHYB have evolved independently several times in dicot angiosperms, and monocot angiosperms may lack orthologs of PHYD and PHYE. Nonetheless, for studies of organismal evolution, the phytochrome gene family is potentially useful because "the loci occur as single copy sequences, and the data suggest that the various loci are evolving independently. In two plant families, dicotyledonous Fabaceae (legumes) and monocotyledonous Poaceae (grasses), phytochrome data provided phylogenetic resolution. In addition to nucleotide substitutions, phylogenetically informative insertions and deletions characterize the phytochrome data sets. Furthermore, together with data obtained from public databases, the data detected in this study suggest that the differential distribution of phytochrome loci among flowering plant groups may be phylogenetically informative. The presence of a legume-specific locus most closely related to PHYA may be informative once its phylogenetic distribution is known; likewise, the apparent absence of PHYD and PHYE from monocots and some dicot plant groups potentially resolves relationships among major angiosperm lineages.

EVOLUTION OF THE PHYTOCHROME GENE FAMILY IN LAND PLANTS AND ITS UTILITY FOR PHYLOGENETIC ANALYSES OF FLOWERING PLANTS

by

Sarah L. Mathews

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Biological Sciences

Montana State University
Bozeman, Montana

August 1995

D378
M4257

APPROVAL

of a thesis submitted by

Sarah L. Mathews

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citation, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

July 14, 1995 — Date

Robert A Sharrock — Chairperson, Graduate Committee

Approved for the Major Department

July 14, 1995 — Date

E R Vyse — Head, Major Department

Approved for the College of Graduate Studies

8/26/95 — Date

R L Brown — Graduate Dean

EVOLUTION OF THE PHYTOCHROME GENE FAMILY IN LAND PLANTS AND ITS UTILITY FOR PHYLOGENETIC ANALYSES OF FLOWERING PLANTS

Sarah Mathews

Advisor: Robert A. Sharrock, Ph.D.

Montana State University
1995

Abstract

The phytochrome nuclear gene family encodes photoreceptor proteins that mediate diverse developmental responses to red and far red light throughout the life of a plant. From studies of the angiosperm *Arabidopsis thaliana*, the family has been modeled as comprising five loci, *PHYA-PHYE*. In gymnosperms, two loci have been detected, while in other nonangiosperm groups just one locus has been observed. A polymerase chain reaction (PCR) based strategy to sample plant DNAs was developed in order to test the gene family model inferred from *Arabidopsis* in other plant species and to evaluate the utility of phytochrome DNA sequence data for phylogenetic studies. Results presented here indicate that the *Arabidopsis* model is not completely appropriate for all

angiosperm groups because additional *PHY* loci related to *PHYA* and *PHYB* have evolved independently several times in dicot angiosperms, and monocot angiosperms may lack orthologs of *PHYD* and *PHYE*. Nonetheless, for studies of organismal evolution, the phytochrome gene family is potentially useful because the loci occur as single copy sequences, and the data suggest that the various loci are evolving independently. In two plant families, dicotyledonous Fabaceae (legumes) and monocotyledonous Poaceae (grasses), phytochrome data provided phylogenetic resolution. In addition to nucleotide substitutions, phylogenetically informative insertions and deletions characterize the phytochrome data sets. Furthermore, together with data obtained from public databases, the data detected in this study suggest that the differential distribution of phytochrome loci among flowering plant groups may be phylogenetically informative. The presence of a legume-specific locus most closely related to *PHYA* may be informative once its phylogenetic distribution is known; likewise, the apparent absence of *PHYD* and *PHYE* from monocots and some dicot plant groups potentially resolves relationships among major angiosperm lineages.

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a doctoral degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library. I further agree that copying of this thesis is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for extensive copying or reproduction of this thesis should be referred to University Microfilms International, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom I have granted "the exclusive right to reproduce and distribute my dissertation in and from microform along with the non-exclusive right to reproduce and distribute my abstract in any format in whole or in part."

Signature Sarah L. Mathews

Date 18 August 1995

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The phytochrome nuclear gene family encodes photoreceptor proteins that mediate diverse developmental responses to red and far red light throughout the life of a plant. From studies of the angiosperm *Arabidopsis thaliana*, the family has been modeled as comprising five loci, *PHYA-PHYE*. In gymnosperms, two loci have been detected, while in other nonangiosperm groups just one locus has been observed. A polymerase chain reaction (PCR) based strategy to sample plant DNAs was developed in order to test the gene family model inferred from *Arabidopsis* in other plant species and to evaluate the utility of phytochrome DNA sequence data for phylogenetic studies. Results presented here indicate that the *Arabidopsis* model is not completely appropriate for all angiosperm groups because additional *PHY* loci related to *PHYA* and *PHYB* have evolved independently several times in dicot angiosperms, and monocot angiosperms may lack orthologs of *PHYD* and *PHYE*. Nonetheless, for studies of organismal evolution, the phytochrome gene family is potentially useful because the loci occur as single copy sequences, and the data suggest that the various loci are evolving independently. In two plant families, dicotyledonous Fabaceae (legumes) and monocotyledonous Poaceae (grasses), phytochrome data provided phylogenetic resolution. In addition to nucleotide substitutions, phylogenetically informative insertions and deletions characterize the phytochrome data sets. Furthermore, together with data obtained from public databases, the data detected in this study suggest that the differential distribution of phytochrome loci among flowering plant groups may be phylogenetically informative. The presence of a legume-specific locus most closely related to *PHYA* may be informative once its phylogenetic distribution is known; likewise, the apparent absence of *PHYD* and *PHYE* from monocots and some dicot plant groups potentially resolves relationships among major angiosperm lineages.

CHAPTER 1

INTRODUCTION

Phytochromes

The phytochromes are photoreceptors for red and far-red light in all green plants and some green algae (reviewed in Quail, 1991; Furuya, 1993). Each subunit of these large cytoplasmic receptors comprises a protein of 1100 to 1200 amino acids and a covalently attached linear tetrapyrrole chromophore. Existing in two continuously interconvertible forms, Pr, the red light-absorbing form, and Pfr, the far-red light-absorbing and biologically active form, phytochromes mediate diverse developmental responses throughout the plant's life cycle. While the mechanisms whereby phytochromes participate in cellular signalling remain unknown, regions of the polypeptide required for photosensory and regulatory activities and for dimerization have been identified (Cherry et al., 1993; Edgerton & Jones, 1992; Quail et al., 1995).

Several reports have described the presence of only a single *PHY* gene in certain nonangiosperms (Hanelt et al., 1992; Kolukisaoglu et al., 1993; Morand et al., 1993; Okamoto et al., 1993; Thümmler et al., 1992; Winands et al., 1992), while evidence of two *PHY* genes is reported for other nonangiosperms. For example, Maucher et al. (1992) refer to a putative second sequence in the fern *Dryopteris filix-mas* L., although the fragment remains uncharacterized. Two unpublished *PHY* sequence fragments from *Psilotum nudum* (L.)Griseb. (GenBank accessions X74930, X74931) differ from one another in the region of overlap; and two *PHY* cDNAs from *Pinus palustris* Mill. reportedly have been cloned and partially sequenced (Furuya, 1993), while a single *PHY* cDNA from *Gingko biloba* L. is cited in the same report. However, in angiosperms, five related sequences encoding phytochrome proteins designated *PHYA-PHYE* have been characterized from *Arabidopsis thaliana* (L.)Schur (Sharrock & Quail, 1989; Clack et al., 1994). The genes for these five phytochromes have been mapped to *Arabidopsis* chromosomes 1, 2, 4, and 5, (unpublished), and no evidence for *PHY* pseudogenes was found. Homologs of *Arabidopsis PHYA* and *PHYB* have been characterized in other angiosperms (Adam et al., 1993;

Christensen & Quail, 1989; Cordonnier-Pratt et al., 1994; Dehesh et al., 1991; Hershey et al., 1985; Heyer & Gatz, 1992a, 1992b; Kay et al., 1989; Sato, 1988; Sharrock et al., 1986), as have homologs of *PHYC* and *PHYE* (Cordonnier-Pratt et al., 1994). A putative pseudogene most similar to *PHYA* has been reported in *Pisum* (Sato, 1990), and a cDNA clone from *Zea* containing a partial *PHY* fragment has been interpreted as a pseudogene (Christensen & Quail, 1989). Overall, these studies suggest that the gene family increases in complexity from nonangiosperms to angiosperms. This suggestion is consistent with data recently submitted to GenBank in which the dicot *Piper* is represented by three distinct sequences, but additional nonangiosperm taxa are represented by just single sequences (Kolukisaoglu et al., unpublished).

Nearly all *PHY* genes that are fully characterized share high sequence identity (App. A) and structural similarity with the *Arabidopsis* loci (example in Fig. 1). Peptide fragments from the nonangiosperms *Anemia phyllitidis* (L.) Sw. and *Dryopteris filix-mas* (Maucher et al., 1992) share high sequence identity with the *Arabidopsis* phytochromes in their N-termini, as do sequence fragments from the N-termini of phytochromes from the major nonangiosperm taxa (Mathews et al., 1995:App. 2); small internal PHY peptides from the alga *Mesotaenium caldariorium* (Lagerh.)Hansg. are highly similar to both N- and C-terminal peptides of other phytochromes (Morand et al., 1993). Two exceptional *PHY* genes have been described in nonangiosperms. The *PHY* gene from the alga *Mougeotia scalaris* Hässel (Winands et al., 1992) contains additional introns in the N-terminal coding sequence, and in the *PHY* gene from the moss *Ceratodon purpureus* (Hedw.)Brid., the conserved N-terminal region is combined with a highly divergent C-terminal coding region (Fig. 1) which encodes a putative light-regulated protein kinase (Thümmler et al., 1992). However, in another moss, *Physcomitrella patens* (Hedw.)B.S.G., the C-terminal coding region is similar to all other *PHY* genes (Kolukisaoglu et al., 1993). In angiosperms, the *PHYC* locus from *Arabidopsis* lacks the third intron found in all other fully characterized angiosperm loci (Cowl et al., 1994).

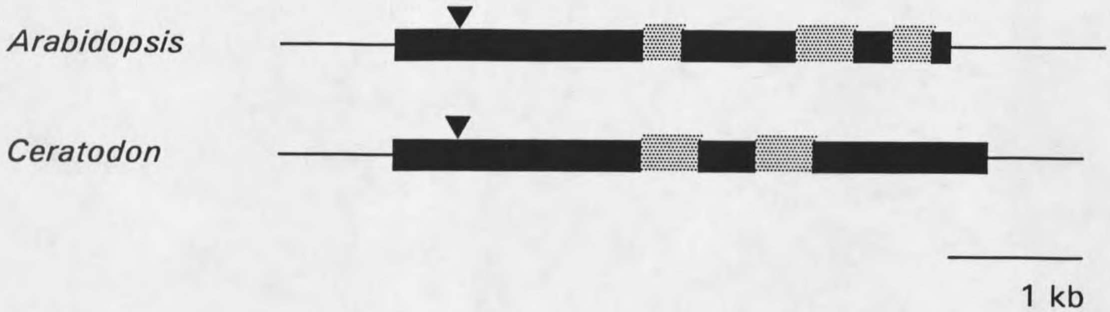


Figure 1. Phytochrome gene structure of *Arabidopsis* (Clack et al., 1994) and *Ceratodon* (Thümmler et al., 1992), from N-terminus (left) to C-terminus (right) showing untranslated regions (lines), exons (filled rectangles), introns (shaded rectangles), and the approximate site of chromophore attachment (triangle).

Phytochrome mediated responses that are characterized in green algae, mosses, and ferns include reorientation of chloroplasts (algae and ferns), rhizoid induction (algae, mosses, ferns), changes in cell membrane potential (algae and ferns) and water permeability (algae), cell elongation (ferns), reorientation of growth in protonemata in response to polarized light (mosses and ferns), germination of spores (ferns), organization of the cytoskeleton and cell cycle control (fern gametophytes), and phototropism in young leaves of some ferns (Wada & Kadota, 1989; Morand et al., 1993; Kraml, 1994; Haupt & Häder, 1994; Wada & Sugai, 1994). Neither the shade avoidance nor the deetiolation responses that are important in angiosperms have been detected in nonflowering plants (Smith, 1994).

In flowering plants, phytochrome mediated responses include germination, seedling hypocotyl elongation, stem cell differentiation, plastid development, flavonoid pigment synthesis, and floral induction in response to photoperiod. Modulation of plant gene expression by phytochrome is well documented (Batschauer et al., 1994). In the dicot *Arabidopsis*, the *PHYA-E* genes are differentially expressed in response to the light environment (Sharrock & Quail, 1989; Somers et al., 1991; Clack et al., 1994), and unique, contrasting photosensory roles are attributed to *PHYA* and *PHYB* that cannot be accounted for by *PHYC*, *PHYD*, or *PHYE* (summarized in Quail, 1995).

Goals of this project

1. The primary model plant species used for investigations of phytochrome function in angiosperms is *Arabidopsis thaliana*. A critical consideration in evaluating the importance of the *PHY* photoreceptor family to plants in general is whether homologs of the various *PHY* genes are present in a wide variety of flowering plant species. Therefore the first goal of this project was to develop a method for detecting the various *PHY* genes in genomic DNA from diverse plant species, to clone and

sequence portions of those genes, and subsequently to estimate the relationships among the detected *PHY* genes using phylogenetic analyses.

2. The functional divergence observed among PHY loci in *Arabidopsis*, together with high sequence divergence (approximately 50% among the *PHYA*, *PHYB*, and *PHYC* loci) suggests that nonhomologous recombination is infrequent among *PHY* genes of *Arabidopsis*. If the loci are evolving independently, distinguishing orthologous genes from paralogous genes should not be difficult, thus predicting that these molecules might be useful tools for plant phylogeneticists (see MATERIALS AND METHODS). The second goal of this project was to use the phytochrome sequence data obtained to test the hypothesis that the genes are independently evolving and to ascertain the taxonomic level at which *PHY* data might be useful for phylogenetic studies of flowering plants.

CHAPTER 2

MATERIALS AND METHODS

DNA sequence detection

Total DNA was isolated from fresh, lyophilized, or dried herbarium material of taxa listed in Table 1 by standard methods (Doyle & Doyle, 1987). Aliquots were extracted once with phenol:chloroform-isoamyl alcohol (1:1 volume), and the aqueous portions were purified over sepharose CL-6B (Pharmacia, Piscataway, NJ) columns. DNAs were sampled from different subclasses of angiosperms (sensu Cronquist, 1981) and, two angiosperm plant families, Fabaceae (legumes) and Poaceae (grasses) were extensively sampled; to assess phytochrome gene and nucleotide diversity in Poaceae, DNAs were sampled from genera of the five major subfamilies Arundinoideae, Bambusoideae, Chloridoideae, Panicoideae, and Pooideae. DNAs of *Bambusa*, *Dianthus*, *Equisetum*, and *Quercus* were kindly provided by Elizabeth

Kellogg, Randy Woodson, Pamela Soltis, and Paul Manos respectively, and DNAs of *Flagellaria*, *Joinvillea*, and *Thamnochortus* were kindly provided by Jerrold Davis.

Table 1. Sources of *PHY* sequences determined in this study. Taxonomic arrangement follows Cronquist (1981) and Clayton & Renvoise (1986).

Subclass	Species	Source/Voucher
Sphenophyta	*Equisetum arvense* L.	P. Soltis (no voucher)
Pinophyta	*Gingko biloba* L.	*S. Mathews 365* MONT
	Pseudotsuga menziesii (Mirb.) Franco	*S. Mathews s.n.* MONT
Magnoliophyta		
Dicots		
MAGNOLIIDAE	*Ceratophyllum demersum* L.	*S. Mathews s.n.* MONT
	Aquilegia L. sp.	S. Mathews (no voucher)
HAMAMELIDAE	*Urtica dioca* L.	*S. Mathews 330* MONT
	Quercus turbinella Greene	*J. M. Tucker 4491* UCD
CARYOPHYLLIDAE	*Dianthus caryophyllus* L.	R. Woodson (no voucher)
	Spinacia oleracea L.	S. Mathews (no voucher)
DILLENIIDAE	*Arabidopsis thaliana* (L.) Schur	S. Mathews (no voucher)
ASTERIDAE	*Lycopersicon esculentum* Mill.	S. Mathews (no voucher)
	Antirrhinum majus L.	*S. Mathews 301* MONT
ROSIDAE	*Daucus carota* L.	S. Mathews (no voucher)
Monocots		
ALISMATIDAE	*Elodea* Michx. sp.	S. Mathews (no voucher)
ARECIDAE	*Lemna gibba* L.	Silverthorne (no voucher)
ZINGIBERIDAE	*Billbergia nutans* H. Wendl	*S. Mathews 351* MONT
LILIIDAE	*Muscari* Mill. sp.	S. Mathews (no voucher)
COMMELINIDAE		
Flagellariaceae		
	Flagellaria indica	*J. I Davis s.n.* BH
Joinvilleaceae		
	Joinvillea ascendens Gaudich.	*J. I Davis s.n.* BH
Restionaceae		
	Thamnochortus P. Bergins sp.	*J. I Davis s.n.* BH
Poaceae		
ARUNDINOIDEAE		
	Aristida purpurea Nutt.	*Lavin s.n.* MONT
	Danthonia unispicata (Thurber) Munro ex Macoun	*Lavin s.n.* MONT
	Phragmites australis (Cav.) Trin. ex Steudel	*Lavin s.n.* MONT
BAMBUSOIDEAE		
	Bambusa Schreb. sp.	*E. A. Kellogg V6 A*
	Oryza sativa L.	*Lavin s.n.* MONT
CHLORIDOIDEAE		
	Bouteloua gracilis (Willd. ex H.B.K.) Lag. ex Griffiths	*Lavin s.n.* MONT
	Calamovilfa longifolia (Hook.) Scribn.	*Lavin s.n.* MONT

Table 1, continued.

Eragrostis cilianensis (All.) Mosher	*Lavin s.n.* MONT
PANICOIDEAE	
Panicum capillare L.	*Lavin s.n.* MONT
Pennisetum setaceum (Forsk.)Chiov.	*Lavin s.n.* MONT
Zea mays L.	*Lavin s.n.* MONT
POOIDEAE	
Bromus inermis Leyss.	*Lavin s.n.* MONT
Poa pratensis L.	*Lavin s.n.* MONT
Stipa viridula Trin.	*Lavin s.n.* MONT

A region of the *PHY* gene that encodes a peptide including and proximal to the chromophore attachment site was amplified using PCR, resulting in a target of 270-350 bp (See Appendix A for region of amplification). Oligonucleotides with equimolar mixtures of nucleotide pairs at two-fold degenerate sites and inosines (I) at three- to four-fold degenerate sites were designed to amplify all possible target sequences in template DNAs flanked by the conserved upstream peptide HYPATDIP (5'-CAYTAYYCIGCIACIGAYATHCC-3') and downstream PFPLRYAC (5'-CRCAIGCRTAICKIARIGGRWAIGG-3'). These peptide sequences are conserved in all *Arabidopsis* phytochromes and in the amino acid sequences inferred from other fully sequenced dicot and monocot genes, and they flank a region comprising variation likely to be phylogenetically informative. Conversely, to restrict the number of phytochrome loci amplified in single experiments locus-specific downstream oligonucleotides were paired with HYPATDIP in some experiments. These included *PHYA*-specific 5'-ACRTGIAYIGCRAAIACYTGIGC-3' at AQVFAI/VHV, *PHYB*-specific 5'-ASYTGIARICCRAAIGCYTGCAT-3' at MQAFGLQL, and *PHYC*-specific 5'-ATYTGIACICCRAAIACYTGIGT-3' at TQVFGVQI. Additionally, oligonucleotides at the upstream peptide GYDRVM (5'-GGNTAYG AYMGNGTNATG-3') and the downstream peptide KVLDMI (5'-YTTNAC NARRTCCATDAT-3') were designed to amplify a larger *PHY* fragment (ca. 600 bp) inclusive of the target sequences detected in these investigations. Used in combination with HYPATDIP and PFPLRYAC, GYDRVM and KMLDMI potentially provide the opportunity to detect loci that are mutated at HYPATDIP or PFPLRYAC; however, this strategy was not fully tested in this study. Standard PCR protocols (Perkin-Elmer, Norwalk, CT) were modified to include an initial 5 cycles in which annealing temperatures were less stringent (e.g., 45-49^{o} C).

The PCR products were converted to blunt-end fragments with T4 DNA polymerase (BRL, Gaithersburg, MD) and were ligated to *Eco*RV-cut bacteriophage M13KRV8.2. M13KRV8.2 carries an *Eco*K cassette that facilitates screening of nonrecombinants in an *E. coli* strain which is $r_k^+m_k^+$ (Waye et al., 1985). Transformation of *E. coli* with the ligation product yielded a population of M13*PHY* clones containing amplified genomic *PHY* sequences. Individual clones were cultured, and double-stranded phage DNA was isolated from bacterial pellets by alkaline-lysis minipreparation. Inserts cut from M13 vectors using *Eco*RI and *Hind*III were resolved on 3% NuSieve (FMC, Rockland, ME), or 2% standard, agarose gels. Up to 108 clones were screened per individual DNA and in many cases, restriction enzyme digestion of PCR products or clones was used to aid in detection of less abundantly amplified or cloned sequences. Single-stranded DNAs for Sanger dideoxy sequencing (Sequenase version 2.0, USB, Cleveland, OH) were isolated from recombinants carrying putative *PHY* inserts. In most cases, sequences of both orientations were determined, and multiple PCR products from two accessions or genera were sequenced to detect possible contaminants and PCR errors.

Alignment of DNA Sequences

Peptide sequences were multiply aligned using ALIGN (Scientific & Education Software, State Line, PA) and GDE 2.2 (Steven Smith & University of Illinois) and were adjusted by eye at sites that were not accurately aligned by the computer algorithms; peptide alignments were the basis for multiple nucleotide sequence alignments. Appendices A through F comprise the peptide and nucleotide alignments. For sequence comparisons, just gaps that could be identified as homologous were retained and were coded as single characters. All other gaps were deleted from the data matrices, as were nonhomologous 3′ and 5′ nucleotide sites. Four sequences that were included in the full length nucleotide data set (App. B) were not included in Appendix A because they were obtained later and did not significantly alter the consensus sequence. They are the *PHY* sequence from *Psilotum* (GenBank accession X74931), the *PHY* sequence from the moss *Physcomitrella* (Kolukisaoglu et al., 1993), and the *PHY* sequences from the angiosperm *Nicotiana* (GenBank accessions X66784, L10114). To assess phytochrome diversity in early land plants, DNA sequences from different nonangiosperm phyla available from GenBank were aligned with angiosperm sequences (App. C).

Phylogenetic analyses

Phylogenetic analyses of the phytochrome data were used to evaluate the relationships among newly obtained sequences and the genes characterized from *Arabidopsis*. The orthology of fully sequenced *PHY* genes from various species to individual *PHY* loci of *Arabidopsis* commonly has been established by overall similarity (Dehesh et al., 1991; Heyer and Gatz, 1992a, 1992b; Quail, 1991; Furuya, 1993). Similarities in gene expression and regulation have been used secondarily to imply orthology (Furuya, 1993). However, overall similarity may not reflect phylogeny, and phylogenetically related loci may differ in function due to mutations in *cis*-regulatory regions (e.g., Doyle, 1991; Li & Noll, 1994). Since orthology is best determined by shared ancestry, as evidenced by homologies, cladistic analysis was used to determine the orthology of all available full length *PHY* sequences to those characterized from *Arabidopsis*. Likewise, evidence from cladistic analysis was used to assess orthology of the sequences detected in this study to the loci from *Arabidopsis*.

The assumption of phylogenetic analysis is explicit: given the evolutionary model that sequences diverged from a common ancestor by descent with modification, the goal is to discern the branching pattern among all sequences under consideration, successively grouping those together that most recently shared a common ancestor. Specifically, results presented here are phylogenies inferred from the DNA sequences by maximum parsimony analysis, with one exception (see below). Starting with a raw data matrix of sequences, aligned such that homologous nucleotide sites form columns of characters and genes (or taxa) are rows, maximum parsimony trees are constructed by optimizing the distribution of character state changes (nucleotide mutations) on a tree such that the fewest changes are required; the minimal length (ML) tree is chosen as the best explanation of the character data, and thus, the best estimate of phylogenetic relationships among the genes. Maximum parsimony algorithms ignore characters for which all taxa share the same character state because they do not provide specific evidence of relationship among subsets of the observed taxa; such characters are said to be uninformative. Conversely, if taxa A, B, and C in a matrix share the character state adenine at a nucleotide site, and D and E share a guanine at the same position, that character is informative in that it provides evidence that A, B, and C are most closely related to one another. Thus, maximum parsimony analysis discriminates between two types of similarity observed among taxa, overall similarity that reflects ancestral states and derived similarity that is due to more recent common ancestry, and to use only the latter in formulating historical hypotheses. Throughout this thesis, the terms "maximum parsimony" and "cladistic" will be used interchangeably. Indices that accompany the maximum parsimony trees that are presented here provide an estimate of how many of the changes on the tree are due to the independent evolution of the same character state (homoplasy) and how many are inferred to be unique, and thus a sound basis for inferring homology; these include the CI (consistency index, Kluge & Farris, 1969; Farris, 1989), RI (retention index, Farris, 1989), and the RC (rescaled consistency index, Farris, 1989). CI is the minimum possible number of steps over all characters divided by the actual number inferred from the tree; CI = 1.0 if the actual number of steps equals the minimum number possible, if the characters in a data set are perfectly congruent with each other and the tree; the CI is thus an expression of the amount of homoplasy as a proportion of total change. Alternatively, the RI expresses the observed amount of homoplasy as a proportion of total possible homoplasy. The RC is the product of the CI and RI. When more than one ML tree is found for the data observed in this study, strict consensus trees that include only those components that occur in all ML trees (Sokal & Rohlf, 1981) are used to reveal consistently resolved groups.

In contrast to trees from maximum parsimony analysis, phylogenetic trees derived from matrices of pairwise distances or from maximum likelihood methods do not discriminate between ancestral and derived similarity. However, all methods make assumptions about evolutionary change, either that substitution frequencies fit a very specific model (maximum likelihood and distance methods), or that characters are evolving independently (maximum parsimony). Criteria for comparing the accuracy of the methods (e.g., Hillis, 1995) include consistency (the correct tree is converged upon as the data become infinite), efficiency (a measure of how quickly a method converges upon the correct tree as more data are available for analysis), and robustness (the degree to which performance of the method is affected by violations of the assumed model). These criteria have been applied to results from computer simulations (e.g., Nei, 1991; Kuhner & Felsenstein, 1994; Huelsenbeck, 1995), providing information about the expected performance of the different methods under idealized conditions; the results are predictions about how the methods will perform during analysis of real data sets. A general conclusion from simulation studies and from tests of their predictions in investigations of known phylogenies, is that parsimony, maximum likelihood, and additive distance methods perform similarly, especially when the data are corrected for multiple substitutions per nucleotide site (e.g., Huelsenbeck, 1995). Parsimony methods are known to be misleading when parallel changes among sequences exceed informative nonparallel changes (Felsenstein, 1978) because long branches attract one another. Thus, the more conservatively evolving amino acid characters were analyzed in comparisons of nonangiosperm with angiosperm phytochrome sequences and hypotheses inferred from distance analyses are included.

Maximum parsimony algorithms used for sequence comparisons were those available in PHYLIP 3.5c (Felsenstein, 1993), Hennig86 (Farris, 1988), PAUP 3.1 (Swofford, 1993), and RNA (Farris, 1994). Minimal length trees resulted from heuristic search options available in either Hennig86 (mh*, bb* with no upper limit set), PHYLIP (DNAPARS), or in PAUP (CLOSEST or RANDOM data addition sequence, HOLD option set for 5 trees when applicable, STEEPEST DESCENT, MULPARS, and TBR branch swapping options activated, with branch swapping on nonminimal trees, and MAXTREES set at 10,000 or 20,000). The PROTDIST and NEIGHBOR algorithms in PHYLIP 3.5c were used to reconstruct a phylogeny from pairwise distances among amino acid sequences from nonangiosperms and angiosperms; and the DNAML algorithm in the same program was used to infer a maximum likelihood phylogeny from the grass data. Graphical output of trees is from COMPONENT (Page, 1993a) and PAUP 3.1

(Swofford, 1993).

For the cladistic analysis of the full length sequences, trees were rooted by designating *PHY* sequences from *Physcomitrella*, *Selaginella* and *Adiantum capillus-veneris* L. (Okamoto et al., 1993) as the outgroups, because they are the only fully characterized *PHY* genes from nonangiosperms. For analysis of partial sequences in angiosperms, *Selaginella* was retained as an outgroup, along with the *PHY* sequences from the gymnosperms *Gingko* and *Pseudotsuga* that were determined during this analysis. For cladistic analysis of sequences from grass genera, trees were rooted by designating *PHY* sequences from *Flagellaria indica, Joinvillea ascendens* Gaudich. and *Thamnochortus* P. Bergins as outgroups; these taxa represent families inferred from morphological (Campbell and Kellogg, 1987; Linder and Rudall, 1993) and molecular data (Doyle et al., 1992) to be among the closest relatives of Poaceae. Cladograms rooted at *Muscari* Mill. allow detection of phylogenetic structure within outgroup species, but do not differ in the details of grass relationships from those rooted at *Joinvillea*. The *PHY* sequence from *Selaginella* was designated as the outgroup for the cladogram of all *PHY* sequences from grass genera because it is not likely to be more closely related to one angiosperm *PHY* paralog than to another. Finally, analyses that addressed relationships among nonangiosperm and angiosperm *PHY* loci were rooted by designating *Mougeotia* as the outgroup.

Tests of phylogenetic accuracy

Phylogenetic hypotheses inferred from phytochrome data were evaluated for robustness using a subset of the statistical and congruence approaches that are available (e.g., Felsenstein, 1988; Hillis et al., 1993; Li & Zharkikh, 1995); support for monophyly of clades was evaluated using the bootstrap resampling, Bremer support, and total support tests described below. Congruence approaches were used to evaluate the agreement among individual *PHY* gene trees and among *PHY* gene trees and trees inferred from other data sets.

Statistical approaches

Bootstrap resampling. The use of the non-parametric bootstrap resampling technique to place confidence limits on phylogenies was proposed first by Felsenstein (1985); it is perhaps the method most frequently used by systematists to assess the robustness of phylogenetic hypotheses. The test "involves inferring the variability in an unknown distribution from which your data were drawn by resampling from the data" (Felsenstein, 1985:784). A single bootstrap sample in a test of a phylogenetic hypothesis maintains the original set of species, but draws characters with replacement from the original matrix; consensus trees are used to show the monophyletic groups that occur in a majority of, for example, 500 or 1000 bootstrap replicates. The technique relies on several assumptions (Felsenstein, 1985) that are probably reasonable for DNA sequence data. A more serious limitation is that for proper hypothesis testing a null model should be specified in advance. However, in phylogenetic studies, the null hypothesis is the topology that has been inferred from a data sample (Felsenstein, 1985; Li & Zharkikh, 1994). The bias that this introduces to the outcome of bootstrap testing has been explored (Li & Zharkikh, 1994, 1995) and the CP (complete-and-partial) bootstrap has been developed to compensate for bias. The method is not yet available (Li & Zharkikh, 1995), thus the bootstrap values reported in this investigation are uncorrected.

Permutation methods. Maddison and Slatkin (1991) suggested that the appropriate null model for a statistical test of a "known" tree (i.e., the tree inferred from observed data) is one in which characters are randomized. The PTP (Faith & Cranston, 1991) and total support (Källersjö et al., 1992) tests compare the observed data to randomizations of those data; character states are randomly reassigned to taxa in the observed data such that congruence among character state distributions is produced by chance alone. The PTP measures character congruence (assumed to result from common ancestry) by comparing maximum parsimony trees from the randomized and observed data; minimal length trees inferred from randomized data are expected to be longer than trees inferred from observed data because randomized data sets should comprise fewer nonrandomly covarying characters. In preliminary analyses of phytochrome DNA data, all data sets were shown to be significantly structured based on results of PTP tests. However, Källersjö et al. (1992) demonstrated that the PTP can imply significant structure in ambiguous data. Thus, further permutation tests of the phytochrome data sets were performed using the total support test.

The total support test measures departure from random character covariation differently than the PTP; first the

Bremer support (Bremer, 1988) for each group in the observed tree is calculated (i.e., the number of steps that must be added to a tree before the group is lost in a strict consensus tree). Total support is the sum of group supports, which is assumed to be greater in well-structured data than in randomized data. The error rate on concluding that a data set is significantly structured is the upper tail probability $\alpha'_s = (X + 1)/(W + 1)$, where X is a number of W total randomizations yielding total support no less than that of the observed data (Källersjö et al., 1992). The phylogenetic program RNA (Farris, 1994) was used to calculate group support values. Additionally, Bremer support was investigated manually for some phylogenies by examining all trees up to ten or twenty steps longer than the minimal length tree(s).

Congruence approaches

Congruence approaches potentially address special concerns associated with inferring organismal relationships from molecular phylogenies. Various biological processes such as introgressive hybridization and lineage sorting from polymorphic ancestry may result in discordance among gene trees and/or among gene and species trees (e.g., Harrison et al., 1987; Rieseberg & Brunsfeld, 1992; Soltis et al., 1992). Such differences also may result from lack of resolution in one of the data sets (e.g., Olmstead, 1989; Olmstead & Sweere, 1994), or from mistaken orthology (e.g., Goodman et al., 1979; Doyle, 1992). Thus, determining organismal relationships requires that evolutionary hypotheses derived from single genes be tested with further data (e.g., Pamilo and Nei, 1988; Takahata, 1989), as well as methods for reconciling differences.

Congruence approaches assume that similar patterns of relationships observed among trees derived from multiple independent data sets are evidence of both the veracity of the shared components and the accuracy of the phylogenetic method (summarized in Hillis, 1995). The debate between advocates of combining all data in a single analysis, the total evidence approach (Kluge, 1989), and summarizing topological features of trees derived from data partitioned into different types in a consensus tree (Adams, 1972; Carpenter, 1988) is ongoing (e.g., Barrett et al., 1991; de Quieroz, 1993; Chippindale & Wiens, 1994; Page, 1990, 1993b, 1994). The major argument against combining data for analysis is that subsets of characters may have been subject to different evolutionary processes (e.g., Bull et al., 1993). However, character weighting schemes can be used to incorporate assumptions about evolutionary models (e.g., Chippindale & Wiens, 1994). Advocates of combining data for analysis object to consensus techniques because information about individual results is lost. A distinct advantage of combining the data is that it allows character congruence (the degree to which all available characters make a unified, internally consistent statement about relationships, Swofford, 1991:314) rather than taxonomic congruence to be evaluated. Nonetheless, consensus trees are useful for expressing areas of agreement among trees and need not be viewed as phylogenetic hypotheses. Furthermore, discordance in results from separate analyses may be informative regarding nonindependence of characters (e. g., Swofford, 1991). Thus, it seems sensible to do both when possible (e.g., Doyle et al., 1994; Olmstead & Sweere, 1994).

Combining data. Phytochrome data from individual loci were not combined in broad comparisons (i.e., those comprising sequences from all angiosperm subclasses) because the main goal was to assess homology of individual sequences to *PHY* loci from *Arabidopsis*, and to assess the degree to which the phytochrome family comprises monophyletic gene lineages. However, in order to compare the degree and strength of resolution in phylogenies comprising phytochrome sequences from grass genera, the data were analyzed both separately (with individual loci comprising individual terminals in a data set) and together (with data from all loci that were sampled combined for each genus).

Consensus analysis. As noted above, strict consensus trees were used to combine multiple minimal length trees from individual parsimony analyses. Furthermore, because the phytochrome data from grass genera were used to infer a species phylogeny, consensus techniques available in the computer package COMPONENT (Page, 1993a) were used to measure agreement among the phylogeny inferred from phytochrome data and grass phylogenies inferred from other data sets. For example, they were used to compare species phylogenies from phytochrome, *rbc*L, chloroplast DNA (cpDNA) restriction site variation, and morphological data from grasses.

Combining trees. In response to the suggestion that it is desirable to use as many genes as possible to infer an organismal phylogeny from molecular data (Pamilo & Nei, 1988; Takahata, 1989), or to combine molecular with morphological data when possible (e.g., Hillis, 1987), Baum (1992) proposed a protocol for combining the trees from different analyses rather than combining the data. Following the method that Brooks (1981, 1990) proposed for recoding trees as single multistate characters in order to study coevolution, additive binary coding matrices are derived for single trees and subsequently combined for cladistic analysis. Doyle (1992) proposed a similar approach because he postulated that genes might behave as single characters rather than as a set of independent nucleotide characters; according to Doyle, such a set of nonindependent nucleotide characters could potentially "swamp" the signal from morphological data in a set of combined molecular and morphological data. Phylogenies inferred from phytochrome, *rbc*L, chloroplast DNA (cpDNA) restriction site variation, and morphological data from grasses were compared in this manner.

Tree mapping. The assumption of parsimony analyses that evolution is divergent is violated by convergence through such events as nonhomologous recombination among related loci. Of special concern relative to using sequences from a multigene family in phylogenetic reconstruction are potential problems related to concerted evolution (sensu Zimmer et al., 1980). For example, an analysis of *rbc*S nucleotide sequences (Meagher et al., 1989) indicated that gene conversions among *rbcS* loci have occurred in each genus examined, leading to regions of "partial homology" (Patterson, 1987) in a locus and thus, to the possibility of mistaken orthology of genes. Gene conversion involving a complete locus is a gene loss because one gene is lost at the expense of another; loss of a gene through other nonhomologous recombination events, or through gene inactivation, also may result in inadvertent comparison of paralogous sequences. Comparison of paralogous rather than orthologous sequences potentially results in discordant gene and species phylogenies (Goodman et al., 1979; Doyle, 1992). Sanderson and Doyle (1992), however, suggest that it is possible to reconstruct a reliable organismal phylogeny from DNA sequences of multigene families in which concerted evolution is infrequent, and preliminary data indicate that nonhomologous recombination events are infrequent among phytochrome genes, (Sharrock & Quail, 1989; Dehesh et al., 1991; Heyer & Gatz, 1992a, 1992b; Clack et al., 1994; Adam et al., 1993). Nonetheless, tree mapping procedures available in COMPONENT (Page, 1993a) based on the hemoglobin phylogenetic model of Goodman et al. (1979), which evaluate whether incongruence of gene and species trees could be due to sampling error (Page, 1993b, 1994), were used to compare the phylogeny of phytochrome sequences from nonangiosperms and angiosperms with a green plant phylogeny from other data (Donoghue, 1994).

Weighting

Schemes for differentially weighting characters during phylogenetic analysis are used to incorporate explicit assumptions about character evolution. For example, parsimony algorithms used to search the phytochrome amino acid data set incorporated a step matrix that allows only amino acid changes that are consistent with the genetic code, or that specifies the number of steps required to change from one amino acid to another amino acid consistently with the code (Felsenstein, 1993). Likewise, assigning third codon position nucleotides a weight of zero during phylogenetic analyses is based on the assumption that these sites are more likely to have undergone multiple substitutions, and thus are possibly less informative due to parallel mutations.

In the cladistic analyses of nucleotide data from the broad survey of angiosperms and from legumes (i.e., those of Mathews et al., 1995), first, second, and third codon positions were equally weighted for the following reasons. First, empirically determined transition/transversion ratios were close to 1.0 for most comparisons. Second, results from cladistic analyses under certain differential weighting schemes are apparently the same as those from analyses under equal weighting schemes when taxonomic sampling is adequate (Albert et al., 1993; Cracraft & Helm-Bychowski, 1991). Third, all codon positions may exhibit similar levels of homoplasy (see Chase et al., 1993). Thus a rationale for excluding or differentially weighting codon positions is difficult to define. In the analyses of Mathews et al. (1995) third codon positions, and perhaps many of the synonymous substitutions, were determined by bootstrap resampling analyses to be phylogenetically very informative, with confidence intervals for just the third codon position of between 90-100%, or at least as high as the values obtained for the 1st or 2nd position.

While the observed transition to transversion ratio was used by Mathews et al. (1995) as a rationale for not giving transversion mutations greater weight than transition mutations in parsimony analyses, it might more properly be interpreted as evidence that nucleotide sites, possibly many of the third codon position sites, have undergone multiple substitutions (see next paragraph). Furthermore, it should be noted that when highly divergent sequences are compared, the bootstrap is not necessarily a valid method for evaluating the informativeness of codon positions because a statistically significant result can be obtained for an erroneous phylogeny (Felsenstein, 1985; Li & Guoy, 1990).

Nucleotide sites that have undergone multiple substitutions may be less informative phylogenetically (e.g., Mindell, 1991) than those sites that have undergone fewer changes, thus, providing the rationale for transversion parsimony, which assigns a weight of zero to transition substitutions. It is further suggested that calculation of just third codon position mutation ratios is a useful test for saturation of transition substitutions (e.g., Hillis et al., 1993; Mindell et al., 1995) because it is expected that if nucleotide sites are saturated with change the number of inferred transition mutations among sequences compared will equal the number of transversions (but see Holmquist [1983] for an argument that this is an overly simple assumption). Therefore, if as many third codon position transitions are observed in pairwise comparisons of sequences from closely related grass taxa as in comparisons of sequences from grass taxa and their outgroup species, these sites putatively are saturated with mutations and should be given less weight in phylogenetic analyses. Comparisons of third codon position mutations in *PHY* sequences from grass and outgroup genera did not indicate that a such a weighting scheme was called for.

However, contrary to the observations of Chase et al. (1993) that all codon positions may be similarly homoplasious, removing third codon position nucleotides from the grass data set in which *PHY* sequences were combined resulted in the CI changing from 0.56 to 0.62, thus indicating that there is slightly less homoplasy in first and second codon position nucleotides in this data set. Results from analyses applying both weighting schemes are presented.

Absolute and relative evolutionary rates

Evolutionary rates of phytochrome sequences were investigated to address two questions. First, at what taxonomic level are the phytochrome data likely to be useful? Calculation of absolute evolutionary rates provides an estimate of the taxonomic level at which the data will be most reliable for phylogenetic inference (e.g., Ritland & Eckenwalder, 1992). Second, are phytochrome sequences evolving in a clocklike fashion? Unequal evolutionary rates are a confounding factor in parsimony analysis because they can result in spurious attraction of long branches (Felsenstein, 1978; Hendy & Penny, 1989).

The proportions of Jukes-Cantor corrected synonymous or nonsynonymous, and Kimura 2-parameter, differences within and among gene lineages were estimated using the program

MEGA (Kumar et al., 1993). Absolute evolutionary rates were calculated by the method of Kimura (1981), in which the base substitution rate per nucleotide site per year, k, equals $K/2T$, where K is the substitution rate observed between sequences and T is divergence time. Relative evolutionary rates were estimated by the method of Wu & Li (1985), in which the rates of nucleotide substitutions in two lineages of interest are compared relative to a reference sequence. The sequences in a relative rate test are related as two sister taxa (e.g., PHY1 and PHY2) with an outgroup taxon (reference sequence). The distances between PHY1 and the reference (d_{13}) and PHY2 and the reference (d_{23}) are compared and the standardized normal test is used to assess the significance of the difference. A minimum of twenty substitutions between two sequences is required for the test to be sensitive; an alternative, more sensitive maximum likelihood relative rate test (Muse & Weir, 1992) was not implemented because observed substitutions in pairwise comparisons exceeded twenty.

CHAPTER THREE

RESULTS

Phylogenetic analyses

Analysis of full length sequences

A single most parsimonious tree (Fig. 2) was generated in analysis of the fully sequenced phytochrome genes and it resolved the following monophyletic clades with strong (90-100%) bootstrap support: all monocot *PHYA*s, all dicot *PHYA*s, all *PHYA*s, all *PHYA*s + *Arabidopsis PHYC*, just *PHYB* and *PHYD* of *Arabidopsis*, just *PHYB*s and *Arabidopsis PHYD*, *Arabidopsis PHYE* + all *PHYB*s and *Arabidopsis PHYD*, all angiosperm *PHY*s, all angiosperm *PHY*s + *Psilotum*, and angiosperm *PHY*s + *Psilotum* + *Adiantum*. Seventy-eight trees were found by keeping all trees that were ≤ 30 steps longer than the most parsimonious one; all clades were retained in all trees that are 20 steps longer, except for *Arabidopsis PHYC* + all

*PHYA*s. The two trees that were one step longer than the minimal length tree varied in their placement of *PHYC* as the sister group of either the *PHYA* or *PHYB/D/E* clade. These results thus suggest that, for example, the dicot and monocot *PHYA*s are orthologous, as are the dicot and rice *PHYB*s. Additionally, evidence is provided for the sister group relationship of *PHYE* with *PHYB* + *PHYD*, and for a later duplication giving rise to *Arabidopsis PHYB* or *PHYD*.

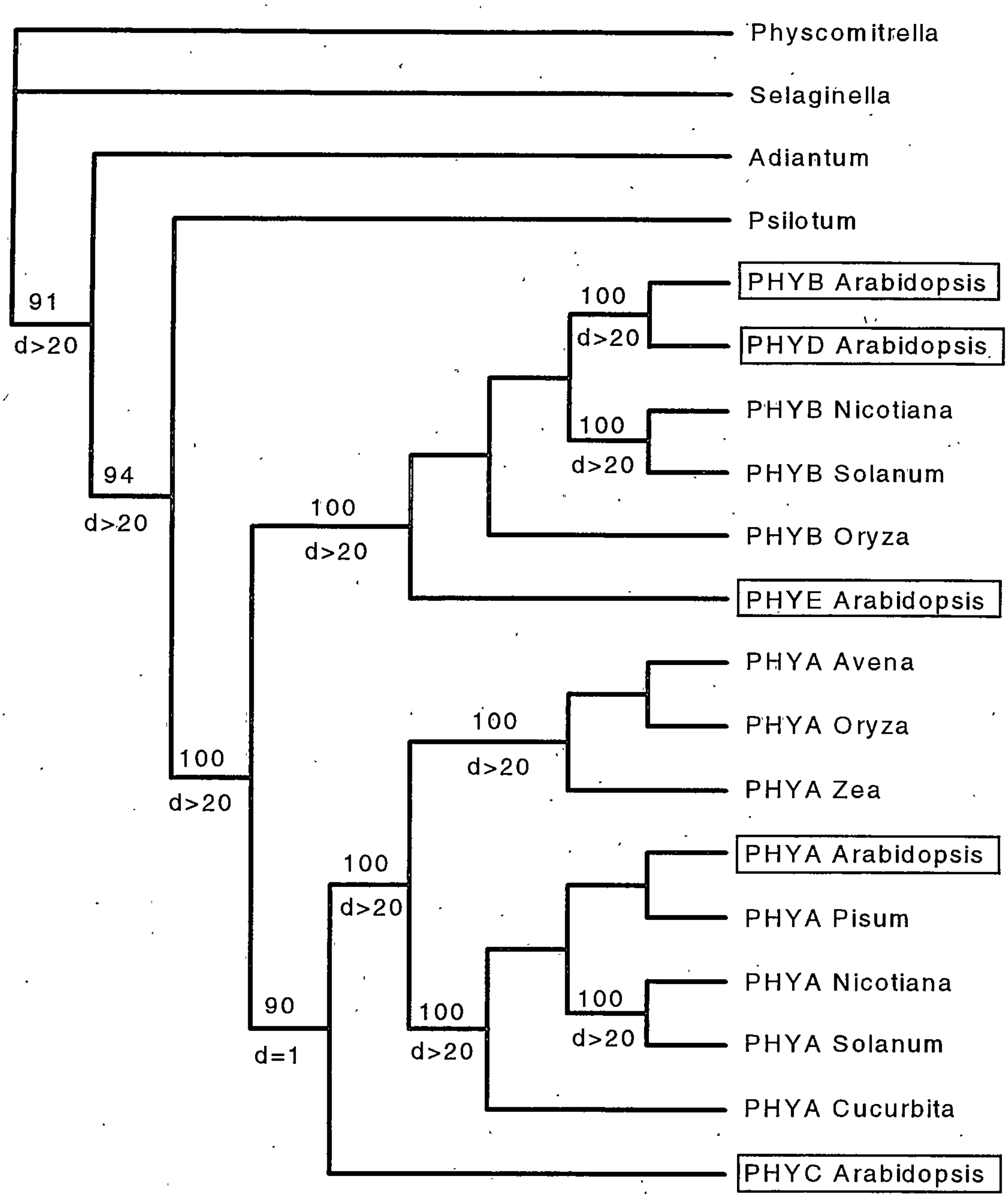


Figure 2. Single most parsimonious tree from analysis of 2637 variable nucleotide sites from the full length phytochrome sequence data set (App. A). The length is 11376, the CI=0.459, & the RI= 0.502. Bootstrap (from 500 replications) and Bremer support values are included on the best supported clades. The *Nicotiana* sequences were obtained from GenBank accessions X66784 & L10114.

<u>Analysis of sequences detected in this study</u>

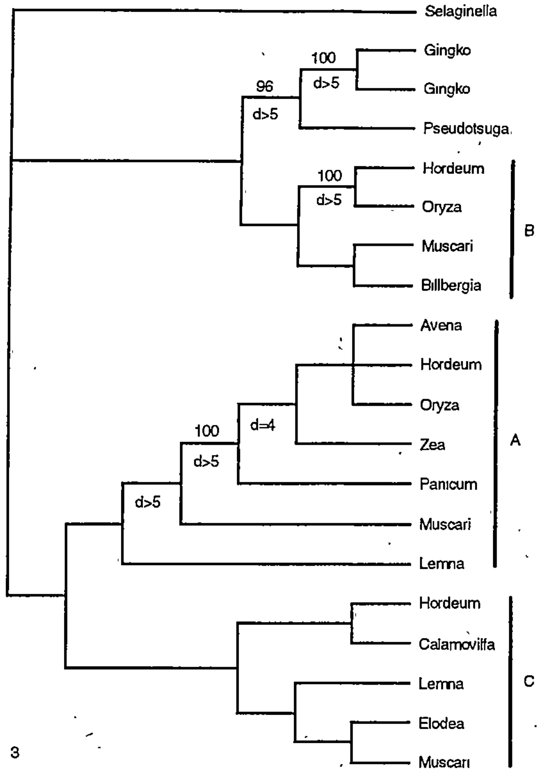
Using highly degenerate primers (HYPATDIP and PFPLRYAC, see MATERIALS AND METHODS) and amplification by PCR, target sequences from all five *Arabidopsis* genes, as well as from multiple *PHY* genes of other angiosperms, were recovered in single cloning experiments (App. D). Single *PHY* sequences were obtained from the nonangiosperms *Equisetum* and *Pseudotsuga* and two were obtained from *Gingko*. In tests of the locus-specific amplification primers (see MATERIALS AND METHODS), exclusively *PHYB*-related sequences were amplified from *Arabidopsis*, *Daucus*, *Quercus*, and *Spinacia* using the *PHYB/D/E*-specific amplification primer, and orthologs of just *PHYA* and *PHYC* were amplified from *Arabidopsis* using the *PHYA*- and *PHYC*-specific oligonucleotides, respectively. Inserts varied from 270 to 350 bp, and a region of insertion and deletion corresponding to residues 398 to 415 (App. A) was eliminated from broad comparisons because nucleotide site homologies could not be determined. However, this region could be retained in narrower comparisons, where site homologies were more readily established, as in the Fabaceae (Mathews et al., 1995) and Poaceae data sets (App. E).

Similarly to the analysis of full length sequences, angiosperm sequences determined in this study were cladistically analyzed to determine their orthology to the *PHY* loci of *Arabidopsis* (Figs. 3-5). Each sequence occurred in a monophyletic clade that included a single, specific *PHY* locus of *Arabidopsis*, providing evidence for distinct *PHY* subfamilies. Retention of a clade in a strict consensus tree (Figs. 4,5) resulting from the mhennig and branch-and-bound search options in Hennig86 or from heuristic options available in PAUP (see above), was considered as good evidence of monophyly. Results from bootstrap resampling and total support analyses revealed that some clades were strongly supported (>95%, d>5-20), and that all data sets were characterized by significant phylogenetic signal ($\alpha' = 0.001$).

The *Arabidopsis PHYA* sequence was included in a distinct monophyletic lineage in the dicot cladogram (Fig. 4). In the phylogenetic analyses of monocot sequences (Figs. 3,5), monocot orthologs of *PHYA* (Fig. 2) were substituted for *Arabidopsis PHYA*. A notable finding was that from three dicot plant taxa, Ceratophyllaceae, Caryophyllaceae, and Fabaceae, two different PCR products were amplified that were determined to be most closely related to *Arabidopsis PHYA*. These are interpreted to be duplicated *PHYA* loci, and in legumes, the additional locus has been designated *PHYA*'' (Mathews et al., 1995). These additional *PHYA*-related sequences appear to have arisen independently in the dicot plant groups (Fig. 4), but were not observed in monocot taxa, except in *Panicum* (Figs. 3,5). For example, a legume phytochrome phylogeny depicts this monophyletic *PHYA*' clade as being derived from within the legume *PHYA* lineage (which is thus paraphyletic); further, it is well supported by a bootstrap value of 95%, and in a global analysis of legume *PHYA*' with all other angiosperm loci, it is most closely related to legume *PHYA* (Mathews et al., 1995). It thus appears that the evolution of the phytochrome gene family in the Fabaceae has involved the duplication of the *PHYA* locus. A similar argument can be made for the duplicated *PHYA* genes in Ceratophyllaceae and Caryophyllaceae (Fig. 4). In the *PHYA* subfamily, and in other cases described below, this pattern of diversification is attributed to the evolution of a new locus rather than to allelic diversity. With the exception of genes that are under frequency-dependent selection, such as alleles of the S-locus (Ioerger et al., 1990) and MHC-loci (Klein et al., 1993), levels of divergence among alleles at most loci are much lower (e.g., Gaut & Clegg, 1993; Thomas et al., 1993) than those observed among *PHYA* and the duplicated *PHYA* loci.

The duplicate *PHYA* locus observed in the grass *Panicum* is highly divergent from all other grass *PHYA* sequences (Fig. 5) and never occurs as a sister group to the *PHYA* homolog in *Panicum*. This degree of divergence could be indicative of gene silencing following polyploidization (see DISCUSSION). The sequence is treated here as an unknown and called *PHYU*. While a branching pattern similar to the sequences from *Panicum* is noted for the *PHYA* sequences from *Ceratophyllum* in Figure 4, in local analyses comprising less divergent taxa, the two sequences from *Ceratophyllum* occur as sister lineages (cladogram not shown).

Sequences homologous to *Arabidopsis PHYC* were amplified commonly in monocots (Figs. 3,5). In dicots, a sequence homologous to *Arabidopsis PHYC* was detected in DNA from just *Dianthus*, perhaps indicating divergence at an amplification primer sequence. The homologs of *PHYC* in monocots were identified by their close relationship with just *Arabidopsis PHYC* in a global analysis (cladogram not shown). The *PHYC* homolog in *Dianthus* was identified by its sister group relationship with *Arabidopsis PHYC* (Fig. 4).



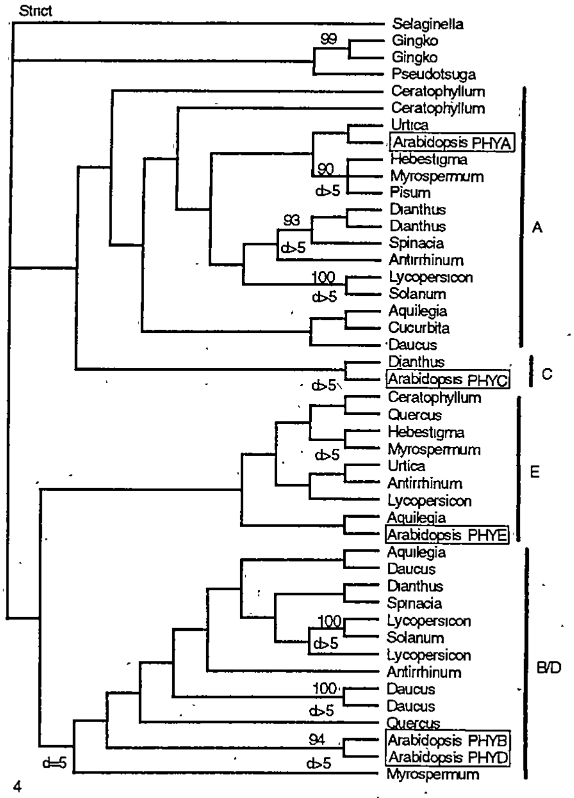


Figure 3. Single most parsimonious tree (left) from analysis of monocot DNA sequence data, which comprised 169 informative sites; length is 799, CI = 0.44, & RI = 0.52. Bootstrap (from 500 replications) and Bremer support values are included on the best supported clades. Single upper case letters to the right of the generic names are the names of the homologous *Arabidopsis PHY* loci

Figure 4. Strict consensus (right) of four most parsimonious trees from analysis of all dicot DNA sequence data, which comprised 172 informative sites; length is 1743, CI=0.23, & RI=0.49.

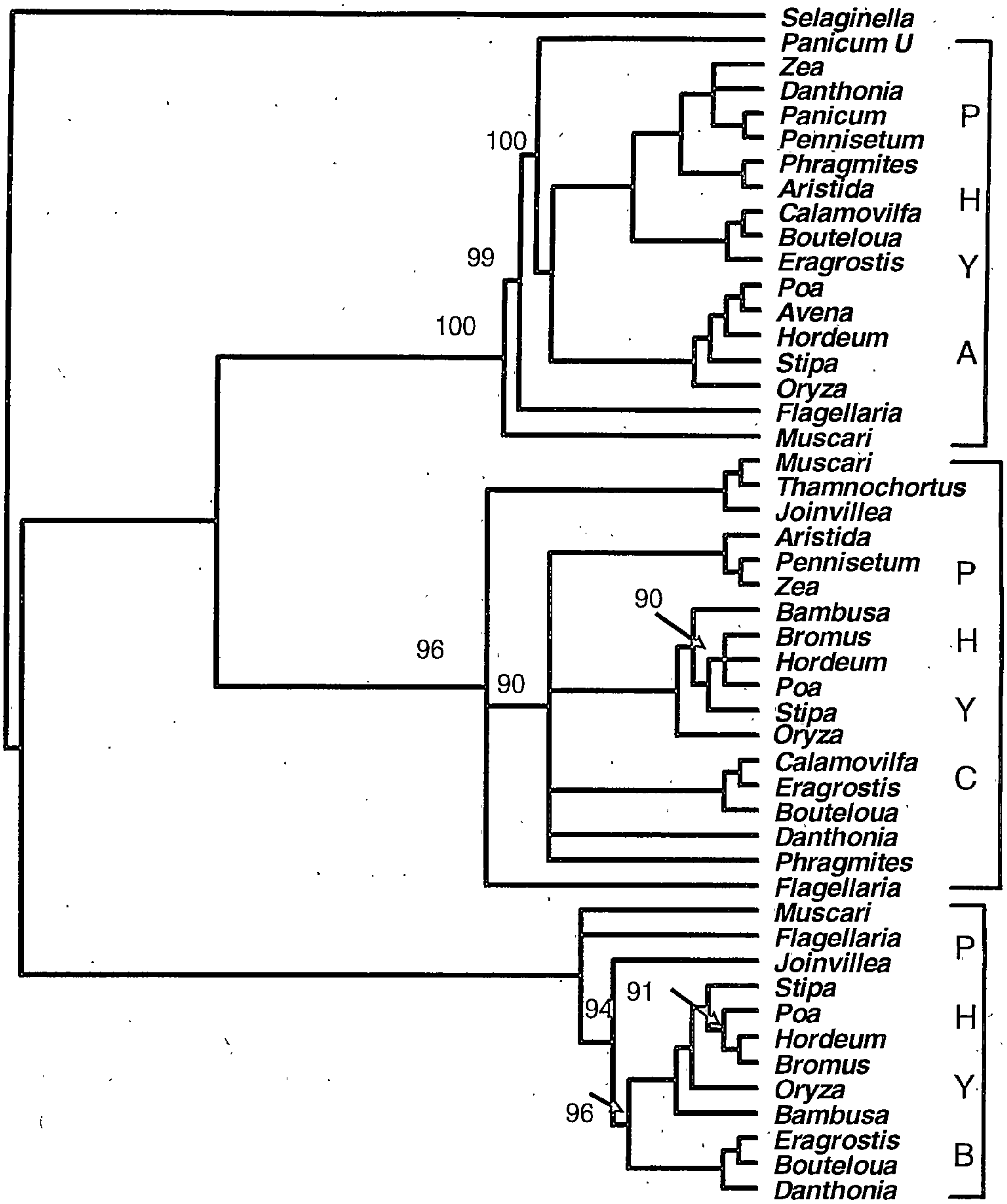


Figure 5. Strict consensus of the five most parsimonious trees from analysis of phytochrome DNA sequences from grasses, which comprised 203 informative sites. Length = 1276; CI = 33; RI = 71. Bootstrap values (from 100 replications) are included on the best supported clades. The names of the homologous *Arabidopsis PHY* loci are to the right of the generic names.

Sequences homologous to *Arabidopsis PHYE* were not amplified in any of the subclasses of monocots, nor in extensive sampling of the monocot family Poaceae, using the primer set described above (Figs. 3,5). However, such homologs were commonly amplified in dicots, including in *Ceratophyllum*, and the homology of these sequences to *PHYE* was readily established by the inclusion of *Arabidopsis PHYE* in monophyletic gene lineages (e.g., Fig. 4). In the legume phytochrome gene phylogeny, the bootstrap value for the *PHYE* clade was 100%, thus revealing how strongly this lineage is supported by the data in narrow comparisons at the taxonomic level of the family (Mathews et al., 1995).

The evolution of genes related to *Arabidopsis PHYB* has been more complex, with the apparently independent duplication and divergence of *PHYB*-related genes in some dicot lineages, but perhaps not in monocot lineages (Figs. 3-5). The notable pattern here is that the *Arabidopsis PHYB* and *PHYD* sequences are sister groups in comparisons including dicots (Figs. 2,4), and together with the sequence from *Myrospermum* are the sister group of the clade comprising the *PHYB-PHYD* subfamily. Note that two *PHYB/D*-related sequences occur in *Lycopersicon*, forming a monophyletic clade, with a *PHYB*-related sequence from

Solanum, that is separate from the clade containing *Arabidopsis PHYB* and *PHYD*, as do two of the *PHYB/D*-related sequences from *Daucus* (Fig. 4). This pattern could result from nonhomologous recombination between loci, but the hypothesis of recent divergence is consistent with the putative absence of additional *PHYB*-like sequences from monocots. Additionally, *PHYD* in *Arabidopsis* is apparently functionally distinct, as evidenced by its failure to compensate for the loss of PHYB function in *phyB* null mutants of *Arabidopsis* (Reed et al., 1993; Wester et al., 1994).

In the two trees with dicots (e.g., Figs. 2,4), *PHYE* is the sister group to the *PHYB/PHYD* clade. Since *PHYD* and *PHYE* have not been amplified from monocots, the diversification of this part of the phytochrome gene family may have taken place only during the diversification of dicots. Alternatively, *PHYE* may be ancestral to *PHYB* and *PHYD* with its absence from monocots representing a loss. In this regard, the possible absence of *PHYB* from *Ceratophyllum* is notable (see DISCUSSION), and further sampling from Nymphaeales, Piperales, Winterales, Laurales, and Magnoliales should address the question of whether the presence of just *PHYA*, *PHYB*, and *PHYC*, or the presence of just *PHYA*, *PHYE*, and *PHYC*, is the ancestral condition in angiosperms. It should be noted, however, that the inability to amplify *PHYD* and *PHYE* from monocots could mean that the oligonucleotide primers used in these studies do not recognize and amplify all *PHYD* and *PHYE* homologs. This alternative explanation should be evaluated in subsequent studies of the phytochrome gene family in monocots and in magnoliids with uniaperaturate pollen.

Analysis of combined data from Poaceae

The gene tree of phytochrome sequences from grasses illustrates a certain degree of discordance among individual gene trees from grasses (Fig. 5). For example, the *PHYA* portion of the tree resolves a clade comprising genes from panicoid, chloridoid, and arundinoid genera as the sister group of a clade comprising sequences from pooid and bambusoid genera, a degree of resolution not observed in the *PHYC* clade. The *PHYB* clade is resolved similarly to the *PHYA* clade, but there are no sequences representing panicoid genera and the sequence from *Danthonia* (an arundinoid) is the sister group of the chloridoid sequences. The combined data (Fig. 6) resolve two major lineages that are well supported, one comprising genes from bambusoid and pooid genera in which bambusoid sequences are paraphyletic to those from pooid grasses, and another that resolves a highly supported clade of sequences from chloridoid grasses as the sister group of the sequences from panicoids plus arundinoids, similarly to the *PHYA* gene phylogeny. Thus, apparent discordance among separate *PHY* gene trees simply may result from a lack of informativeness in the individual data sets that is overcome when the data are joined; for example, the *PHYA*, *PHYB*, and *PHYC* data sets have 51, 69, and 54 informative sites respectively (for the same set of taxa). A maximum likelihood analysis of the same data results in identical topology (tree not shown). The major features of this phylogeny are maintained when third codon position nucleotides are excluded from the data set, resulting in a more consistent set of characters (e.g., CI =0.62, as opposed to 0.56 when nucleotide sites are equally weighted).

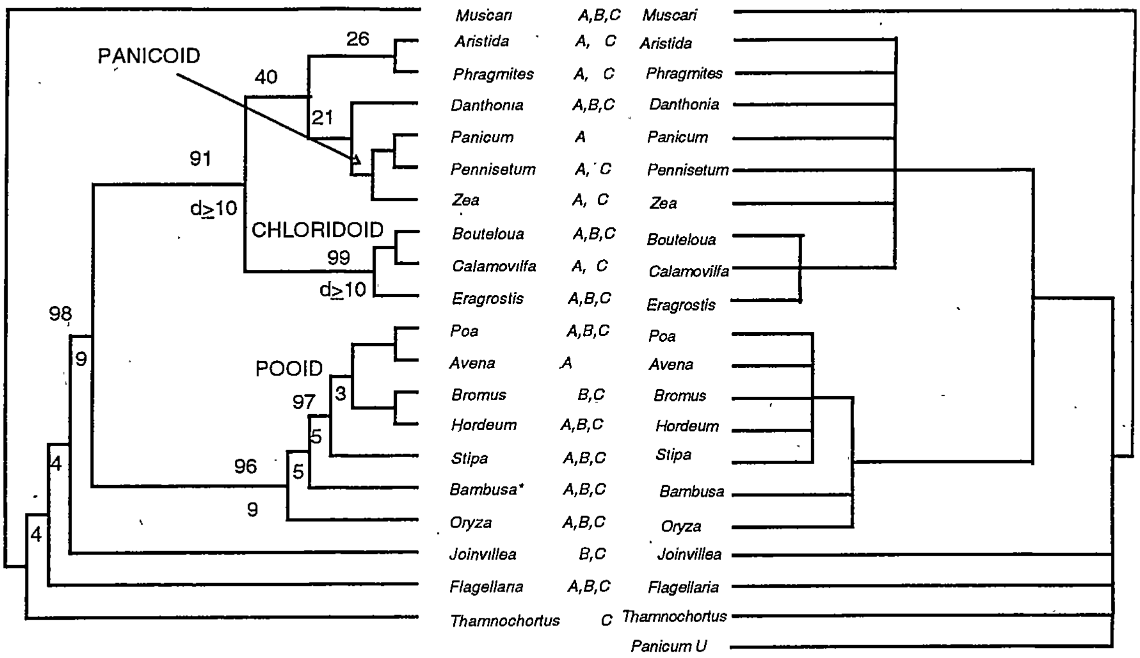


Figure 6. Single most parsimonious tree (left) from analysis of combined phytochrome DNA sequence data from grasses, which comprised 299 informative sites. Length is 907; CI = 0.56, RI = 0.53. Bootstrap values from 100 replications are above, and Bremer support values are below branches. Single uppercase letters are the *PHY* loci that were sampled from each genus; * *Bambusa PHYA* is a fragment. Strict consensus (right) of 7110 trees from analysis of the same data set with third codon position sites deleted. Length is 208 CI = 0.62; RI = 0.64.

Comparison of grass phylogenies from phytochrome and other data sets

The phytochrome sequence from *Joinvillea* is the sister group to sequences from grasses, consistent with other evidence (Campbell & Kellogg, 1987; Doyle et al., 1992). However, taken as an estimate of organismal relationships, this phylogeny differs from those inferred from other data (Fig. 7). For example, in the *rbc*L phylogeny, bambusoids are the sister group to all other grasses, while in the phylogeny inferred from cpDNA restriction site variation, they are the sister group of the panicoids, arundinoids, and chloridoids. Placement of arundinoid taxa also varies among the phylogenies, exemplified by their distribution among both panicoid- and chloridoid-containing clades in the chloroplast trees. Furthermore, morphological data place the arundinoid *Aristida* basally to all taxa but pooids. Taxa that were sampled for chloroplast data did not strictly correspond to taxa sampled for phytochrome data; thus, in some cases, taxon labels in the cpDNA and *rbc*L trees (Fig. 7) are substitutes for closely related sister taxa. Further, panicoid genera are unresolved in the morphology tree because in the analyses of morphological data they were represented by a single placeholder (Kellogg & Campbell, 1987). Strict consensus trees of the four phylogenies retain just the pooid and panicoid clades, while components of less conservative consensus trees vary (Fig. 8). For example, bambusoid genera are variously unresolved, are a monophyletic sister taxon of a panicoid + arundinoid + chloridoid clade, or are a monophyletic clade that is included in a basal polytomy; a pooid + bambusoid component is not a component of any consensus tree.

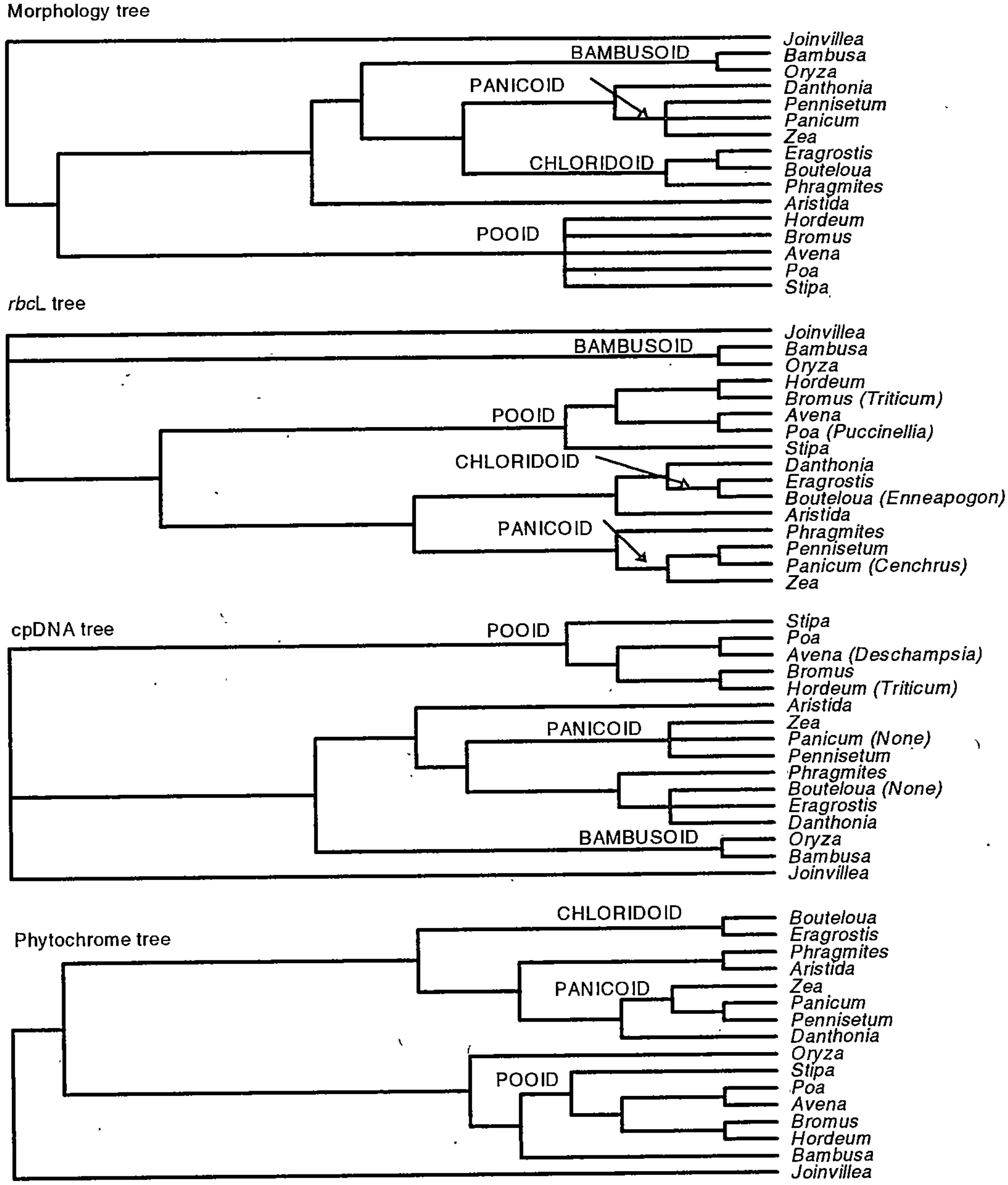


Figure 7. Comparison of the phylogeny inferred from phytochrome DNA data with organismal phylogenies inferred from other data sets for Poaceae by Kellogg & Campbell, (1987), Davis & Soreng, (1993), N. Barker (pers. comm.). Taxa that were actually sampled for the given type of data are in parentheses.

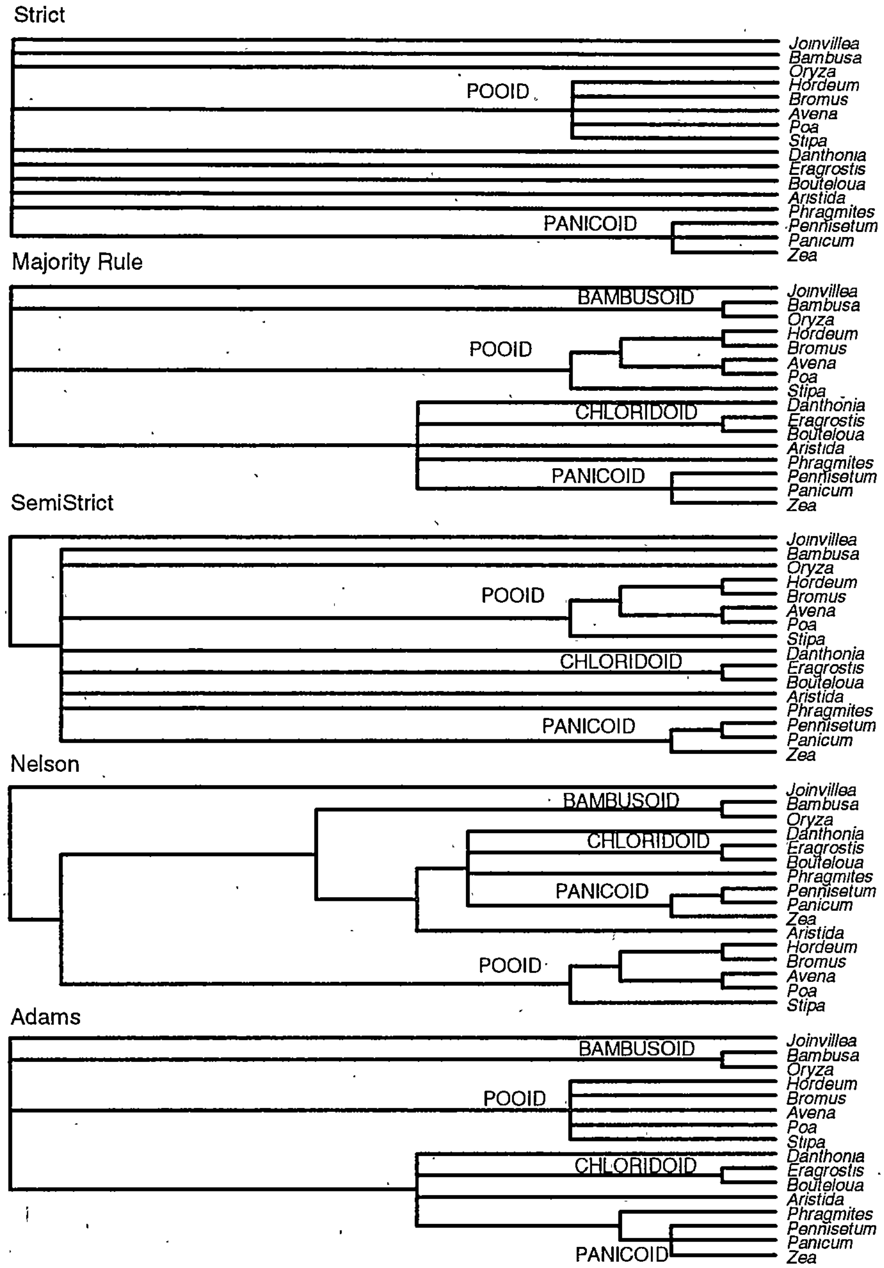


Figure 8. Different consensus trees of the four grass phylogenies in Fig. 7. All trees were calculated using algorithms in COMPONENT (Page, 1993a).

Notably, placement of the arundinoid genera is not well supported in any of the rival phylogenies (Fig. 6; Davis & Soreng, 1993;Kellogg & Campbell, 1987; N. Barker, pers. comm.). In the phytochrome phylogeny, low support may result from the lack of *PHYB* sequences from both panicoid and arundinoid genera, except *Danthonia*, because *PHYB* sequences contribute over half the informative sites to the combined phytochrome data set. Conversely, *PHYB* sequences commonly were detected in the chloridoid, bambusoid, and pooid genera, perhaps resulting in the well supported placement of bambusoid genera.

Additive binary coding matrices (App. G) were derived from the individual trees depicted in Figure 7 by recoding them as single character state trees (e.g., Brooks, 1990). This approach allows trees to be combined for analysis when the data from which they are inferred is not available (Baum, 1992) and effectively reduces each data set to a single character (e.g., Doyle, 1992) Cladistic analyses of the combined matrices resulted in two different trees (Fig. 9). If just three of the four trees were combined based on the assumption that the chloroplast data sets are not independent of one another, the resulting tree was exactly the same as the majority rule consensus tree (Fig. 8), whereas combining all four trees resulted in a tree that was most similar to the chloroplast trees (Fig. 7).

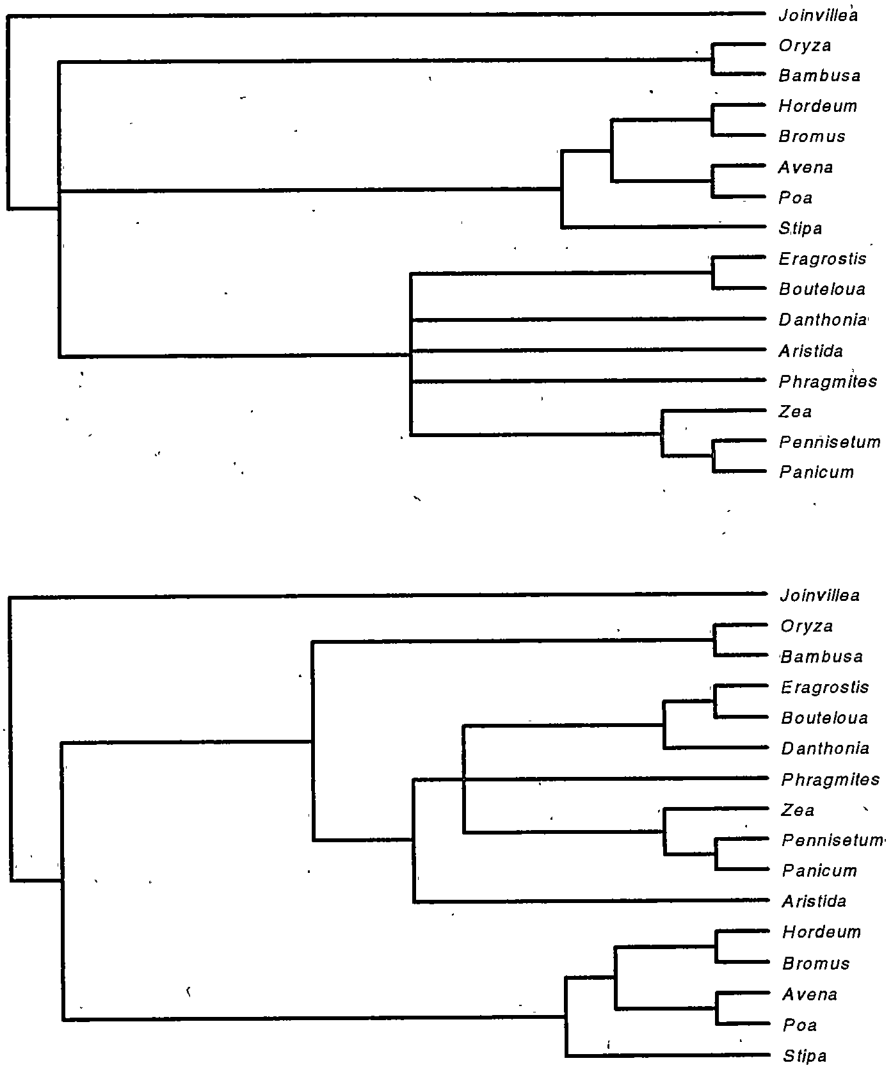


Figure 9. The two trees from analyses of combined additive binary matrices of grass phylogenies inferred from (upper) morphological, *rbc*L, and *PHY* data, and (lower) from morphological, cpDNA, *rbc*L, and *PHY* data.

<u>Analysis of nonangiosperm sequences from GenBank</u>

Relationships among angiosperm and nonangiosperm *PHY* lineages were evaluated in parsimony and distance analyses of first intron protein sequences obtained from GenBank (App. C; Figs. 10,11). The number of protein sites compared was determined by the length of the sequence from *Pseudotsuga*, which required that a number of sequences have sites coded as missing (App. C). In the strict consensus of the four minimal length trees found in parsimony analysis (Fig. 10), *PHY* sequences occur in two major clades that are designated the *PHYB/D/E* and *PHYA/C* clades according to the angiosperm loci they comprise. This analysis identifies three sequences from *Piper* as orthologs of *PHYA*, *PHYB*, and *PHYC* of *Arabidopsis*, the same subset of eudicot phytochrome loci that have been observed in monocots. Notably, nonangiosperm sequences are distributed between the two clades, rather than being basal to both, as would be expected if nonangiosperm taxa did not have duplicated *PHY* loci. The bootstrap support for splitting the gymnosperm taxa between the clades is high (91% and 94%). However, the support for placement of the other nonangiosperm taxa is minimal. Furthermore, in the neighbor-joining tree (Fig. 11), just gymnosperms are members of clusters that include angiosperm sequences, but the placement of just *Ephedra* and *Pseudotsuga* in the *PHYA/C* cluster is highly supported, with a bootstrap proportion of 96%. The presence of duplicate *PHY*s in gymnosperms is consistent with other findings (see INTRODUCTION), but duplicate *PHY*s have not been characterized in other nonangiosperms. The costliness with regard to the amount of sampling error required to explain the alternative tree topologies was compared by mapping (Page, 1993b, 1994) phytochrome trees onto the current best phylogenetic hypothesis for green plants (Fig. 12, summarized from various types of data by Donoghue [1994]). The reconciled tree depicted in Figure 13 resulted from mapping the *PHY* phylogeny (Fig. 10) onto the green plant phylogeny. Just open circles represent putative gene duplications (4), while filled circles represent putative duplications that are required to reconcile different branching orders between the *PHY* and green plant trees (7). This tree implies that a total of eleven gene duplications have occurred in the evolution of the *PHY* gene family and requires thirty-three losses (i.e., sampling error due to either gene loss or lack of detection). Two modified *PHY* phylogenies were also compared with the green plant phylogeny; in both of the modified phylogenies, the branching orders that were not well supported by phytochrome data were changed to agree with the green plant tree. However, in the first modification, just the branching orders were changed, while in the second modification, branching orders were changed *and* all nonangiosperm sequences except gymnosperm sequences were removed from the *PHYA/C* and the *PHYB/D/E* clades. Each of the reconciled trees from these subsequent comparisons (Figs. 14,15) hypothesizes just four gene duplications, but they differ in the amount of sampling error that is implied. For example, the early duplication event in Figure 14 suggests that mosses, liverworts, club mosses, and ferns and fern allies have, or have had, multiple phytochrome genes, while the later duplication in Figure 15 suggests that they do, or did, not. Each of the three reconciled trees require significantly fewer losses than do reconciliations of phytochrome trees with 1000 trees randomly generated according to the equiprobable model (i.e., trees are randomly drawn from the set of all possible tree topologies for *n* taxa, Figs. 13-15).

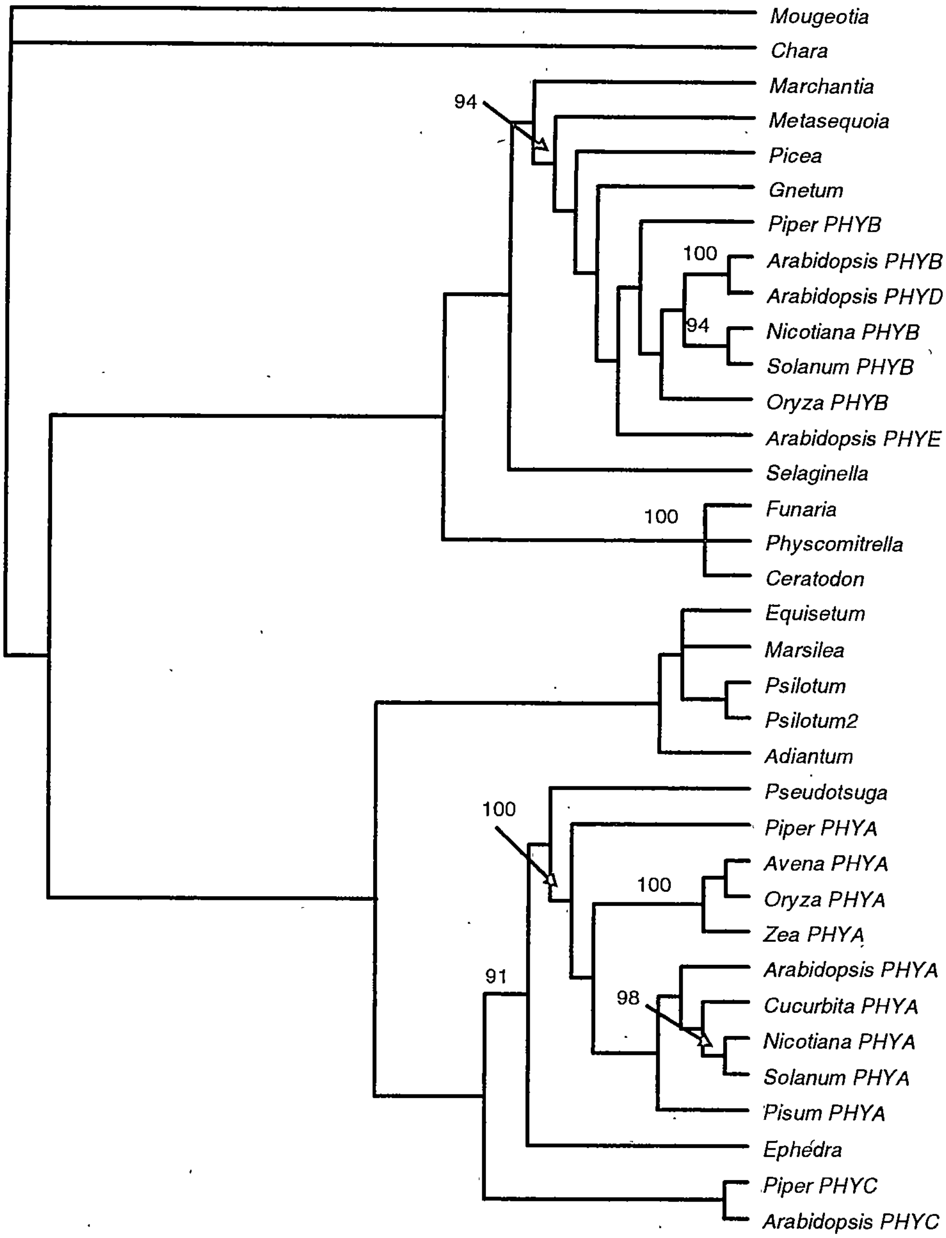


Figure 10. Strict consensus of four minimal length trees from analysis of 216 informative amino acid sites. The length is 1295. Bootstrap values from 100 replicates are included on the best supported clades.

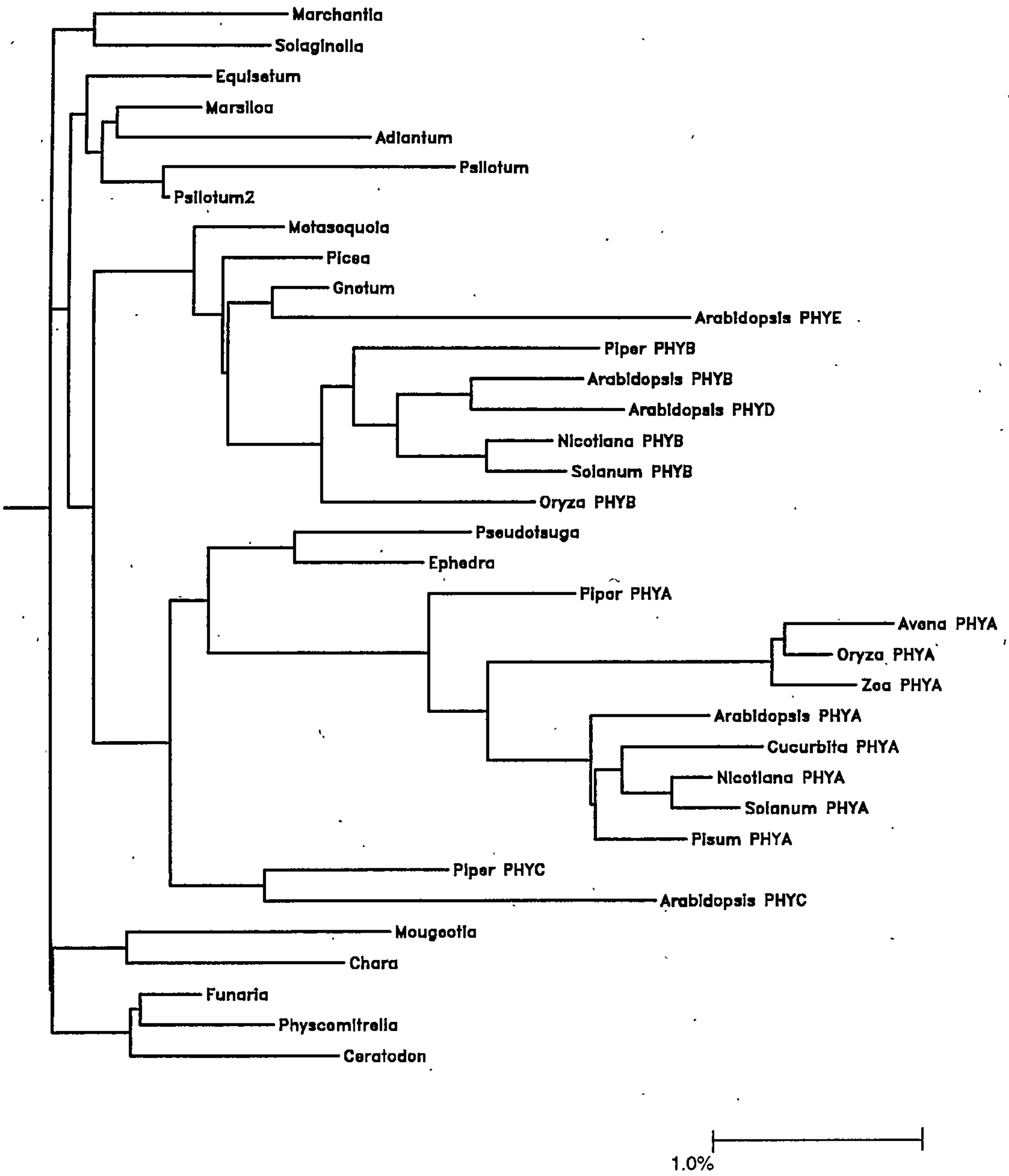


Figure 11. The neighbor-joining topology of relationships among phytochrome protein sequences of nonangiosperms and angiosperms reconstructed from Kimura 2-parameter pairwise distances.

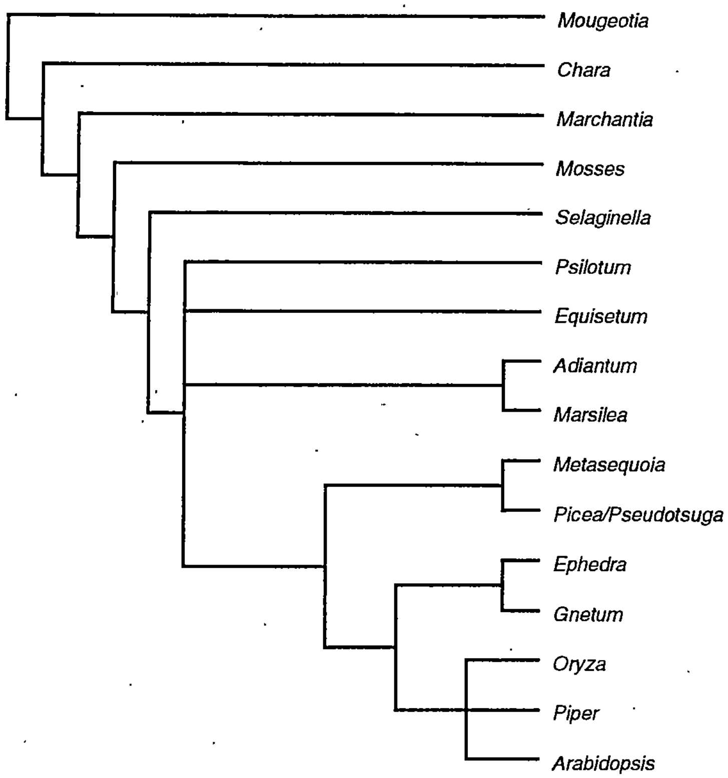


Figure 12. The branching order of nonangiosperm and angiosperm taxa (Donoghue, 1994) that were compared for phytochrome amino acid data.

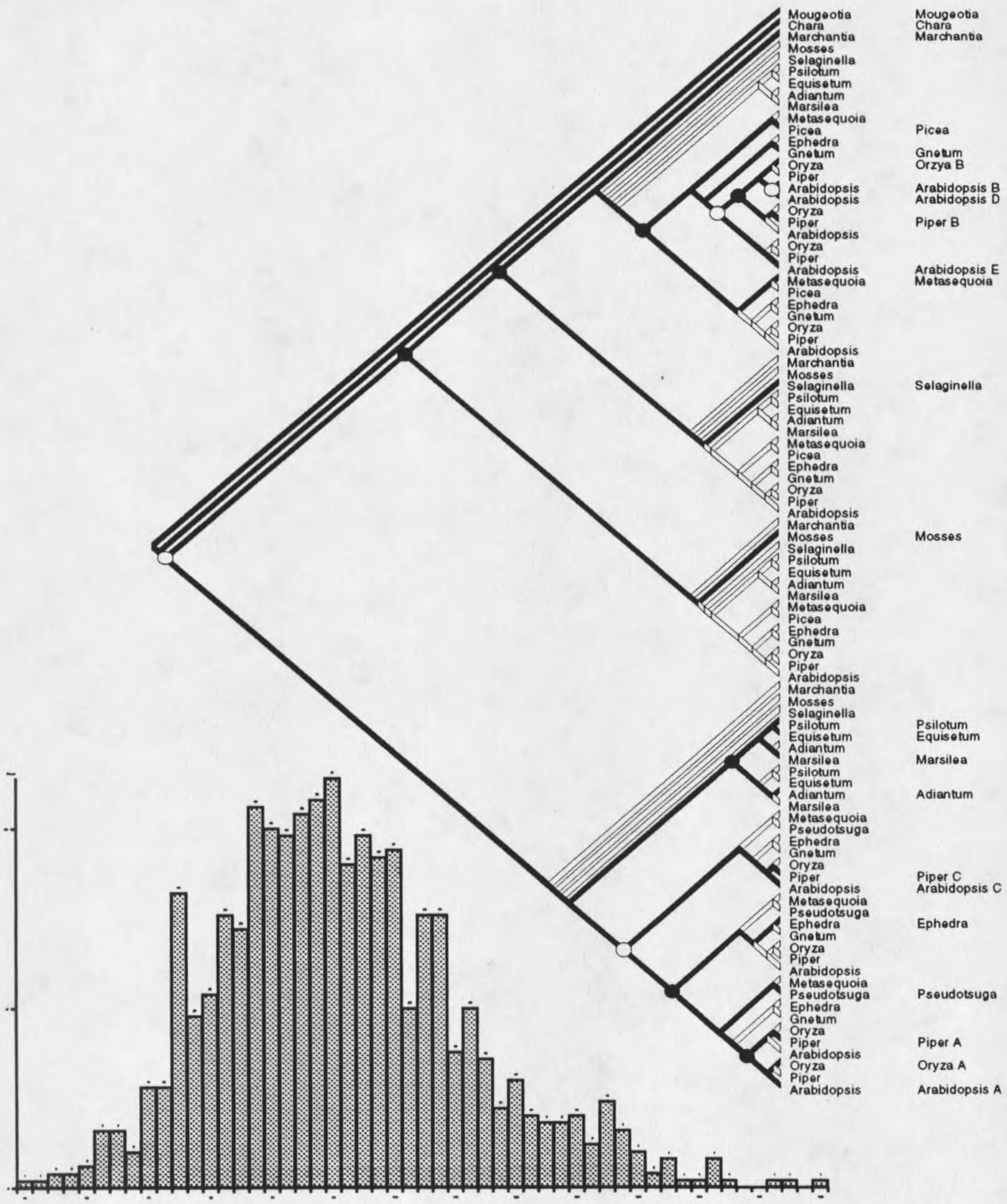


Figure 13. The reconciled tree (above) from mapping the unmodified phytochrome phylogeny (Fig. 10) onto the green plant phylogeny (Fig. 12) requires 11 gene duplications and 33 losses. Solid branches are the gene tree, open branches are inferred losses, open circles are gene duplications, and filled circles are gene duplications required to reconcile branching orders of the green and phytochrome phylogenies. The frequency distribution of losses (below) from mapping the same tree onto 1000 randomly generated trees; the mean is 97.7 and the standard deviation is 16.2.

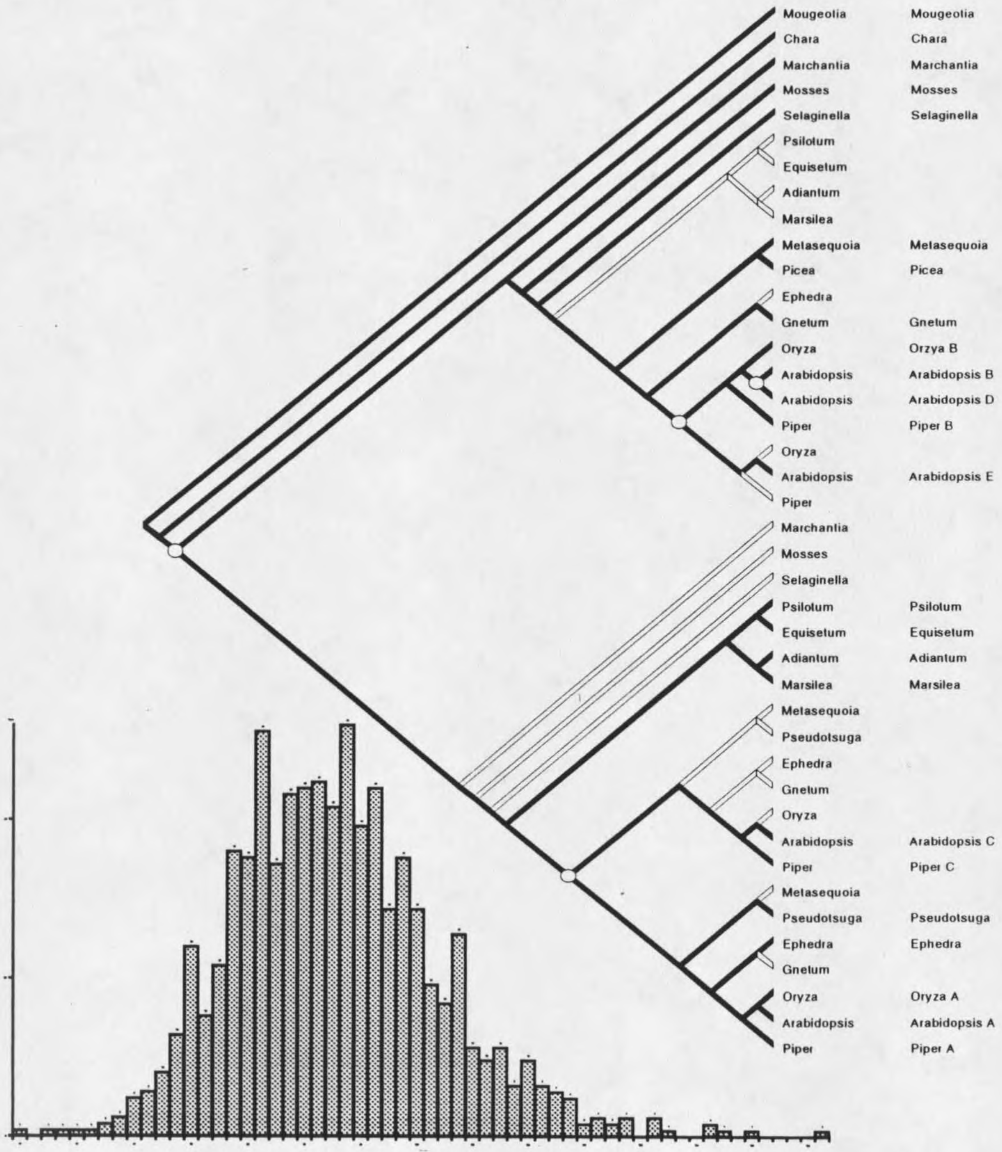


Figure 14. The reconciled tree from mapping a phytochrome tree that was modified to agree with branching order depicted in Figure 12, but which retained ferns in the *PHYA/C* clade and mosses, etc. in the *PHYB/D/E* clade (App. H) requires 4 gene duplications and 12 losses. Solid branches are the gene tree, open branches are inferred losses, open circles are gene duplications, and filled circles are gene duplications required to reconcile branching orders of the green and phytochrome phylogenies. The frequency distribution of losses (below) from mapping the same tree onto 1000 randomly generated trees; the mean is 93.4 and the standard deviation is 15.2.

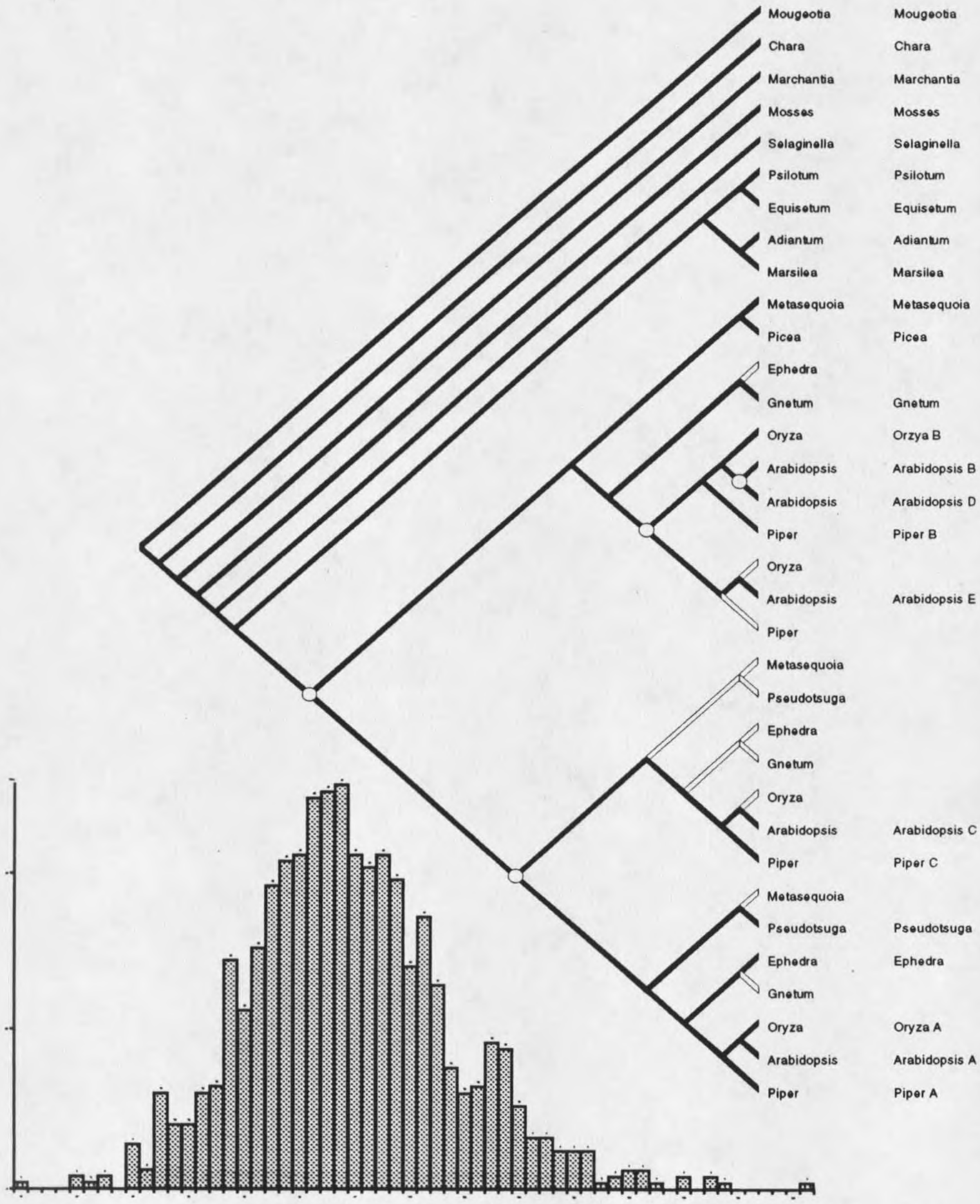


Figure 15. The reconciled tree from mapping a phytochrome tree that was modified to agree with branching order depicted in Figure 12, but which retained just gymnosperms in the *PHYA/C* & *PHYB/D/E* clades (App. H) requires 4 gene duplications and 8 losses. Solid branches are the gene tree, open branches are inferred losses, and open circles are gene duplications. The frequency distribution of losses (below) from mapping the same tree onto 1000 randomly generated trees; the mean is 95.4 and the standard deviation is 15.4.

Analyses of evolutionary rates

Absolute rates

Using the 2-parameter model of Kimura (1981) to estimate distances among all pairs of full length coding sequences (App. A), and a divergence time for *Selaginella* of 300 million years (Ma) ago (Townrow, 1968), the estimated overall rate of evolution of *PHY* lineages is 0.9 to 1.5 x 10^{-9} substitutions per site per year, or about ten times as fast as *rbc*L (Chase et al., 1993). In contrast, the rate of Jukes-Cantor corrected synonymous substitutions (K_S) among *PHY* sequences from pooid and panicoid grasses, with an estimated divergence time of 50 Ma (Doebley et al., 1990), and among tropical woody tribes of Fabaceae, with an estimated divergence time of 40 Ma (Herendeen, 1992; Wheeler & Baas, 1992) is four to five times as fast as *rbc*L (Zurawski et al., 1984; Doebley et al., 1990), or about 3.7 to 6.1 x 10^{-9} substitutions per site per year. Rates of Jukes-Cantor corrected nonsynonymous substitutions (K_A) estimated from pairwise comparisons with *Selaginella* for different portions of full length phytochrome molecules (App. I) indicate that the 594 bp including and proximal to the chromophore attachment site is the most conserved portion of the molecule (K_A = 3.2 to 4.6 x 10^{-10} subst./site/year), followed by the 2400 bp encoding the N-terminus (K_A = 4.0 to 5.4 x 10^{-10} subst./site/year), followed by 3384 bp comprising nearly the complete coding region (K_A = 4.3 to 6.2 x 10^{-10} subst./site/year). Among the full length sequences, percentage sequence divergence ranged from 12 between *PHYA*s of *Avena* and *Oryza* to 45 between *PHYE* of *Arabidopsis* and *PHYA* of *Zea*. Within dicots (Figs. 4), sequence fragments were 0% (between *PHYA*s of legumes) to 75% divergent (between *PHYB* of *Nicotiana* and *PHYE* of *Urtica* or *PHYB* of *Daucus*), while within all monocots (Fig. 3), sequence fragments were 9% (between *PHYA*s of *Avena* and *Oryza*) to 58% (between *PHYA* of *Avena* and *PHYB* of *Hordeum*) divergent. Within grasses (Fig. 5), *PHYA*s were from 5 to 18% divergent, *PHYB*s were from 6 to 20% divergent, and *PHYC*s were from 7 to 18% divergent.

Relative rates

In 72 relative rate tests (Wu & Li, 1985; App. I) used to evaluate the hypothesis that rates within and among the *PHY* loci are clocklike, twenty rate differences were significantly different ($P < 0.05$ or 0.01), given a model of rate constancy. All of these significant differences were among, rather than within, *PHY* lineages (Table 2). For example, in comparisons of sequences from dicots, evolution in *PHYB* and *PHYD* lineages is significantly slower relative to *PHYA*, *PHYC*, or *PHYE* sequences. However, in grasses, it is mostly *PHYC* lineages that are evolving at a significantly slower rate relative to *PHYA* and *PHYB* sequences, while nearly all comparisons between *PHYA*s and *PHYB*s reveal no significant rate differences; the single exception is that *Oryza PHYB* is significantly less divergent than *Oryza PHYA* when full length sequences are compared. Elsewhere in monocots, this pattern is noted in sequences from *Muscari* but not *Lemna*, where evolutionary rates between *PHYA* and *PHYC* are similar. In the dicot *Piper*, *PHYC* is significantly less divergent relative to *PHYA* and to *PHYB*, similarly to the pattern observed in monocots. The sequence from *Panicum* designated *PHYU* is significantly divergent relative to all other *PHY*s from grasses.

Table 2. Relative rate tests (Wu and Li, 1985) to detect rate assymetry. $*P < 0.05$; $**P < 0.01$. d_{13} and d_{23} are the number of nonsynonymous (or synonymous in legume comparisons) substitutions per site between species 1 and 3, and species 2 and 3, respectively; under the null hypothesis $d_{13} = d_{23}$. SE is standard error.

Species 1	Species 2	Species 3 (reference)	$d_{13} - d_{23}$ + SE
CHROMOPHORE REGION ONLY (330-594 bp) COMPARED			
Within lineage tests			
ArabidopsisA	CucurbitaA	Selaginella	0.0082 ± 0.0194
ArabidopsisA	SolanumA	Selaginella	0.0024 ± 0.0415
ArabidopsisA	OryzaA	Selaginella	-0.0033 ± 0.0416
AvenaA	ZeaA	Selaginella	-0.0081 ± 0.0161
PoaA	PennisetumA	Flagellaria	0.0520 ± 0.0276
BouteIouaB	BambusaB	Flagellaria	0.0286 ± 0.0387
BambusaC	AristidaC	Flagellaria	0.0465 ± 0.0103
OryzaC	PoaC	Flagellaria	0.0024 ± 0.0103
OryzaC	BambusaC	Flagellaria	-0.0232 ± 0.0315
ZeaC	AristidaC	Flagellaria	0.0413 ± 0.0289
ZeaC	BambusaC	Flagellaria	-0.0052 ± 0.3302
ZeaC	OryzaC	Flagellaria	0.0180 ± 0.0310
ZeaC	PhragmitesC	Flagellaria	0.0343 ± 0.0300
MyrospermumA	HebestigmaA	PisumA	0.0661 ± 0.1679
MilletiaA	SesbaniaA	MyrospermumA	0.1337 ± 0.1350
PisumA	HebestigmaA	MyrospermumA	0.2012 ± 0.1466
HebestigmaE	MilletiaE	MyrospermumE	-0.0536 ± 0.1002
Among lineage tests			
ArabidopsisA	ArabidopsisB	Selaginella	0.0451 ± 0.0395
ArabidopsisA	ArabidopsisC	Selaginella	0.0761 ± 0.0379*
ArabidopsisA	ArabidopsisD	Selaginella	0.0537 ± 0.0390
ArabidopsisA	ArabidopsisE	Selaginella	0.0709 ± 0.0387
SolanumA	SolanumB	Selaginella	0.0847 ± 0.0373*
OryzaA	OryzaB	Selaginella	0.0639 ± 0.0387
OryzaA	OryzaC	Selaginella	0.0852 ± 0.0547
OryzaB	OryzaC	Selaginella	0.0890 ± 0.0554
BambusaB	BambusaC	Selaginella	0.0919 ± 0.0549
PhragmitesA	PhragmitesC	Selaginella	0.1179 ± 0.0540*
DanthoniaA	DanthoniaB	Selaginella	0.0012 ± 0.0634
DanthoniaA	DanthoniaC	Selaginella	0.1238 ± 0.0548*
DanthoniaB	DanthoniaC	Selaginella	0.1226 ± 0.0552*
AristidaA	AristidaC	Selaginella	0.1350 ± 0.0530*

Table 2, continued.

ZeaA	ZeaC	Selaginella	0.1176 ± 0.0566*
HordeumA	HordeumB	Selaginella	0.0590 ± 0.0676
HordeumA	HordeumC	Selaginella	0.1735 ± 0.0597**
HordeumB	HordeumC	Selaginella	0.1145 ± 0.0560*
PanicumU	PennisetumA	Flagellaria	0.1350 ± 0.0366**
PanicumU	PoaA	Flagellaria	0.0830 ± 0.0409*
LemnaA	LemnaC	Selaginella	0.0084 ± 0.0525
MuscariA	MuscariB	Selaginella	-0.0057 ± 0.0634
MuscariA	MuscariC	Selaginella	0.1465 ± 0.0052**
MuscariB	MuscariC	Selaginella	0.1522 ± 0.0528**
PiperA	PiperB	Selaginella	-0.0414 ± 0.0641
PiperA	PiperC	Selaginella	0.0762 ± 0.0783
PiperB	PiperC	Selaginella	0.1176 ± 0.0585*

N-TERMINAL ENCODING SEQUENCE (2400 bp) COMPARED
Within lineage tests

ArabidopsisA	CucurbitaA	Selaginella	0.0064 ± 0.0216
ArabidopsisA	SolanumA	Selaginella	-0.0107 ± 0.0220
ArabidopsisA	OryzaA	Selaginella	0.0053 ± 0.0216
ArabidopsisB	SolanumB	Selaginella	0.0346 ± 0.0195
ArabidopsisB	OryzaB	Selaginella	-0.0008 ± 0.0204

Among lineage tests

ArabidopsisA	ArabidopsisB	Selaginella	0.0272 ± 0.0253
ArabidopsisA	ArabidopsisC	Selaginella	-0.0115 ± 0.0214
ArabidopsisA	ArabidopsisD	Selaginella	0.0182 ± 0.0213
ArabidopsisA	ArabidopsisE	Selaginella	-0.0228 ± 0.0225
ArabidopsisB	ArabidopsisC	Selaginella	-0.0387 ± 0.0221
ArabidopsisB	ArabidopsisE	Selaginella	-0.0500 ± 0.0218*
SolanumA	SolanumB	Selaginella	0.0725 ± 0.0205**
OryzaA	OryzaB	Selaginella	0.0211 ± 0.0210

FULL LENGTH CODING SEQUENCE (3384 bp) COMPARED
Within lineage tests

ArabidopsisA	CucurbitaA	PisumA	-0.0112 ± 0.0111
ArabidopsisA	SolanumA	Selaginella	-0.0145 ± 0.0199
ArabidopsisA	OryzaA	Selaginella	-0.0226 ± 0.0202
ArabidopsisB	SolanumB	Selaginella	0.0326 ± 0.0234
ArabidopsisB	OryzaB	Selaginella	-0.0085 ± 0.0179

Among lineage tests

ArabidopsisA	ArabidopsisB	Selaginella	0.0476 ± 0.0187*

Table 2, continued.

ArabidopsisA	ArabidopsisC	Selaginella	-0.0346 ± 0.0206
ArabidopsisA	ArabidopsisD	Selaginella	0.0339 ± 0.0190
ArabidopsisA	ArabidopsisE	Selaginella	-0.0273 ± 0.0204
ArabidopsisE	ArabidopsisB	Selaginella	0.0749 ± 0.0194**
ArabidopsisE	ArabidopsisD	Selaginella	0.0612 ± 0.0197**
ArabidopsisC	ArabidopsisB	Selaginella	0.0822 ± 0.0194**
ArabidopsisC	ArabidopsisD	Selaginella	0.0685 ± 0.0197**
SolanumA	SolanumB	Selaginella	0.0947 ± 0.0183**
OryzaA	OryzaB	Selaginella	0.0617 ± 0.0194**

CHAPTER 4

DISCUSSION

Evolution of Phytochrome Genes

Origin of phytochromes

Phytochromes are known from just green plants. Homologies of all or portions of the molecules with currently identified nonphytochrome proteins remain unknown despite the observation that phytochrome C-termini exhibit limited similarity to the transmitter component of bacterial two-component signalling systems (e.g., Schneider-Poetsch et al, 1991). The bacterial systems comprise an N-terminal sensory domain and a (usually) C-terminal transmitter domain (Parkinson & Kofoid, 1992). In combination with the observed protein sequence similarity (just 22 to 27%), the role of the phytochromes to sense (at the N-terminus) and respond to environmental stimuli is used to infer homology of phytochromes with the bacterial molecules (e.g., Schneider-Poetsch et al., 1991). Two-component signalling systems have now been described from eukaryotic cells (Hughes, 1994), and include the ethylene-response gene (*ETR1*, Chang et al., 1993) of *Arabidopsis*. However, the canonical consensus sequence that typifies all known bacterial transmitter components (e.g., Parkinson & Kofoid, 1992), and the plant and yeast systems, is not highly conserved in the phytochrome sequences. Furthermore, the histidine residue that is autophosphorylated in bacterial and eukaryotic systems is found in only four of the phytochrome sequences. Therefore, while the similarity among phytochrome C-termini and bacterial transmitter components (and to *ETR1* of 23%) may be suggestive of phylogenetic origin of phytochrome C-termini, it is not a sound basis for inferring homologous function.

<u>Evolution of phytochromes</u>

The evolutionary pattern inferred from phytochrome gene studies is that *PHY* gene diversity appears to be limited in nonangiosperms, where often a single gene is found, while diversity is much greater in angiosperms, where orthologs of the *PHYA*, *PHYB/D*, *PHYC*, and *PHYE* genes discovered in *Arabidopsis* are present (Figs. 2, 3-5). The most conservative hypothesis regarding early duplication events from these data suggests that two gene duplications occurred in a gymnosperm group, leading to distinct *PHYA*, *PHYB/D/E*, and *PHYC* lineages prior to the origin of flowering plants. The phylogenetic evidence that conifers (*Metasequoia*, *Picea*, and *Pseudotsuga*) and Gnetales (*Gnetum* and *Ephedra*) have two loci is high (Fig. 10), and there is empirical evidence supporting this hypothesis; two distinct sequences are reported from *Pinus* (see INTRODUCTION). Furthermore, the sequence detected from *Pseudotsuga* in this investigation apparently is a homolog of *PHYB* of *Arabidopsis* and is distinct from the sequence in GenBank that apparently is a homolog of *PHYA* (compare Figs. 3 & 10). Tree mapping results indicate the potential presence of a third gene in these taxa (Fig. 15), but suggest that diversification in the *PHYB/D/E* lineage is unique to angiosperms in agreement with empirical data. A more speculative hypothesis places the initial phytochrome gene duplication earlier, predating the origin of mosses, liverworts, and club mosses (Fig. 14). However, aspects of the phytochrome phylogenies that suggest this pattern lack support (Fig. 10). Empirical support also is lacking. Thus, the putative absence of duplicate loci in mosses, liverworts, club mosses, and ferns and fern allies would require explanations such as gene loss or gene conversion events.

In angiosperms, the model of a five member phytochrome gene family developed for *Arabidopsis* is probably not completely appropriate for all groups. For example, though the PCR primers developed in this study annealed to and amplified dicot orthologs of *PHYA*, *PHYB/D*, *PHYC*, and *PHYE* of *Arabidopsis* they annealed to and amplified only three genes in monocots, orthologs of *PHYA*, *PHYB/D*, and *PHYC*, including in an extensive sampling of DNAs from the monocot family Poaceae. The same primers applied to DNAs from the Fabaceae most commonly amplified *PHYA*, *PHYA'*, and *PHYE*; rarely did they amplify *PHYB/D* homologs, and they have yet to amplify *PHYC* (Mathews et al., 1995). It is very possible that sequence divergence at the primer sites precludes the amplification of all loci present in some genomes, or that bias toward certain gene family members has occurred during amplification cycles; i.e., PCR selection or drift (sensu Wagner et al., 1994). For example, multiple strategies applied to DNA of tomato detected only four of the homologs of *Arabidopsis*, while other scientists, using different primers, apparently detected the complete set (Cordonnier-Pratt, 1994). Nonetheless, preliminary results indicate that the same loci are obtained from genera in Fabaceae when primers differing in GC content are used (Lavin, unpublished). Likewise, certain variations of initial amplification conditions have not altered the set of loci detected in other angiosperms (Mathews, unpublished). Thus, it is likely that all five genes characterized from *Arabidopsis* did not precede the early diversification of angiosperms. Indeed, data presented here showing independent evolution of multiple *PHYB/D*-related sequences in *Arabidopsis*, *Lycopersicon*, and *Daucus* indicate that the divergence of the *PHYB* and *PHYD* loci in *Arabidopsis* occurred sometime well after the diversification of dilleniid families (Fig. 4). Recent diversification of the phytochrome gene family in angiosperms is also suggested by the occurrence of *PHYA*-related sequences that have independently evolved in Ceratophyllaceae, Caryophyllaceae, and Fabaceae (Fig. 4; Mathews et al., 1995). Furthermore, the same subset of homologs of *Arabidopsis* that were detected in monocots was detected in *Piper*, a basal dicot group that potentially is more closely related to monocots than it is to higher dicots; these sequences were observed by other workers using a different set of amplification primers (Kolukisaoglu et al., 1993, unpublished).

All of the angiosperm phytochrome sequences detected in this study and that are available from the public databases fall into the three distinct *PHY* lineages characterized here, the *PHYA*, *PHYB/D/E* and *PHYC* subfamilies. Reports that additional lineages may exist in tomato, sorghum, and *Arabidopsis* (Pratt, 1995) are not substantiated by sequence data but are inferred from unpublished blot hybridization results; furthermore, the argument that a partial sequence for 125 kDA oat phytochrome is putatively "unlike the comparable sequence encoded by known *PHYA*, *PHYB* and *PHYC*" and thus indicates that "at least a fourth *PHY* must exist" (Pratt, 1995:15) merely is based on percent similarity of sequences. By this criterion, most of the *PHYC* homologs from grasses, such as *Zea* (74% similar to *Arabidopsis PHYC*) and *Hordeum* (72% similar to *Arabidopsis PHYC*) would be considered members of unknown *PHY* lineages, a conclusion that is unequivocally disputed by the results from phylogenetic analysis of phytochromes that are presented here.

Tempo of sequence evolution

Results from relative rate tests indicate that evolutionary rates among *PHY* lineages are not clocklike, as evidenced by the slower evolution of *PHYB* and *PHYD* in dicot flowering plants, and of *PHYC* in *Piper*, grasses, and *Muscari*. Furthermore, the rapid evolution of *Panicum PHYU* relative to all other phytochrome sequences from grasses is consistent with the notion that it might be a nonfunctional gene. A pseudogene has been observed in *Pisum* (Sato, 1990), and a sequence in *Zea* has been interpreted as a pseudogene (Christensen & Quail, 1989). Moreover, *Panicum* DNA potentially came from a polyploid individual; molecular data are accumulating that are consistent with putatively extensive gene silencing in polyploid genomes of angiosperms and ferns, including in the grasses *Zea* and *Sorghum* (reviewed in Soltis & Soltis, 1993).

Implications for organismal phylogenetic analyses

These phytochrome gene studies suggest that *PHY* loci occur in distinct monophyletic lineages and thus, that organismal phylogenies inferred from phytochrome data are free from the misleading effects of mistaken orthology of genes. Additionally, equal evolutionary rates *within PHY* lineages indicate that phylogenies inferred from phytochrome data are likely free of spurious long branch attractions (resulting from combinations of very long and very short branches in trees, e.g., Hendy & Penny, 1989) because it is the branching order among orthologous sequences that are the source of inference about organismal relationships. Furthermore, the tempo of sequence evolution predicts that the data are potentially informative at the subfamilial taxonomic level (e.g., about 5 to 10 times as fast as the chloroplast locus *rbc*L); additionally, at the subfamilial level, sequence divergence is from 1 to 20%, which encompasses the optimal range for inferring phylogeny from DNA data (Ritland & Eckenwalder, 1992).

Detailed investigations of phytochrome gene and nucleotide diversity in two plant families provide empirical data supporting these predictions. In dicotyledonous Fabaceae (legumes), phytochrome sequence data are providing a high degree of phylogenetic resolution to a taxonomically very difficult tribe of tropical genera; useful characters include nucleotide substitutions and phylogenetically informative insertions and deletions (Mathews et al., 1995; M. Lavin, pers. comm.).

In contrast to phytochrome studies in Fabaceae that comprised genera from closely related tribes, phytochrome studies in monocotyledonous Poaceae (grasses) included genera from each of the major subfamilies, Bambusoideae, Pooideae, Chloridoideae, Panicoideae, and Arundinoideae (e.g., Watson et al., 1985; Watson & Dallwitz, 1988; Clayton & Renvoise, 1986) that are inferred from anatomical, chemical, cytological, morphological, and physiological data. The subfamilies, except for the Arundinoideae, comprise cores of genera whose affinities are reasonably well defined, while placement of numerous anomalous genera remains controversial (e.g., *Aristida*). Explicitly phylogenetic studies of the phenetic subfamilies have provided data supporting the monophyly of some taxa, but their interrelationships are unresolved. For example, morphological data suggest that Bambusoideae and Pooideae (when broadly defined), and Chloridoideae, and Panicoideae (Kellogg & Campbell, 1987; Kellogg & Watson, 1993) are monophyletic, while Arundinoideae are interpreted to be a polyphyletic "assemblage of basal groups" (Kellogg & Campbell, 1987:28) from which the Bambusoideae, Chloridoideae, and Panicoideae are derived. Likewise, chloroplast DNA restriction site data (Davis & Soreng, 1993, 1994) provide evidence that the Pooideae and Panicoideae are monophyletic and the Arundinoideae paraphyletic. However, Davis and Soreng (1994) concluded that Bambusoideae is a paraphyletic taxon from which the Pooideae and a monophyletic clade that includes panicoid, arundinoid, centothecoid , and chloridoid genera (their PACC clade) are derived. An *rbc*L phylogeny differs from both the morphology and cpDNA trees in that bambusoid genera comprise a monophyletic clade that is the sister taxon to all other grasses; it is similar to the cpDNA phylogeny in its resolution of a monophyletic pooid clade and a monophyletic PACC clade in which arundinoid genera are paraphyletic (N. Barker, pers. comm.). A nuclear ribosomal DNA phylogeny (Hamby & Zimmer, 1992) is similar to the morphological phylogeny in that rice (Bambusoideae) is the sister taxon of a monophyletic panicoid clade, and together these taxa are the sister group of a monophyletic pooid clade, with the single arundinoid basal to all clades (chloridoid genera were not sampled). Discordance among the phylogenies may be more apparent than real and could be due, for example, to differences in taxonomic sampling. The possibility of a paraphyletic Bambusoideae is discussed by Kellogg and Watson (1993) and depends in part on whether bambusoid or arundinoid taxa are considered to be most basal among grasses. Alternatively, the lack of resolution within and among grass phylogenies has been attributed to rapid speciation events that were not accompanied by significant molecular and morphological divergence (Kellogg & Watson, 1993; N. Barker, pers. comm.), or to extensive parallel evolution (Kellogg & Watson, 1993). These may well be valid explanations of morphological evolutionary events; however, lack of strong phylogenetic signal in molecular data sets (evidenced by low statistical support in phylogenies) also could result from the choice of the molecular tool (e.g., the more slowly evolving chloroplast or ribosomal DNA molecules).

In contrast to other molecular data sets, the combined phytochrome nucleotide data from Poaceae resolve several highly supported lineages. For example, phytochrome data support the monophyly of Chloridoideae, a broadly defined Pooideae that includes *Stipa*, and a clade comprising Panicoideae, Chloridoideae, and several arundinoid taxa, as do the chloroplast data. However, phytochrome data unambiguously place bambusoid genera with Pooideae (bootstrap support of 96%), including when third codon position sites are dropped from the analysis, while the chloroplast phylogenies differ in their placement of bambusoid taxa. Furthermore, bootstrap support for the *Stipa* + Pooideae and "PACC" clades is higher in the phytochrome data set (67 vs. 97% and 84 vs. 91%, respectively, when compared with the *rbc*L phylogeny). In the phytochrome tree, placement of the arundinoid taxa is weakly supported, as it is in trees from all of the other data sets; yet this merely may result from the lack of the more phylogenetically informative *PHYB* sequences from arundinoid and panicoid taxa in the data set. Thus, these analyses predict that phytochrome nucleotide variation is potentially very informative regarding phylogenetic relationships among grasses, particularly when data from multiple loci are combined. Furthermore, similarly to the phytochrome data set from Fabaceae, a homologous deletion shared by chloridoid genera potentially provides a tool to more definitively circumscribe Chloridoideae (App. E).

Comparison of the phylogeny of grasses inferred from combined phytochrome data with taxonomic hypotheses inferred from other grass data sets in consensus analysis reveals a significant degree of discordance among the phylogenies. Strict consensus trees were highly unresolved and the less conservative consensus methods result in trees that conflict in some way with individual components of rival trees. No consensus tree retains the robust bambusoid + pooid component observed in the phytochrome phylogeny. However, the observed incongruence among independent phylogenetic hypotheses may be informative. For example, it might be evidence of differential evolutionary history of chloroplast and nuclear genomes; cytoplasmic gene flow apparently occurs much more readily than nuclear gene flow in plants, and at a variety of taxonomic levels (Rieseberg & Soltis, 1991).

Further data from nuclear genes of grasses would potentially clarify the history of both genomes. Additionally, discordance among trees inferred from different molecular data sets at the very least suggests further sampling strategies. Phylogenetic results are strongly impacted by species sampling (e.g., LeCointre et al., 1993) and additional sampling from Bambusoideae and Arundinoideae potentially will modify phylogenetic hypotheses of grasses. Alternatively, disagreement between phylogenies inferred from phytochrome and morphological data may provide insight as to which morphological characters are least subject to parallel evolution (e.g., perhaps embryo characters will be more reliable than leaf anatomical characters).

Combining gene trees by recoding them as character state trees potentially is informative if trees from numerous, *independent* data sets are available. The emphasis is placed on independence because the method appears to be very sensitive to components that are in the majority.

CHAPTER 5

CONCLUSIONS

Despite cautions regarding the use of molecular data from multigene families for phylogenetic inference, specifically because orthology potentially is mistaken due to events such as gene conversions, phytochrome nucleotide data are shown here to be potentially very useful for evolutionary investigations. This is largely because results from cladistic analysis of phytochrome DNA sequences strongly suggest that loci are evolving independently, as evidenced by the distribution of all *PHY* sequences in monophyletic gene lineages; such a pattern is unlikely if nonhomologous recombination is frequent. Thus, nucleotide variation allowed the gene family model inferred from *Arabidopsis* to be tested in other species because orthologous sequences could be readily identified. Additionally, two empirical investigations of gene and nucleotide diversity in Fabaceae and Poaceae, dicot and monocot families respectively, resulted in strongly supported phylogenetic hypotheses that may resolve current taxonomic questions. Moreover, a distinct advantage of the phytochrome gene family for phylogenetic inference, and potentially of other gene families, is that the distribution of loci is likely to be very informative. For instance, the presence of a legume-specific locus related to *PHYA* should prove to be phylogenetically informative once its taxonomic distribution is better known. Structural mutations have proven very useful for resolution of certain phylogenetic relationships, exemplified by the presence of tRNA introns that identify certain algal groups as close relatives of land plants (Manhart & Palmer, 1990), by chloroplast DNA inversions that identify Joinvilleaceae as the sister group of Poaceae (Doyle et al., 1992), and by the loss of the chloroplast DNA inverted repeat in tribes of Fabaceae (Lavin et al., 1990).

The apparent differential distribution of phytochrome loci among major lineages of angiosperms should similarly be useful for addressing evolutionary questions. Specifically, the origin and interrelationships of major angiosperm groups remain unresolved (e.g., Fig. 16). Alternate hypotheses (e.g., Doyle et al., 1994) place the origin of angiosperms either among woody Magnoliales or among herbaceous "paleoherbs" (e.g., Piperales, Nymphaeales, monocots). Likewise, rival hypotheses differently explain origins of monocots (e.g., Dahlgren et al., 1985). Phytochrome data may provide resolution to these issues specifically because the groups in question appear to be characterized by different *PHY* loci; for, example, just homologs of *PHYA*, *PHYB*, and *PHYC* are known from monocots and *Piper*, while homologs of *PHYA*, *PHYB*, *PHYC* and *PHYE* of *Arabidopsis* occur in eudicots. Moreover, the paleoherb *Ceratophyllum* (Nymphaeales) has a *PHYE* that apparently is orthologous with both eudicot *PHYE*s and sequences from gymnosperms.

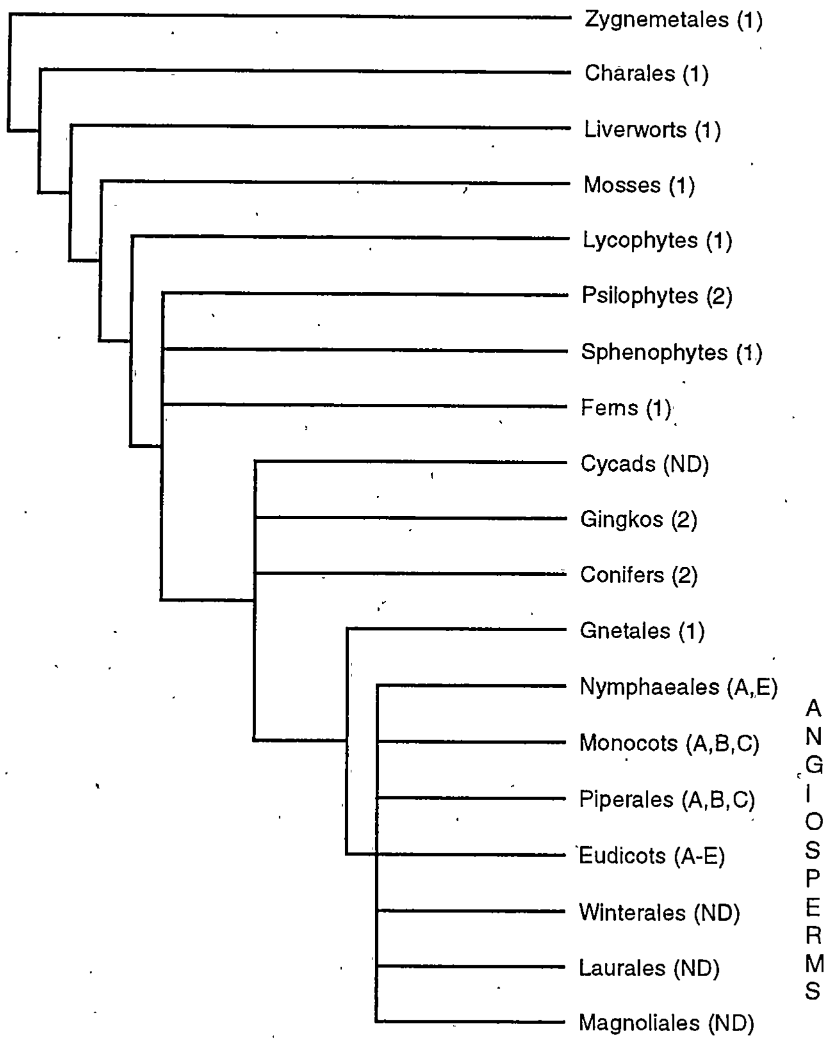


Figure 16. The hypotheses of green plant phylogeny summarized by Donoghue (1994) showing the lack of resolution in the angiosperm clade. Numbers of phytochrome genes that have been detected in nonangiosperms, and homologs of the *Arabidopsis* genes that have been detected in angiosperms, are in parentheses. ND is not determined.

In summary, the method of sampling DNA for molecular data via the polymerase chain reaction developed in this work has proven useful for characterization of the phytochrome gene family in angiosperms. While limitations exist with respect to inferring molecular evolutionary events because just gene fragments have been sampled, the data are adequate for phylogenetic studies and for estimating a model of phytochrome evolution that can be tested with further data.

LITERATURE CITED

LITERATURE CITED

Adam, E., M. Deak, S. Kay, N. H. Chua, & F. Nagy. 1993. Sequence of tobacco (*Nicotiana tabacum*) gene coding for type A phytochrome. Plant Physiol. 101: 1407-1408.

Adams, E. N. 1972. Consensus techniques and the comparison of taxonomic trees. Syst. Zool. 21: 390-397.

Albert, V. A., M. W. Chase, & B. D. Mishler. 1993. Character-state weighting for cladistic analysis of protein-coding DNA sequences. Ann. Missouri Bot. Gard. 80: 752-766.

Barrett, M., M. J. Donoghue, & E. Sober. 1991. Against consensus. Syst. Zool. 40: 486-493.

Batschauer, A., P. M. Gilmartin, F. Nagy, & E. Schäfer. 1994. The molecular biology of photoregulated genes. Pp. 559-599 *in* R. E. Kendrick & G. H. M. Kronenberg (editors), Photomorphogenesis in plants. Kluwer Academic Publishers, Dordrecht.

Baum, B. 1992. Combining trees as a way of combining data sets for phylogenetic inference, and the desirability of combining gene trees. Taxon 41: 3-10.

Bremer, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. Evolution 42: 795-803.

Brooks, D. R. 1981. Hennig's parasitological method: A proposed solution. Syst. Zool. 30: 229-249.

Brooks, D. R. 1990. Parsimony analysis in historical biogeography and coevolution: Methodological and theoretical update. Syst. Zool. 39: 14-30.

Bull, J. J., J. P. Huelsenbeck, C. W. Cunningham, D. L. Swofford, & P. J. Waddell. 1993. Partitioning and combining data in phylogenetic analysis. Syst. Biol. 42: 384-397.

Campbell, C. S., & E. A. Kellogg. 1987. Sister group relationships of the Poaceae. Pp. 217-224 *in* T. R. Soderstrom, K. W. Hilu, C. S. Campbell, & M. E. Barkworth (editors), Grass systematics and evolution. Smithsonian Institution, Washington, D. C.

Carpenter, J. M. 1988. Choosing among multiple equally parsimonious cladograms. Cladistics 4: 291-296.

Chang, Caren, S. F. Kwok, A. B. Bleecker, & E. M. Meyerowitz. 1993. *Arabidopsis* ethylene-response gene *ETR1*: Similarity of product to two-component regulators. Science 262: 539-544.

Chase, M. W., et al. 1993. Phylogenetics of seed plants: An analysis of nucleotide sequences from the plastid gene *rbc*L. Ann. Missouri Bot. Gard. 80: 528-580.

Cherry, J. R., D. Hondred, J. M. Walker, J. M. Keller, H. P. Hershey, & R. D. Vierstra. 1993. Carboxy-terminal deletion of oat phytochrome A reveals the presence of separate domains required for structure and biological activity. Plant Cell 5: 565-575.

Chippindale, P. T., & J. J. Wiens. 1994. Weighting, partitioning, and combining characters in phylogenetic analysis. Syst. Biol. 43: 278-287.

Christensen, A. H., & P. H. Quail. 1989. Structure and expression of a maize phytochrome-encoding gene. Gene 85: 381-390.

Clack, T., S. Mathews, & R. A. Sharrock. 1994. The phytochrome apoprotein family in *Arabidopsis* is encoded by five genes: the sequences and expression of *PHYD* and *PHYE*. Plant Molec. Biol. 25: 413-427.

Clayton, W. D. & S. A. Renvoise. 1986. Genera Graminum. Her Majesty's Stationery Office, London.

Cordonnier-Pratt, M.-M., L. H. Pratt, B. Hauser, G. Kochert, and M. Caboche. 1994. Comparative analysis of the phytochrome gene family in tomato (*Lycopersicon esculentum* Mill.) and sorghum ([L.]Moench). Pl. Physiol. 105 (Supplement): Abstract 72.

Cowl, J. S., N. Hartley, D.-X. Xie, G. C. Whitelam, G. P. Murphy, & N. P. Harberd. 1994. The *PHYC* gene of *Arabidopsis*. Pl. Physiol. 106: 813-814.

Cracraft, J., & K. Helm-Bychowski. 1991. Parsimony and methodological strategies. Pp. 184-220 *in* M. M. Miyamoto & J. Cracraft (editors), Phylogenetic analysis of DNA sequences. Oxford University Press, New York.

Cronquist, A. 1981. An Integrated System of Classification of Flowering Plants. Columbia Univ. Press, New York.

Dahlgren, R. M. T., H. T. Clifford, & P. F. Yeo. 1985. The families of monocotyledons. Springer-Verlag, Berlin.

Davis, J. I, & R. J. Soreng. 1993. Phylogenetic structure in the grass family (Poaceae) as inferred from chloroplast DNA restriction site variation. Amer. J. Bot. 80: 1444-1454.

Dehesh, K., J. Tepperman, A. H. Christensen, & P. H. Quail. 1991. *phyB* is evolutionarily conserved and constitutively expressed in rice seedling shoots. Mol. Gen. Genet. 225: 305-313.

de Queiroz, A. 1993. For consensus (sometimes). Syst. Biol. 42: 368-372.

Doebley, J., M. Durbin, E. M. Golenberg, M. T. Clegg, & D. P. Ma. 1990. Evolutionary analysis of the large subunit of carboxylase (*rbcL*) nucleotide sequence among the grasses (Gramineae). Evolution 44: 1097-1108.

Donoghue, M. J. 1994. Progress and prospects in reconstructing plant phylogeny. Ann. Missouri Bot. Gard. 81: 405-418.

Doyle, J. A., M. J. Donoghue, & E. A. Zimmer. 1994. Integration of morphological and ribosomal RNA data on the origin of angiosperms. Ann. Missouri Bot. Gard. 81: 419-450.

Doyle, J. J. 1991. Evolution of higher-plant glutamine synthetase genes: tissue specificity as a criterion for predicting orthology. Mol. Biol. Evol. 8:366-377.

Doyle, J. J. 1992. Gene trees and species trees: molecular systematics as one-character taxonomy. Syst. Bot. 17: 144-163.

Doyle, J. J., J. I Davis, R. J. Soreng, D. Garvin, & M. Anderson. 1992. Chloroplast DNA inversions and the origin of the grass family (Poaceae). Proc. Natl. Acad. Sci. USA 89: 7722-7726.

Doyle, J. J. & J. L. Doyle. 1987. A rapid DNA isolation procedure for small amounts of fresh leaf tissue. Phytochem. Bull. 19: 11-15.

Doyle, J. J. & J. L. Doyle. 1993. Chloroplast DNA phylogeny of the papilionoid legume tribe Phaseoleae. Syst. Bot. 18: 309-327.

Edgerton, M. D., & A. M. Jones. 1992. Localization of protein-protein interactions between subunits of phytochrome. Plant Cell 4:161-171.

Farris, J. S. 1988. Hennig86. Computer program and documentation. Port Jefferson Station, New York.

Farris, J. S. 1989. The retention index and the rescaled consistency index. Cladistics 5: 417-419.

Farris, J. S. 1994. RNA. Computer program and documentation. Naturhistoriska riksmuseet, Molekylärsystematiska laboratoriet, Stockholm.

Felsenstein, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. 27: 401-410.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783-791.

Felsenstein, J. 1988. Phylogenies from molecular sequences: inference and reliability. Annu. Rev. Genet. 22: 521-565.

Felsenstein, J. 1993. PHYLIP 3.5c - Phylogeny Inference Package, version 3.52c. Distributed by the author. Department of Genetics, University of Washington, Seattle.

Furuya, M. 1993. Phytochromes: their molecular species, gene families, and functions. Annu. Rev. Plant Physiol. Plant Mol. Biol. 44: 617-645.

Gaut, B. S., & M. T. Clegg. 1993. Nucleotide polymorphism in the *Adh1* locus of pearl millet (*Pennisetum glaucum*) (Poaceae). Genetics 135:1091-1097.

Goodman, M., J. Czelusniak, G. W. Moore, A. E. Romero-Herrera, & G. Matsuda. 1979. Fitting the gene lineage into its species lineage, a parsimony strategy illustrated by cladograms constructed from globin sequences. Syst. Zool. 28: 132-163.

Hamby, R. K., & E. A. Zimmer. 1992. Ribosomal RNA as a phylogenetic tool in plant systematics. Pp. 50-91 *in* P. S. Soltis, D. E. Soltis, & J. J. Doyle (editors), Molecular systematics of plants. Chapman and Hall, New York.

Hanelt, S., B. Braun, S. Marx, & H. A. W. Schneider-Poetsch. 1992. Phytochrome evolution: a phylogenetic tree with the first complete sequence of phytochrome from a cryptogamic plant (*Selaginella martensii* Spring). Photochem. Photobiol. 56: 751-758.

Harrison, R. G., D. M. Rand, & W. C. Wheeler. 1987. Mitochondrial DNA variation in field crickets across a narrow hybrid zone. Mol. Biol. Evol. 4: 144-158.

Haupt, W., & D.-P. Häder. 1994. Pp. 707-732 *in* R. E. Kendrick & G. H. M. Kronenberg (editors), Photomorphogenesis in plants. Kluwer Academic Publishers, Dordrecht.

Hendy, M. D., & D. Penny. 1989. A framework for the quantitative study of evolutionary trees. Syst. Zool. 38: 297-309.

Herendeen, P. S. 1992. The fossil record of the Leguminosae from the Eocene of southeastern North America. Pp. 85-160 *in* P. S. Herendeen & D. L. Dilcher (editors), Advances in legume systematics, part 4. Royal Botanic Gardens, Kew.

Hershey, H. P., R. F. Barker, K. B. Idler, J. L. Lissemore, & P. H. Quail. 1985. Analysis of cloned cDNA and genomic sequences for phytochrome: amino acid sequences for two gene products expressed in etiolated *Avena*. Nucleic Acids Res. 13: 8543-8559.

Heyer, A., & C. Gatz. 1992a. Isolation and characterization of a cDNA clone coding for potato type A phytochrome. Plant Mol. Biol. 18: 535-544.

Heyer, A., & C. Gatz. 1992b. Isolation and characterization of a cDNA clone coding for potato type B phytochrome. Plant Mol. Biol. 20: 589-600.

Hillis, D. M. 1987. Molecular versus morphological approaches to systematics. Annu. Rev. Ecol. Syst. 18: 23-42.

Hillis, D. M. 1995. Approaches for assessing phylogenetic accuracy. Syst. Biol. 44: 3-16.

Hillis, D. M., M. W. Allard, & M. M. Miyamoto. 1993. Analysis of DNA sequence data: phylogenetic inference. Methods Enzymol. 224: 456-487.

Holmquist, R. 1983 Transitions and transversions in evolutionary descent: an approach to understanding. J. Mol. Evol. 19: 134-144.

Huelsenbeck, J. P. 1995. Performance of phylogenetic methods in simulation. Syst. Biol. 44: 17-48.

Hughes, D. A. 1994. Histidine kinases hog the limelight. Nature 369: 187-188.

Ioerger, T. R., A. G. Clark, & T. Kao. 1990. Polymorphism at the self-incompatibility locus in Solanaceae predates speciation. Proc. Natl. Acad. Sci. U.S.A. 87: 9732-9735.

Iwabe, N., K. Kuma, M. Hasegawa, S. Osawa, & T. Miyata. 1989. Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. Proc. Natl. Acad. Sci. U.S.A. 86: 9355-9359.

Källersjö, M., J. S. Farris, A. G. Kluge, & C. Bult. 1992. Skewness and permutation. Cladistics 8: 275-287.

Kay, S. A., B. Keith, K. Shinozaki, & N.-H. Chua. 1989. The sequence of the rice phytochrome gene. Nucleic Acids Res. 17: 2865-2866

Kellogg, E. A. & C. S. Campbell. 1987. Phylogenetic analyses of the Gramineae. Pp. 310-322 *in* T. R. Soderstrom, K. W. Hilu, C. S. Campbell, and M. E. Barkworth (editors), Grass systematic and evolution. Smithsonian Institution, Washington, D. C.

Kellogg, E. A., & L. Watson. 1993. Phylogenetic studies of a large data set. I. Bambusoideae, Andropogonodae, and Pooideae (Gramineae). Botanical Review 59: 273-343.

Kimura, M. 1981. Estimation of evolutionary distances between homologous nucleotide sequences. Proc. Natl. Acad. Sci. U.S.A. 78: 454-458.

Klein, J., N. Takahata, & F. J. Ayala. 1993. MHC polymorphism and human origins. Sci. Amer. December 1993: 78-83.

Kluge, A. G. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). Syst.Zool. 38: 7-25.

Kluge, A. G., & J. S. Farris. 1969. Quantitative phyletics and the evolution of the anurans. Syst. Zool. 18:1-32.

Kolukisaoglu, H. U., B. Braun, W. F. Martin, & H. A. W. Schneider-Poetsch. 1993. Mosses do express conventional, distantly B-type-related phytochromes. Phytochrome of *Physcomitrella patens* Hedw. FEBS Lett. 334: 95-100.

Kraml, M. 1994. Light direction and polarization. Pp. 417-445 *in* R. E. Kendrick & G. H. M. Kronenberg (editors), Photomorphogenesis in plants. Kluwer Academic Publishers, Dordrecht.

Kuhner, M. K., & J. Felsenstein. 1994. A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. Mol. Biol. Evol. 11: 459-468.

Kumar, S., K. Tamura, & M. Nei. 1993. MEGA: Molecular Evolutionary Genetics Analysis, version 1.01. The Pennsylvania State University, University Park, Pennsylvania.

Lavin, M., J. J. Doyle, & J. D. Palmer. 1990. Evolutionary significance of the loss of the chloroplast-DNA inverted repeat in the Leguminosae subfamily Papilionoideae. Evolution 44: 390-402.

Lecointre, G., H. Philippe, H. L. V. Lê, & H. Le Guyader. 1993. Species sampling has a major impact on phylogenetic inference. Molec. Phyl. Evol. 2: 205-224.

Li, W.-H., & A. Zharkikh. 1994. What is the bootstrap technique? Syst. Biol. 43: 424-430.

Li, W.-H., & A. Zharkikh. 1995. Statistical test of DNA phylogenies. Syst. Biol. 44: 49-63.

Li, W.-H., & M. Guoy. 1990. Statistical tests of molecular phylogenies. 645-659 Methods Enzymol. 183: 645-659.

Li, X., & M. Noll. 1994. Evolution of distinct developmental functions of three *Drosophila* genes by acquisition of different *cis*-regulatory regions. Nature 367: 83-87.

Linder, H. P., & P. J. Rudall. 1993. The megagametophyte in *Anarthria* (Anarthriaceae, Poales) and its implications for the phylogeny of the Poales. Amer. J. Bot. 80: 1455-1464.

Maddison, W.P. & M. Slatkin. 1991. Null models for the number of evolutionary steps in a character on a phylogenetic tree. Evolution 45: 1184-1197.

Manhart. J. R. & J. D. Palmer. 1990. The gain of two chloroplast tRNA introns marks the green algal ancestor of land plants. Nature 345: 268-270.

Mathews, S., M. Lavin, & R. A. Sharrock. 1995. Evolution of the phytochrome gene family and its utility for phylogenetic analyses of angiosperms. Ann. Missouri Bot. Gard. 82: 296-321.

Maucher, H. P., R. Scheuerlein, & H. Schraudolf. 1992. Detection and partial sequence of phytochrome genes in the ferns *Anemia phyllitidis* (L.)Sw. (Schizaeaceae) and *Dryopteris filix-mas* L. (Polypodiaceae) by using the Polymerase-chain reaction technology. Photochem. Photobiol. 56: 759-763.

Meagher, R. B., S. Berry-Lowe, & K. Rice. 1989. Molecular evolution of the small subunit of ribulose bisphosphate carboxylase: nucleotide substitution and gene conversion. Genetics 123: 845-863.

Mindell, D. P. 1991. Aligning DNA sequences: homology and phylogenetic weighting. Pp. 73-89 *in* M. M. Miyamoto & J. Cracraft (editors), Phylogenetic analysis of DNA sequences. Oxford University Press, New York.

Mindell, D. P., J. W. Shultz, & P. W. Ewald. 1995. The AIDS pandemic is new, but is HIV new? Syst. Biol. 44: 77-92.

Morand, L. Z., D. G. Kidd, & J. C. Lagarias. 1993. Phytochrome levels in the green alga *Mesotaenium caldariorum* are light-regulated. Plant Physiol. 101: 97-103.

Muse, S. V., & B. S. Weir. 1992. Testing for equality of evolutionary rates. Genetics 132: 269-276.

Nei, M. 1991. Relative efficiencies of different tree-making methods for molecular data. Pp. 90-128 *in* M. M. Miyamoto & J. Cracraft (editors), Phylogenetic analysis of DNA sequences. Oxford University Press, New York.

Nei, M., & A. L. Hughes. 1991. Polymorphism and evolution of the major histocompatibility complex loci in mammals. Pp. 222-247 *in* R. K. Selander, A. G. Clark, & T. S. Whittam (editors), Evolution at the molecular level. Sinauer Associates, Sunderland, Massachusetts.

Okamoto, H. Y. Hirano, H. Abe, K. Tomizawa, M. Furuya, & M. Wada. 1993. The deduced amino acid sequence of phytochrome from *Adiantum* includes consensus motifs present in phytochrome B from seed plants. Plant Cell Physiol. 34: 1329-1334.

Olmstead, R. G. 1989. Phylogeny, phenotypic evolution, and biogeography of the *Scutellaria angustifolia* complex (Lamiaceae): inference from morphological and molecular data. Syst. Bot. 14: 320-338.

Olmstead, R. G., & J. A. Sweere. 1994. Combining data in phylogenetic systematics: an empirical approach using three molecular data sets in the Solanaceae. Syst. Biol. 43: 467-481.

Pamilo, P., & M. Nei. 1988. Relationships between gene trees and species trees. Mol. Biol. Evol. 5: 568-583.

Page, R. D. M. 1990. Component analysis: a valiant failure? Cladistics 6: 119-136.

Page, R. D. M. 1993a. Component, version 2.0. Biogeography and Conservation Laboratory. The Natural History

Museum, London.

Page, R. D. M. 1993b. Genes, organisms, and areas: The problem of multiple lineages. Syst. Biol. 42: 77-84.

Page, R. D. M. 1994. Maps between trees and cladistic analysis of historical associations among genes, organisms, and areas. Syst. Biol. 43: 58-77.

Parkinson, J. S., & E. C. Kofoid. 1992. Communication modules in bacterial signalling proteins. Annu. Rev. Genet. 26: 71-112.

Patterson, C. 1987. Introduction. Pp.1-22, *in* C. Patterson (editor), Molecules and morphology in evolution: conflict or compromise. Cambridge University Press, Cambridge.

Pratt, L. H. 1995. Phytochromes: Differential properties, expression patterns and molecular evolution. Photochem. Photobiol. 61: 10-21.

Quail, P. H. 1991. Phytochrome: a light-activated molecular switch that regulates plant gene expression. Ann. Rev. Genet. 25: 389-409.

Quail, P. H., M. T. Boylan, B. M. Parks, T. W. Short, Y. Xu, & D. Wagner. 1995. Phytochromes: photosensory perception and signal transduction. Science 268: 675-680.

Reed, J. W., P. Nagpal, D. S. Poole, M. Furuya, & J. Chory. 1993. Mutations in the gene for the red/far-red light receptor phytochrome B alter cell elongation and physiological responses throughout *Arabidopsis* development. Plant Cell 5: 147-157.

Rieseberg, L. H., & S. J. Brunsfeld. 1992. Molecular evidence and plant introgression. Pp. 151-176 *in* P. S. Soltis, D. E. Soltis, & J. J. Doyle (editors), Molecular systematics of plants. Chapman and Hall, New York.

Rieseberg, L. H., & D. E. Soltis. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. Evol. Trends Plants 5: 65-84.

Ritland, K., & J. E. Eckenwalder. 1992. Polymorphism, hybridization, and variable rate in molecular phylogenies. Pp. 404-428 *in* P. S. Soltis, D. E. Soltis, & J. J. Doyle (editors), Molecular systematics of plants. Chapman and Hall, New York.

Sanderson, M. J., & J. J. Doyle. 1992. Reconstruction of organismal and gene phylogenies from data on multigene families: concerted evolution, homoplasy and confidence. Syst. Biol. 41: 4-17.

Sato, N. 1988. Nucleotide sequence and expression of the phytochrome gene in *Pisum sativum*: differential regulation by light of multiple transcripts. Plant Mol. Biol. 11: 697-710.

Sato, N. 1990. Nucleotide sequence of a pseudogene for pea phytochrome reminiscent of an incorrect splicing event. Nucleic Acids Res. 18: 3632.

Schneider-Poetsch, H. A. W., B. Braun, S. Marx, & Anke Schaumburg. 1991. Phytochromes and bacterial sensor proteins are related by structural and functional homologies. FEBS Lett. 281: 245-249.

Sharrock, R. A., J. L. Lissemore, & P. H. Quail. 1986. Nucleotide and amino sequence of a *Cucurbita* phytochrome cDNA clone: identification of conserved features by comparison with *Avena* phytochrome. Gene 47: 287-295.

Sharrock, R. A., & P. H. Quail. 1989. Novel phytochrome sequences in *Arabidopsis thaliana*: structure, evolution, and differential expression of a plant regulatory photoreceptor family. Genes Dev. 3: 1745-1757.

Smith, H. 1994. Sensing the light environment: the functions of the phytochrome family. Pp. 377-416 *in* R. E. Kendrick & G. H. M. Kronenberg (editors), Photomorphogenesis in plants. Kluwer Academic Publishers, Dordrecht.

Sokal, R. R., & F. J. Rohlf. 1981. Taxonomic congruence in the Leptopodomorpha re-examined. Syst. Zool. 30: 309-325.

Soltis, D. E., & P. S. Soltis. 1993. Molecular data and the dynamic nature of polyploidy. Crit. Rev. Plant Sci.

12: 243-273.

Soltis, D. E., P. S. Soltis, & B. G. Milligan. 1992. Intraspecific chloroplast DNA variation: systematic and phylogenetic implications. Pp. 117-150 *in* P. S. Soltis, D. E. Soltis, & J. J. Doyle (editors), Molecular systematics of plants. Chapman and Hall, New York.

Somers, D. E., R. A. Sharrock, J. M. Tepperman, & P. H. Quail. 1991. The *hy*3 long hypocotyl mutant of *Arabidopsis* is deficient in phytochrome B. The Plant Cell 3: 1263-1274.

Soreng, R. J., & J. I Davis. 1994. Phylogenetic structure of the grass family (Poaceae) as inferred from chloroplast DNA restriction site variation. Amer. J. Bot. 81 (Supplement): abstract 532.

Swofford, D. L. 1991. When are phylogeny estimates incongruent? Pp. 295-333 *in* M. M. Miyamoto & J. Cracraft (editors), Phylogenetic analysis of DNA sequences. Oxford University Press, New York.

Swofford, D. L. 1993. PAUP: Phylogenetic analysis using parsimony version 3.1. Illinois Natural History Survey, Champaign.

Takahata, N. 1989. Gene genealogy in three related populations: consistency probability between gene and population trees. Genetics 122: 957-966.

Thomas, B. R., V. S. Ford, E. Pichersky, & L. D. Gottlieb. 1993. Molecular characterization of duplicate cytosolic phosphoglucose isomerase genes in Clarkia and comparison to the single gene in Arabidopsis. Genetics 135: 895-905.

Thümmler, F., M. Dufner, P. Kreisl, & P. Dittrich. 1992. Molecular cloning of a novel phytochrome gene of the moss *Ceratodon purpureus* which encodes a putative light-regulated protein kinase. Plant Molec. Biol. 20: 1003-1017.

Townrow, J. A. 1968. A fossil *Selaginella* from the Permian of New South Wales. J. Linnean Soc. (Bot.) 61: 13-23.

Wada, M., & A. Kadota. 1989. Photomorphogenesis in lower green plants. Annu. Rev. Plant Physiol. Mol. Biol. 40: 169-191.

Wada, M., & M. Sugai. 1994. Photobiology of ferns. Pp. 783-802 *in* R. E. Kendrick & G. H. M. Kronenberg (editors), Photomorphogenesis in plants. Kluwer Academic Publishers, Dordrecht.

Wagner, A., N. Blackstone, P. Cartwright, M. Dick, B. Misof, P. Snow, G. P. Wagner, J. Bartels, M. Murtha, & J. Pendleton. 1994. Surveys of gene families using polymerase chain reaction: PCR selection and drift. Syst. Biol. 43: 250-261.

Watson, L., & M. J. Dallwitz. 1992. Grass genera of the world. CAB International, Wallingford, UK.

Waye, M. M. Y., M. E. Verhoeyen, P. T. Jones, & G. Winter. 1985. *Eco*K selection vectors for shotgun cloning into M13 and deletion mutagenesis. Nucleic Acids Res. 13: 8561-8571.

Wester, L., D. E. Somers, T. Clack, & R. A. Sharrock. 1994. Transgenic complementation of the *hy3* phytochrome B mutation and response to *PHYB* gene copy number in *Arabidopsis*. Plant J. 5: 261-272.

Wheeler, E., & P. Baas. 1992. Fossil wood of the Leguminosae: a case study in xylem evolution and ecological anatomy. Pp. 281-301 *in* P. S. Herendeen & D. L. Dilcher (editors), Advances in legume systematics, part 4. Royal Botanic Gardens, Kew.

Winands, A., G. Wagner, S. Marx, & H. A. W. Schneider-Poetsch. 1992. Partial nucleotide sequence of phytochrome from the zygnematophycean green alga *Mougeotia*. Photochem. Photobiol. 56: 765-770.

Wu, C.-I., & W.-H. Li. 1985. Evidence for higher rates of nucleotide substitution in rodents than in man. Proc. Natl. Acad. Sci. U.S.A. 82: 1741-1745.

Zimmer, E. A., S. L. Martin, S. M. Beverly, Y. W. Kan, & A. C. Wilson. 1980. Rapid duplication and loss of genes coding for the alpha-chains of hemoglobin. Proc. Natl. Acad. Sci. U.S.A. 77: 2158-2162.

Zurawski, G., M. T. Clegg, & A. H. D. Brown. 1984. The nature of nucleotide sequence substitution between barley and maize chloroplast DNA. Genetics 106: 735-749.

APPENDICES

Appendix A

Alignment Of Amino Acids Of Fully Characterized
Phytochrome Genes

All available full length phytochrome amino acid sequences, and 776 residues from *Ceratodon*. [1] Hanelt et al., 1992; [2] Thümmler et al., 1992; [3] Okamoto, et al., 1993; [4] Sharrock & Quail, 1989; [5] Sharrock et al., 1986; [6] Sato, 1988; [7] Heyer & Gatz, 1992a; [8] Hershey et al., 1985; [9] Kay et al., 1989; [10] Christensen & Quail, 1989; [11] Clack et al., 1994; [12] Heyer & Gatz, 1992b; [13] Dehesh et al., 1991. The triangle denotes the chromophore attachment site. The sequences amplified in this study correspond to residues 329-431.

```
10          20          30          40          50
                                    *           *           *           *           *
Selaginella (sm) [1]          ---------- ---------- ---------- ---------- MSTTKLTYSS
Ceratodon (cp) [2]            ---------- ---------- ---------- ---------- MSATKKTYSS
Adiantum (ac) [3]             ---------- ---------- ---------- ---------- MSSTRHSYSS
Arabidopsis PHYA (atA)[4]     ---------- ---------- ---------- -------MSG SRPT--QSSE
Cucurbita PHYA (cpA)[5]       ---------- ---------- ---------- -------MST SRPS--QSSS
Pisum PHYA (psA)[6]           ---------- ---------- ---------- -------MST TRPS--QSSN
Solanum PHYA (stA)[7]         ---------- ---------- ----MSTSLF ASDSDQLMSS SRPS--QSST
Avena PHYA (asA)[8]           ---------- ---------- ---------- -------MSS SRPA--SSSS
Oryza PHYA (osA)[9]           ---------- ---------- ---------- -------MSS SRPTQCSSSS
Zea PHYA (zmA)[10]            ---------- ---------- ---------- -------MSS SRPAHSSSSS
Arabidopsis PHYB (atB)[4]     --MVSGVGGS GGGRGGGRGG EEEPSSSHTP NNRRGGEQAQ SSGTKSLRPR
Arabidopsis PHYD (atD)[11]    MVSGGGSKTS GGEAASSGHR RSRHTSAAEQ AQSSANKALR WQNQQPQNHG
Arabidopsis PHYE (atE)[11]    ---------- ---------- ---------- ---------- MGFESSSSAA
Solanum PHYB (stB)[12]        ---------- ---------- -------MAS GSRTKHSHHS SSQAQSSGTS
Oryza PHYB (osB)[13]          MASGSRATPT RSPSSARPAA PRHQHHHSQS SGGSTSRAGG GGGGGGGGGG
Arabidopsis PHYC (atC)[4]     ---------- ---------- ---------- ---------- --MSSNTSRS
ANG                           ---------- ---------- ---------- ---------- ----------
CON                           ---------- ---------- ---------- ---------- ----------

             60         70         80         90        100        110        120
              *          *          *          *          *          *          *
sm   GSSAKSKHSV RVAQTTADAK LHAVYEESGE SGDSFDYSKS INATKSTGET IPAQ----AV -TAYLQRMQR
cp   TTSAKSKHSV RVAQTTADAA LEAVYEMSGD SGDSFDYSKS VGQSAE--SV P-----AGAV -TAYLQRMQR
ac   GGSGKSKHGR RIAQTSANAK LYAAYEESSE SGS-FDYSQS VSAGKEGI-- -----SSQLV -TAYLQRMQR
atA  GSRRSRHSAR IIAQTTVDAK LHADFE---E SGSSFDYSTS VRVTGPVV-- ENQPPRSDKV TTTYLHHIQK
cpA  NSGRSRHSTR IIAQTSVDAN VQADFE---E SGNSFDYSSS VRVTSDVS-- GDQQPRSDKV TTAYLHHIQK
psA  NSGRSRNSAR IIAQTTVDAK LHATFE---E SGSSFDYSSS VRVSGSVD-- GDQQPRSNKV TTAYLNHIQR
stA  TSSRSKHSAR IIAQTSIDAK LHADFE---E SGDSFDYSSS VRVTNVAE-- GEQRPKSDKV TTAYLHQIQK
asA  SRNRQSSQAR VLAQTTLDAE LNAEYE---E SGDSFDYSKL VEAQRDGP-- PVQQGRSEKV -IAYLQHIQK
osA  SRTRWSSRAR ILAQTTLDAE LNAEYE---E YGDSFDYSKL VEAQRTTG-- PEQQARSEKV -IAYLHHIQR
zmA  SRTRQSSRAR ILAQTTLDAE LNAEYE---E SGDSFDYSKL VEAQRSTP-- PEQQGRSGKV -IAYLQHIQR
atB  SNTESMSKSK AIQQYTVDAR LHAVFEQSGE SGKSFDYSQS LKTTTYGSSV PEQQ------ ITAYLSRIQR
atD  GGTESTNKNK AIQQYTVDAR LHAVFEQSGE SGKSFDYSQS LKTAPYDSSV PEQQ------ ITAYLSRIQR
atE  SNMKPQPQKS NTAQYSVDAA LFADFAQSIY TGKSFNYSKS VISPPN--HV PDEH------ ITAYLSNIQR
stB  NVNYKDSISK AIAQYTADAR LHAVFEQSGE SGKFFDYSQS VKTTTQ--SV PERQ------ ITAYLTKIQR
osB  GAAAAESVSK AVAQYTLDAR LHAVFEQSGA SGRSFDYTQS LRASPT--PS SEQQ------ IAAYLSRIQR
atC  CSTRSRQNSR VSSQVLVDAK LHGNFE---E SERLFDYSAS INLNM---PS SSCEIPSSAV -STYLQKIQR
ANG  ---------- ---Q---DA- ---------- ----F-Y--- ---------- ---------- ---YL--IQ-
CON  ---------- ---Q----A- ---------- ----F-Y--- ---------- ---------- ---YL---Q-

            130        140        150        160        170        180        190
              *          *          *          *          *          *          *
sm   GGLVQPFGCM LAV-EEGSFR VIAFSDNAGE MLDLMP-QSV PSL-GSGQQD VLTIGTDART LFTAAAS-AL
cp   EGLIQNFGCM VAV-EEPNFC VIAYSENASE FLDLIP-QAV PSM-GEM--D VLGIGTDIRT LFTPSSSAAL
ac   GGLVQQFGCL IAV-EEETFR VLHMCE-APE MLDVAT-QAV PTM-GQY--S RLCIGADVRT LLSPASASAL
atA  GKLIQPFGCL LAL-DEKTFK VIAYSENASE LLTMAS-HAV PSV-GEH--P VLGIGTDIRS LFTAPSASAL
cpA  GKLIQPFGCL LAL-DDKTFK VIAYSENAPE MLTMVS-HAV PSM-GDY--P VLGIGTDVRT IFTAPSASAL
psA  GKQIQPFGCL LAL-DEKTCK VVAYSENAPE MLTMVS-HAV PSV-GDH--P ALGIGTDIRT VFTAPSASAL
stA  GKFIQPFGCL LAL-DEKTLK VIAFSENAPE MLTMVS-HAV PSV-GEH--P VLGIGIDIRT IFTGPSGAAL
asA  GKLIQTFGCL LAL-DEKSFN VIAFSENAPE MLTTVS-HAV PSVD---DPP RLGIGTNVRS LFSDQGATAL
osA  AKLIQPFGCL LAL-DEKTFN VIALSENAPE MLTTVS-HAV PSVD---DPP KLRIGTNVRS LFTDPGTTAL
zmA  GKLIQPFGCL LAL-DEKSFR VIAFSENAPE MLTTVS-HAV PNVD---DPP KLGIGTNVRS LFTDPGATAL
atB  GGYIQPFGCM IAV-DESSFR IIGYSENARE MLGIMP-QSV PTLE---KPE ILAMGTDVRS LFTSSSSILL
atD  GGYTQPFGCL IAV-EESTFT IIGYSENARE MLGLMS-QSV PSIE-D-KSE VLTIGTDLRS LFKSSSYLLL
atE  GGLVQPFGCL IAV-EEPSFR ILGLSDNSSD FLGLLSLPST SHS-GEFDKV KGLIGIDART LFTPSSGASL
stB  GGHIQPFGCM IAV-DEASFR VIAYSENACE MLSLTP-QSV PSLE---KCE ILTIGTDVRT LFTPSSSVLL
osB  GGHIQPFGCT LAVADDSSFR LLAYSENTAD LLDLSPHHSV PSLDSSAVPP PVSLGADARL LFAPSSAVLL
atC  GMLIQPFGCL IVV-DEKNLK VIAFSENTQE MLGLIP-HTV PSME---QRE ALTIGTDVKS LFLSPGCSAL
ANG  ----Q-FGC- ---------- ----S-N--- -L-------- ---------- ----G----- -F-------L
CON  ----Q-FGC- ---------- ---------- -L-------- ---------- ----G----- ---------L
```

```
             200        210        220        230        240        250        260
               *          *          *          *          *          *          *
sm   EKAAGAVDLS MLNPIWVQSK TSAKPFYAIV HRIDVGLVMD LEPVKASDTR VGSAAGALQS HKLAAKAISR
cp   EKAAATQDIS LLNPITVHCR RSGKPLYAIA HRIDIGIVID FEAVKMIDVP VSAAAGALQS HKLAARAITR
ac   DRVIGVVDVS MFNPITVQSR SSGKPFYAIL HRNDVGLVID LEPIRPDDAS I-TG-GALQS HKLAAKAIAR
atA  QKALGFGDVS LLNPILVHCR TSAKPFYAII HRVTGSIIID FEPVKPYEVP M-TAAGALQS YKLAAKAITR
cpA  LKALGFGEVT LLNPILVHCK TSGKPFYAIV HRVTGSLIID FEPVKPYEGP V-TAAGALQS YKLAAKAITR
psA  QKALGFAEVS LLNPILVHCK TSGKPFYAII HRVTGSLIID FEPVKPYEVP M-TAAGALQS YKLAAKAITR
stA  QKALGFGEVS LLNPVLVHCK NSGKPFYAIV HRVTGSLIID FEPVKPYEVP M-TAAGALQS YKLAAKAITR
asA  HKALGFADVS LLNPILVQCK TSGKPFYAIV HRATGCLVVD FEPVKPTEFP A-TAAGALQS YKLAAKAISK
osA  QKALGFADVS LLNPILVQCK TSGKPFYAIV HRATGCLVVD FEPVKPTEFP A-TAAGALQS YKLAAKAISK
zmA  QKALGFADVS LLNPILVQCK TSGKPFYAIV HRATGCLVVD FEPVKPTEFP A-TAAGALQS YKLAAKAISK
atB  ERAFVAREIT LLNPVWIHSK NTGKPFYAIL HRIDVGVVID LEPAR-TEDP ALSIAGAVQS QKLAVRAISQ
atD  ERAFVAREIT LLNPIWIHSN NTGKPFYAIL HRVDVGILID LEPAR-TEDP ALSIAGAVQS QKLAVRAISH
atE  SKAASFTEIS LLNPVLVHSR TTQKPFYAIL HRIDAGIVMD LEPAK-SGDP ALTLAGAVQS QKLAVRAISR
stB  ERAFGAREIT LLNPIWIHSK NSGKPFYAIL HRVDVGIVID LEPAR-TEDP ALSIAGAVQS QKLRSEGLFL
osB  ERAFAAREIS LLNPLWIHSR VSSNPFYAIL HRIDVGVVID LEPAR-TEDP ALSIAGAVQS QKLVVRAISR
atC  EKAVDFGEIS ILNPITLHCR SSSKPFYAIL HRIEEGLVID LEPVSPDEVP V-TAAGALRS YKLAAKSISR
ANG  --A------- -LNP------ ----PFYAI- HR-------D -EP------- ----AGA--S -KL-------
CON  ---------- --NP------ ----P-YAI- HR-------D -E-------- -----GA--S -KL-------

             270        280        290        300        310        320        330
               *          *          *          *          *          *          *
sm   LQSLP-GGDI GLLCDTVVEE VRDVTGYDLV MAYKFHEDEH GEVVAEIRRS DLEPYLGLHY PATDIPQASR
cp   LQALP-GGDI ELLCDTIVEE VRELTGYDRV MAFKFHEDEH GEVVAEIRRM DLEPYMGLHY PATDIPQASR
ac   LQSLP-GGDI GLLCDSVVEE VHELTGFDRV MAYKFHEDEH GEVVAEIRRT DLEPYIGLHY PATDIPQAAR
atA  LQSLP-SGSM ERLCDTMVQE VFELTGYDRV MAYKFHEDDH GEVVSEVTKP GLEPYLGLHY PATDIPQAAR
cpA  LQSLP-SGSM ARLCDTMVQE VFELTGYDRV MAYKFHDDDH GEVISEVAKP GLQPYLGLHY PATDIPQAAR
psA  LQSLA-SGSM ERLCDTMVQE VFELTGYDRV MAYKFHEDDH GEVIAEIAKP GLEPYLGLHY PATDIPQAAR
stA  LQSLP-SGSM ERLCDTMVQE VFELTGYDRV MGYKFHDDDH GEVVSEITKP GLEPYLGLHY PATDIPQAAR
asA  IQSLP-GGSM EVLCNTVVKE VFDLTGYDRV MAYKFHEDDH GEVFSEITKP GLEPYLGLHY PATDIPQAAR
osA  IQSLP-GGSM EVLCNTVVKE LFDLTGYDRV MAYKFHEDDH GEVFAEITKP GLEPYLGLHY PATDIPQAAR
zmA  IQSLP-GGSM EALCNTVVKE VFDLTGYDRV MAYKFHEDEH GEVFAEITKP GIEPYIGLHY PATDIPQAAR
atB  LQALP-GGDI KLLCDTVVES VRDLTGYDRV MVYKFHEDEH GEVVAESKRD DLEPYIGLHY PATDIPQASR
atD  LQSLP-SGDI KLLCDTVVES VRDLTGYDRV MVYKFHEDEH GEVVAESKRN DLEPYIGLHY PATDIPQASR
atE  LQSLP-GGDI GALCDTVVED VQRLTGYDRV MVYQFHEDDH GEVVSEIRRS DLEPYLGLHY PATDIPQAAR
stB  ICNHFLVGTL KLLCDTVVES VRELTGYDRV MVYKFHEDEH GEVVAESKRS DLEPYIGLHY PATDIPQASR
osB  LQALP-GGDV KLLCDTVVEH VRELTGYDRV MVYRFHEDEH GEVVAESRRS NLEPYIGLHY PATDIPQASR
atC  LQALP-SGNM LLLCDALVKE VSELTGYDRV MVYKFHEDGH GEVIAECCRE DMEPYLGLHY SATDIPQASR
ANG  -------G-- --LC---V-- ---LTGYDRV M-Y-FH-D-H GEV--E---- ---PY-GLHY -ATDIPQA-R
CON  -------G-- --LC---V-- ----TG-D-V M---FH-D-H GEV--E---- ---PY-GLHY -ATDIPQA-R
                                                                     >TARGET

             340        350        360        370        380        390        400
               *          *          *          *      ▽   *          *          *
sm   FLFMKNRVRM ICDCSAPPVK ITQDKELRQP ISLAGSTLRA PHGCHAQYMG NMGSVASLVM AMIINDNDE-
cp   FLLMKNRVRL IADCYASPVK LIQDPDIRQP VSLAGSTLRA PHGCHAQYMG NMGSIASLVM AVIINDNEE-
ac   FLFMKNRVRM ICDCRLPPVK LIQDKTLSQP MSLTGSTLRA PHGCHTQYMA NMNSISSLVM AVIVNDSDDD
atA  FLFMKNKVRM IVDCNAKHAR VLQDEKLSFD LTLCGSTLRA PHSCHLQYMA NMDSIASLVM AVVVNEEDGE
cpA  FLFMKNKVRM IVDCRAKHLK VLQDEKLQFD LTLCGSTLRA PHSCHLQYME NMNSIASLVM AVVVNEGDEE
psA  FLFMKNKVRM IVDCNAKHVK VLQDEKLPFD LTLCGSTLRA PHSCHLQYMA NMDSIASLVM AVVVNDSDED
stA  FLFMKNKVRM ICDCRAKHVK VVQDEKLPFD LTLCGSTLRA PHYCHLQYME NMNSIASLVM AVVVNDGDEE
asA  LLFMKNKVRM ICDCRARSIK VIEAEALPFD ISLCGSALRA PHSCHLQYME NMNSIASLVM AVVVNENEED
osA  FLFMKNKVRM ICDCRARSIK IIEDESLHLD ISLCGSTLRA PHSCHLQYME NMNSIASLVM AVVVNENEDD
zmA  FLFMKNKVRM ICDCRARSVK IIEDEALSID ISLCGSTLRA PHSCHLKYME NMNSIASLVM AVVVNENEED
atB  FLFKQNRVRM IVDCNATPVL VVQDDRLTQS MCLVGSTLRA PHGCHSQYMA NMGSIASLAM AVIINGNEDD
atD  FLFKQNRVRM IVDCYASPVR VVQDDRLTQF ICLVGSTLRA PHGCHAQYMT NMGSIASLAM AVIINGNEED
atE  FLFKQNRVRM ICDCNATPVK VVQSEELKRP LCLVNSTLRA PHGCHTQYMA NMGSVASLAL AIVVKGKD--
stB  FLFKQNRVRM IVDCHATPVR VTQDESLMQP LCLVGSTLRA PHGCHAQYMA NMGSIASLTL AVIINGNDEE
osB  FLFRQNRVRM IADCHAAPVR VIQDPALTQP LCLVGSTLRS PHGCHGQYMA NMGSIASLVM AVIISSGGDD
atC  FLFMRNKVRM ICDCSAVPVK VVQDKSLSQP ISLSGSTLRA PHGCHAQYMS NMGSVASLVM SVTINGSDSD
ANG  -LF--N-VRM I-DC-A---- ------L--- --L--S-LR- PH-CH--YM- NM-S-ASL-- ----------
CON  -L---N-VR- I-DC------ ---------- --L--S-LR- PH-CH--YM- NM-S--SL-- ----------
```

```
            410        420        430        440        450        460        470
              *          *          *          *          *          *          *
.sm  -PSGGGGGGG QHKGRRLWGL VVCHHTSPRS VPF-LRSACE FLMQVFGLQL NMEAAVAAHV REKHILRTQT
cp   -----YSRGA IQRGRKLWGL VVCQHTSPRT VPFPLRSVCE FLMQVFGMQL NLHVELAAQL REKHILRTQT
ac   ------SAGH SSQGIKLWGL VVCHHTSPRY VPFPVRSACE FLMQVFSLQL NMEVGMAAQV REKHILRTQT
atA  GD-APDATTQ PQKRKRLWGL VVCHNTTPRF VPFPLRYACE FLAQVFAIHV NKEVELDNQM VEKNILRTQT
cpA  NE---GPALQ QQKRKRLWGL VVCHNSSPRF VPFPLRYACE FLAQVFAIHV NKELELENQI IEKNILRTQT
psA  GD--SADAVL PQKKKRLWGL VVCHNTTPRF VPFPLRYACE FLAQVFAIHV NKEIELEYQI LEKNILRTQT
stA  GE--SSDSSQ SQKRKRLWGL VVSHNTTPRF APFPLRYACE FLAQVFAILV NKELELENQF LEKNILRTQT
asA  DEAESEQPAQ QQKKKKLWGL LVCHHESPRY VPFPLRYACE FLAQVFAVHV NREFELEKQL REKNILKMQT
osA  DEVGADQPAQ QQKRKKLWGL LVCHHESPRY VPFPLRYACE FLAQVFAVHV NKEFELERQV REKSILRMQT
zmA  DEPEPEQPPQ QQKKKRLWGL IVCHHESPRY VPFPLRYACE FLAQVFAVHV NKEFELEKQI REKNILRMQT
atB  G----SNVAS GRSSMRLWGL VVCHHTSSRC IPFPLRYACE FLMQAFGLQL NMELQLALQM SEKRVLRTQT
atD  G---NGVNTG GRNSMRLWGL VVCHHTSARC IPFPLRYACE FFMQAFGLQL NMELQLASQV SEKRVLRMQT
atE  ---------- ---SSKLWGL VVGHHCSPRY VPFPLRYACE FLMQAFGLQL QMELQLASQL AEKKAMRTQT
stB  ------AVGG GRNSMRLWGL VVGHHTSVRS IPFPLRYACE FLMQAFGLQL NMELQLASQL SEKHVLRTQT
osB  D--HNIARGS IPSAMKLWGL VVCHHTSPRC IPFPLRYACE FLMQAFGLQL NMELQLAHQL SEKHILRTGT
atC  E-----MNRD LQTGRHLWGL VVCHHASPRF VPFPLRYACE FLTQVFGVQI NKEAESAVLL KEKRILQTQS
ANG  ---------- ------LWGL -V-H----R- -PFPLRYACE F--Q-F---- --E------- -EK-------
CON  ---------- ------LWGL -V------R- -PF--R--CE F--Q-F---- ---------- -EK-------

            480        490        500        510        520        530        540
              *          *          *          *          *          *          *
sm   LLCDMLLRDA -PIGIVSQSP NIMDLVKCDG AALYYGKRFW LLGITPSEAQ IKDIAEWLLE HH-KDSTGLS
cp   LLCDMLMRDA -PLGIVSQTP NIMDLVKCDG AALYYGKRVW LLGTTPTENQ IKEIADWLLE HH-MDSTGLS
ac   LLCDMLLRDA -PIGIVSQSP NIMDLVTCDG AALYYGKKCW LLGTTPTEAQ IVDIAAWLLD CH-KDSTGLS
atA  LLCDMLMRDA -PLGIVSQSP NIMDLVKCDG AALLYKDKIW KLGTTPSEFH LQEIASWLCE YH-MDSTGLS
cpA  LLCDMLMRDA -PLGIVSRSP NIMDLVKSDG AALLYKKKIW RLGLTPNDFQ LLDIASWLSE YH-MDSTGLS
psA  LLCDMLMRDA -PLGIVSQSP NIMDLVKCDG AALFYRNKLW LLGATPTESQ LREIALWMSE YH-TDSTGLS
stA  LLCDMLMRDA -PLGIVSQSP NIMDLIKCDG AALLYKNKIH RLGMNPSDFQ LHDIVSWLCE YH-TDSTGLS
asA  MLSDMLFREA SPLTIVSGTP NIMDLVKCDG AALLYGGKVW RLQNAPTESQ IHDIAFWLSD VH-RDSTGLS
osA  MLSDMLFRES SPLSIVSGTP NIMDLVKCDG AALLYGGKVW RLQNAPTESQ IRDIAFWLSD VH-RDSTGLS
zmA  MLSDMLFKES SPLSIVSGSP NIMDLVKCDG AALLYGDKVW RLQTAPTESQ IRDIAFWLSE VH-GDSTGLS
atB  LLCDMLLRDS -PAGIVTQSP SIMDLVKCDG AAFLYHGKYY PLGVAPSEVQ IKDVVEWLLA NH-ADSTGLS
atD  LLCDMLLRDS -PAGIVTQRP SIMDLVKCNG AAFLYQGKYY PLGVTPTDSQ INDIVEWLVA NH-SDSTGLS
atE  LLCDMLLRDT -VSAIVTQSP GIMDLVKCDG AALYYKGKCW LVGVTPNESQ VKDLVNWLVE NHGDDSTGLT
stB  LLCDMLLRDS -PPGIVTQSP SIMDLVKCDG ALLYYQGKYY PLGVTPTEAQ IKDIVEWLLA YH-GDSTGLS
osB  LLCDMLLRDS -PTGIVTQSP SIMDLVKCDG AALYYHGKYY PLGVTPTEVQ IKDIIEWLTM CH-GDSTGLS
atC  VLCDMLFRNA -PIGIVTQSP NIMDLVKCDG AALYYRDNLW SLGVTPTETQ IRDLIDWVLK SH-GGNTGFT
ANG  -L-DML---- ----IV---P -IMDL-K--G A---Y----- -----P---- ------W--- -H----TG--
CON  -L-DML---- ----IV---P -IMDL----G A---Y----- -----P---- ------W--- -H----TG--

            550        560        570        580        590        600        610
              *          *          *          *          *          *          *
sm   TDSLADAGYP GAASLGDEVC GMAAAKITAK DFLFWFRSHT AKEVKWGGAK HDPDDKDDGR KMHPRSSSKA
cp   TDSLADANYP GAHLLGDAVC GMAAAKITAK DFLFWFRSHT ATEVKWGGAK HDPDEKDDGR KMHPRSSFKA
ac   TDSLAKTGYP EASCLGDAVC GLAAAKITAT DFLFWFRSHT AKEVRWGGAR HDPEERDDGR RMHPRSSFKA
atA  TDSLHDAGFP RALSLGDSVC GMAAVRISSK DMIFWFRSHT AGEVRWGGAK HDPDDRDDAR RMHPRSSFKA
cpA  TDSLYDAGYP GAIALGDEVC GMAAVRITNN DMIFWFRSHT ASEIRWGGAK HEHGQKDDAR KMHPRSSFKA
psA  TDSLSDAGFP GALSLSDTVC GMAAVRITSK DIVFWFRSHT AAEIRWGGAK HEPGDQDDGR KMHPRSSFKA
stA  TDSLYDAGFP GALALGDAVC GMAAVRISDK DWLFWYRSHT AAEVRWGGAK HEPGEKDDGR KMHPRSSFKG
asA  TDSLHDAGYP GAAALGDMIC GMAVAKINSK DILFWFRSHT AAEIRWGGAK NDPSDMDDSR RMHPRLSFKA
osA  TDSLHDAGYP GAAALGDMIC GMAVAKINSK DILFWFRSHT AAEIRWGGAK HDPSDKDDSR RMHPRLSFKA
zmA  TDSLQDAGYP GAASLGDMIC GMAVAKITSK DILFWFRSHT AAEIKWGGAK HDPSDKDDNR RMHPRLSFKA
atB  TDSLGDAGYP GAAALGDAVC GMAVAYITKR DFLFWFRSHT AKEIKWGGAK HHPEDKDDGQ RMHPRSSFQA
atD  TDSLGDAGYP RAAALGDAVC GMAVACITKR DFLFWFRSHT EKEIKWGGAK HHPEDKDDGQ RMNPRSSFQT
atE  TDSLVDAGYP GAISLGDAVC GVAAAEFSSK DYLLWFRSNT ASAIKWGGAK HHPKDKDDAG RMHPRSSFTA
stB  TDSLPDAGYP GAASLGDAVC GMAVAYITSK DFLFWFRSHT AKEIKWGGAK HHPEDKDDGQ RMHPRSSFKA
osB  TDSLADAGYS GAAALGDAVS GMAVAYITPS DYLFWFRSHT AKEIKWGGAK HHPEDKDDGQ RMHPRSSFKA
atC  TESLMESGYP DASVLGESIC GMAAVYISEK DFLFWFRSST AKQIKWGGAR HDPNDR-DGK RMHPRSSFKA
ANG  T-SL---G-- -A--L----- G-A------- D---W-RS-T -----WGGA- -------D-- -M-PR-SF--
CON  T-SL------ -A--L----- G-A------- D---W-RS-T -----WGGA- -------D-- -M-PR-S---
```

```
                 620        630        640        650        660        670        680
                   *          *          *          *          *          *          *
sm   FLEVVKRRSL PWEDVEMDAI HSLQLILRGS FQDIDDSDTK TM-IHAR--- LNDLKLQGMD ELSTVANEMV
cp   FLEVVNKRSP PWEDVEMDAI HSLQLILRGS FRDIADSDTK TM-IHAR--- LNDLKLQGVE ERNALANEMS
ac   FLEVVKQQSL PWEDVEMDAI HSLQLILRGS FQDIDDSNTK TM-IHAR--- LNDLKLQGLD ELSTVASEMV
atA  FLEVVKTRSL PWKDYEMDAI HSLQLILRNA FKDSETTDVN TKVIYSK--- LNDLKIDGIQ ELEAVTSEMV
cpA  FLEVVKTRSL PWKDYEMDAI HSLQLILRNT FKDTDATEIN RKSIQTT--- LGDLKIEGRQ ELESVTSEMV
psA  FLEVVKARSV PWKDFEMDAI HSLQLILRNA SKDTDIIDLN TKAINTR--- LNDLKIEGMQ ELEAVTSEMV
stA  FLEVVKTRSI PWKDYEMDRI HSLQLILRNA FKDADAVNSN TISIHTK--- LNDLKIDGMQ ELEAVTAEMV
asA  FLEVVKMKSL PWSDYEMDAI HSLQLILRGT LNDASKPKRE ASLDNQI--- -GDLKLDGLA ELQAVTSEMV
osA  FLEVVKMKSL PWNDYEMDAI HSLQLILRGT LNDDIKPTRA ASLDNQV--- -GDLKLDGLA ELQAVTSEMV
zmA  FLEVVKTKSL PWSDYEMDAI HSLQLILRGT LNDASKPAQA SGLDNQI--- -GDLKLDGLA ELQAVTSEMV
atB  FLEVVKSRSQ PWETAEMDAI HSLQLILRDS FKESEAAMNS KVVDGVVQPC RDMAGEQGID ELGAVAREMV
atD  FLEVVKSRCQ PWETAEMDAI HSLQLILRDS FKESEAMDSK AAAAGAVQPH GDDMVQQGMQ EIGAVAREMV
atE  FLEVAKSRSL PWEISEIDAI HSLRLIMRES FTSSRPVLSG NGVARDAN-- ---------- ELTSFVCEMV
stB  FLEVVKSRSS PWENAEMDAI HSLQLILRDS FKDAEASNSK AIVHAH---- LGEMELQGID ELSSVAREMV
osB  FLEVVKSRSL PWENAEMDAI HSLQLILRDS FRDSAEGTSN SKAIVNGQVQ LGELELRGID ELSSVAREMV
atC  FMEIVRWKSV PWDDMEMDAI NSLQLIIKGS LQEEH---SK TVVDVP---- LVDNRVQKVD ELCVIVNEMV
ANG  F-E------- PW---E-D-I -SL-LI---- ---------- ---------- ---------- E------EMV
CON  F-E------- PW---E-D-I -SL-LI---- ---------- ---------- ---------- E------EM-

                 690        700        710        720        730        740   INTRON 750
                   *          *          *          *          *          *          *
sm   RLIETATAPI LAVDSSGFIN GWNAKVADVT GLPVTEAMGR SLAKELVLHE SADMVERLLY LALQGDEEQN
cp   RVLETAAAPI LAVDSRGMIN AWNAKIAQVT GLPVEEAMHC SLTKDLVLDE SVVVVERLLS LALQGEEEQN
ac   RLIETATAPI LAVDGQGLIN GWNGKVAELT GLSFETAMGK SLAKELVHEE SKTIVERVLH LALEGEEEQD
atA  RLIETATVPI LAVDSDGLVN GWNTKIAELT GLSVDEAIGK HFLT-LVEDS SVEIVKRMLE NALEGTEEQN
cpA  RLIETATVPI LAVDLDGLIN GWNTKIAELT GLPVDKAIGK HLLT-LVEDS SVEVVRKMLF LALQGQEEQN
psA  RLIETATVPI LAVDVDGTVN GWNIKIAELT GLPVGEAIGK HLLT-LVEDS STDIVKKMLN LALQGEEEKN
stA  RLIETASVPI FAVDVDGQVN GWNTKVAELT GLPVDEAIGK HLLT-LVEDS SVDTVNKMLE LALQGQEERN
asA  RLMETATVPI LAVDGNGLVN GWNQKAAELT GLRVDDAIGR HILT-LVEDS SVPVVQRMLY LALQGKEEKE
osA  RLMETATVPI LAVDSNGLVN GWNQKVAELT GLRVDEAIGR HILT-VVEES SVPVVQRMLY LALQGKEEKE
zmA  RLMETATVPI LAVDGNGLVN GWNQKVAELS GLRVDEAIGR HILT-LVEDS SVSLVQRMLY LALQGREEKE
atB  RLIETATVPI FAVDAGGCIN GWNAKIAELT GLSVEEAMGK SLVSDLIYKE NEATVNKLLS RALRGDEEKN
atD  RLIETATVPI FAVDIDGCIN GWNAKIAELT GLSVEDAMGK SLVRELIYKE YKETVDRLLS CALKGDEGKN
atE  RVIETATAPI FGVDSSGCIN GWNKKTAEMT GLLASEAMGK SLADEIVQEE SRAALESLLC KALQGEEEKS
stB  RLIETATAPI FAVDVEGRIN GWNAKVAELT GVSVEEAMGK SLVHDLVYKE SQETAEKLLY NALRGEEDKN
osB  RLIETATVPI FAVDTDGCIN GWNAKVAELT GLSVEEAMGK SLVNDLIFKE SEETVNKLLS RALRGDEDKN
atC  RLIDTAAVPI FAVDASGVIN GWNSKAAEVT GLAVEQAIGK P-VSDLVEDD SVETVKNMLA LALEGSEERG
ANG  R---TA--PI --VD--G--N GWN-K-AE-- G-----A-G- ---------- --------L- -AL-G-E---
CON  R---TA--PI --VD--G--N -WN-K-A--- G-----A--- ---------- --------L- -AL-G-E---

                 760        770        780        790        800        810        820
                   *          *          *          *          *          *          *
sm   VELKLKTFGG QKDKEAVIL- --VVNACASR DVSDNVVGVC FVGQDVTGQK VVMDKFTRIQ GDYKAIVQNP
cp   VEIKLKTFGT QTTERAVIL- --IVNACCSR DASDFVVGVF FVGQDVTEQR MFMDRFTRIQ GGEKTTVQDP
ac   IEIHLRTYDQ HKQKGVVIL- --IVNTCCSR DVSNNVVGVC FVGQDVTGQK LVLDRFIRIQ GDYKAIVQSL
atA  VQFEIKTHLS RADAGPISL- --VVNACASR DLHENVVGVC FVAHDLTGQK TVMDKFTRIE GDYKAIIQNP
cpA  VQFEIKTHGS HIEVGSISL- --VVNACASR DLRENVVGVF FVAQDITGQK MVMDKFTRLE GDYKAIVQNP
psA  VQFEIKTHGD QVESGPISL- --IVNACASK DLRENVVGVC FVAQDITAQK TVMDKFTRIE GDYKAIVQNP
stA  VEFEIKTHGP SRDSSPISL- --IVNACASK DVRDSVVGVC FIAQDITGQK SIMDKFTRIE GDYRAIIQNP
asA  VRFEVKTHGP KRDDGPVIL- --VVNACASR DLHDHVVGVC FVAQDMTVHK LVMDKFTRVE GDYKAIIHNP
osA  VKFEVKTHGS KRDDGPVIL- --VVNACASR DLHDHVVGVC FVAQDMTVHK LVMDKFTRVE GDYKAIIHNP
zmA  VRFELKTHGS KRDDGPVIL- --VVNACASR DLHDHVVGVC FVAQDMTVHK LVMDKFTRVE GDYKAIIHNP
atB  VEVKLKTFSP ELQGKAVFV- --VVNACSSK DYLNNIVGVC FVGQDVTSQK IVMDKFINIQ GDYKAIVHSP
atD  VEVKLKTFGS ELQGKAMFV- --VVNACSSK DYLNNIVGVC FVGQDVTGHK IVMDKFINIQ GDYKAIIHSP
atE  VMLKLRKFGQ NNHPDYSSDV CVLVNSCTSR DYTENIIGVC FVGQDITSEK AITDRFIRLQ GDYKTIVQSL
stB  VEIKLRTFGA EQLEKAVFV- --VVNACA-R DYTNNIVGVC FVGQDVTGEK VVMDKFINIQ GDYKAIVHSP
osB  VEIKLKTFGP EQSKGPIFV- --IVNACSTR DYTKNIVGVC FVGQDVTGQK VVMDKFINIQ GDYKAIVHNP
atC  AEIRIRAFGP KRKSSPVEL- --VVNTCCSR DMTNNVLGVC FIGQDVTGQK TLTENYSRVK GDYARIMWSP
ANG  ---------- ---------- ---VN-C--- D------GV- F---D-T--K ---------- GDY--I----
CON  ---------- ---------- ---VN-C--- D------GV- F---D-T--- ---------- G---------
```

```
830        840        850        860        870        880        890
             *          *          *          *          *          *          *
sm   NPLIPPIFGA DEFGYCSEWN PAMEKLSGWR REEVLGKMLV GEIFGIQMMY CRLKGQDAVT KFMIVLNSAA
cp   HPLMRPSFDG DEFGRTFKRN SALGGL---- ---------- ---------- ---------- ----------
ac   NPLIPPIFGA DEYGFCSEWN AAMEKLSNWR REEVLGKMLV GEIFGLQMVC CRLQGQDVVT KLMIVLNDAV
atA  NPLIPPIFGT DEFGWCTEWN PAMSKLTGLK REEVIDKMLL GEVFGTQKSC CRLKNQEAFV NLGIVLNNAV
cpA  NPLIPPIFGS DEFGWCSEWN PAMAKLTGWS REEVIDKMLL GEVFGVHKSC CRLKNQEAFV NLGIVLNNAM
psA  NQLIPPIFGT DEFGWCCEWN AAMIKLTGWK REEVMDKMLL GEVFGTQMSC CRLKNQEAFV NFGIVLNKAM
stA  HPLIPPIFGT DQFGWCSEWN SAMTMLTGWR RDDVMDKMLL GEVFGTQAAC CRLKNQEAFV NFGVILNNAI
asA  NPLIPPIFGA DEFGWCSEWN AAMTKLTGWN RDEVLDKMLL GEVFDSSNAS CPLKNRDAFV SLCVLINSAL
osA  SPLIPPIFGA DEFGWCSEWN AAMTKLTGWH RDEVINKMLL GEVFDSTNAS CLVKNKDAFV SLCILINSAL
zmA  NPLIPPIFGA DQFGWCSEWN AAMTKLTGWH RDEVVDKMLL GEVFNSSNAS CLLKSKDAFV RLCIVINSAL
atB  NPLIPPIFAA DENTCCLEWN MAMEKLTGWS RSEVIGKMIV GEVFG---SC CMLKGPDALT KFMIVLHNAI
atD  NPLIPPIFAA DENTCCLEWN TAMEKLTGWP RSEVIGKLLV REVFG---SY CRLKGPDALT KFMIVLHNAI
atE  NPLIPPIFAS DENACCSEWN AAMEKLTGWS KHEVIGKMLP GEVFG---VF CKVKCQDSLT KFLISLYQGI
stB  NPLIPPIFAS DENTCCSEWN TAMEKLTGWS RGEIVGKMLV GEIFG---SC CRLKGPDAMT KFMIVLHNAI
osB  NPLIPPIFAS DENTCCSEWN TAMEKLTGWS RGEVVGKLLV GEVFG---NC CRLKGPDALT KFMIVLHNAI
atC  STLIPPIFIT NENGVCSEWN NAMQKLSGIK REEVVNKILL GEVFTTDDYG CCLKDHDTLT KLRIGFNAVI
ANG  --LIPPIF-- -----C-EWN -AM-KL-G-- ------K--- -E-F------ C--K------ ----------
CON  --L--P-F-- ---------N -A---L---- ------K--- -E-F------ C--------- ----------

           900        910        920        930        940        950        960
             *          *          *          *          *          *          *
sm   DGQ-DTEKFP FAFFDRQGKY VEALLTATKR ADAEGSITGV FCFPHIASAE LQQALTVQRA TEKVALSKLK
cp   ---------- ---------- ---------- ---------- ---------- ---------- ----------
ac   NGQ-ESEKFP LVFYDRNGRR VEALLIASKR TDADGRITGV FCFLHTASPE LLQALIIKRA KEKV----DK
atA  TSQ-DPEKVS FAFFTRGGKY VECLLCVSKK LDREGVVTGV FCFLQLASHE LQQALHVQRL AERTAVKRLK
cpA  CGQ-DPEKAS FGFLARNGMY VECLLCVNKI LDKDGAVTGF FCFLQLPSHE LQQALNIQRL CEQTALKRLR
psA  TGL-ETEKVP FGFFSRKGKY VECLLSVSKK IDAEGLVTGV FCFLQLASPE LQQALHIQRL SEQTALKRLK
stA  TGQ-ESEKIP FGFFARYGKY VECLLCVSKR LDKEGAVTGL FCFLQLASHE LQQALHVQRL SEQTALKRLK
asA  AGE-ETEKAP FGFFDRSGKY IECLLSANRK ENEGGLITGV FCFIHVASHE LQHALQVQQA SEQTSLKRLK
osA  AGD-ETEKAP FSFFDRNGKY IECLLSVNRK VNADGVITGV FCFIQVPSHE LQHALHVQQA SQQNALTKLK
zmA  AGE-EAEKAS FGFFDRNEKY VECLLSVNRK VNADGVVTGV FCFIHVPSDD LQHALHVQQA SEQTAQRKLK
atB  GGQ-DTDKFP FPFFDRNGKF VQALLTANKR VSLEGKVIGA FCFLQIPSPE LQQALAVQRR QDTECFTKAK
atD  GGQ-DTDKFP FPFFDRKGEF IQALLTLNKR VSIDGKIIGA FCFLQIPSPE LQQALEVQRR QESEYFSRRK
atE  AGDNVPESSL VEFFNKEGKY IEASLTANKS TNIEGKVIRC FFFLQIINKE SGLSCPELKE SAQS----LN
stB  GGQ-DTDKFP FSFFDRNGKY VQALLTRNKR VNMEGDTIGA FCFIQIASPE LQQALRVQRQ QEKKCYSQMK
osB  GGQ-DCEKFP FSFFDKNGKY VQALLTANTR SRMDGEAIGA FCFLQIASPE LQQAFEIQRH HEKKCYARMK
atC  SGQKNIEKLL FGFYHRDGSF IEALLSANKR TDIEGKVTGV LCFLQVPSPE LQYALQVQQI SEHAIACALN
ANG  ---------- --F------- ----L----- ----G----- --F------- ---------- ----------
CON  ---------- --F------- ----L----- ----G----- --F------- ---------- ----------

           970        980        990       1000       1010       1020 INTRON 1030
             *          *          *          *          *          *  |       *
sm   ELAYIRQEIK NPLYGIMFTR TLMETTDLSK DQKQYFETGA VCEKQIRKIL DDMDLESIED G--YLELDTT
cp   ---------- ---------- ---------- ---------- ---------- ---------- ----------
ac   ELSYVKEELK KPLEGLAFTR TVLEGTNLTI EQRQLIKTNA WCERQLRKIL -EDDLNNIEE G--YMDLEMS
atA  ALAYIKRQIR NPLSGIMFTR KMIEGTELGP EQRRILQTSA LCQKQLSKIL DDSDLESIIE G--CLDLEMK
cpA  ALGYIKRQIQ NPLSGIIFSR RLLERTELGV EQKELLRTSG LCQKQISKVL DESDIDKIID G--FIDLEMD
psA  VLTYMKRQIR NPLAGIVFSS KMLEGTDLET EQKRIVNTSS QCQRQLSKIL DDSDLDGIID G--YLDLEMA
stA  VLAYIRRQIR NPLSGIIFSR KMLEGTSLGE EQKNILHTSA QCQRQLDKIL DDTDLDSIIE G--YLDLEML
asA  AFSYMRHAIN NPLSGMLYSR KALKNTDLNE EQMKQIHVGD NCHHQINKIL ADLDQDSITE KSSCLDLEMA
osA  AYSYMRHAIN NPLSGMLYSR KALKNTGLNE EQMKEVNVAD SCHRQLNKIL SDLDQDSVMN KSSCLDLEMV
zmA  AFSYMRHAIN KPLSGMLYSR ETLKSTGLNE EQMRQVRVGD NCHRQLNKIL ADLDQDNITD KSSCLDLDMA
atB  ELAYICQVIK NPLSGMRFAN SLLEATDLNE DQKQLLETSV SCEKQISRIV GDMDLESIED G--SFVLKRE
atD  ELAYIFQVIK NPLSGLRFTN SLLEDMDLNE DQKQLLETSV SCEKQISKIV GDMDVKSIDD G--SFLLERT
atE  ELTYVRQEIK NPLNGIRFAH KLLESSEISA SQRQFLETSD ACEKQITTII ESTDLKSIEE G--KLQLETE
stB  ELAYICQEIK SPLNGIRFTN SLLEATNLTE NQKQYLETSA ACERQMSKII RDIDLENIED G--SLTLEKE
osB  ELAYIYQEIK NPLNGIRFTN SLLEMTDLKD DQRQFLETST ACEKQMSKIV KDASLQSIED G--SLVLEKG
atC  KLAYLRHEVK DPEKAISFLQ DLLHSSGLSE DQKRLLRTSV LCREQLAKVI SDSDIEGIEE G--YVELDCS
ANG  ---Y------ -P-------- ---------- -Q-------- -C--Q----- ---------- ------L---
CON  ---Y------ -P-------- ---------- -Q-------- -C--Q----- ---------- ------L---
```

```
1040      1050        1060        1070       1080        1090        1100
             *           *           *          *           *           *          *
sm   EFMMGTVMDA VISQGMITSK EKNLQLIRET PKEIKAMFLY GDQVRLQQVL ADFLLNAIRF TPSSEN----
cp   ---------- ---------- ---------- ---------- ---------- ---------- ----------
ac   EFFMGSVIDA VISQGMAASR GKGVQILTEI PNDVKLMCLF GDQARLQQVL ADLLFCAINH ATTTNEDEKD
atA  EFTLNEVLTA STSQVMMKSN GKSVRITNET GEEVMSDTLY GDSIRLQQVL ADFMLMAVNF TPSGG-----
cpA  EFTLHEVLMV SISQVMLKIK GKGIQIVNET PEEAMSETLY GDSLRLQQVL ADFLLISVSY APSGG-----
psA  EFTLHEVLVT SLSQVMNRSN TKGIRIANDV AEHIARETLY GDSLRLQQVL ADFLLISINS TPNGG-----
stA  EFKLHEVLVA SISQVMMKSN GKNIMISNDM VEDLLNETLY GDSPRLQQVL ANFLLVSVNS TPSGG-----
asA  EFLLQDVVVA AVSQVLITCQ GKGIRISCNL PERFMKQSVY GDGVRLQQIL SDFLFISVKF SPVGG-----
osA  EFVLQDVFVA AVSQVLITCQ GKGIRVSCNL PERYMKQTVY GDGVRLQQIL SDFLFVSVKF SPVGG-----
zmA  EFVLQDVVVS AVSQVLIGCQ AKGIRVACNL PERSMKQKVY GDGIRLQQIV SDFLFVSVKF SPAGG-----
atB  EFFLGSVINA IVSQAMFLLR DRGLQLIRDI PEEIKSIEVF GDQIRIQQLL AEFLLSIIRY APSQE-----
atD  EFFIGNVTNA VVSQVMLVVR ERNLQLIRNI PTEVKSMAVY GDQIRLQQVL AEFLLSIVRY APMEG-----
atE  EFRLENILDT IISQVMIILR ERNSQLRVEV AEEIKTLPLN GDRVKLQLIL ADLLRNIVNH APFPNS----
stB  DFFLGSVIDA VVSQVMLLLR EKGVQLIRDI PEEIKTLTVH GDQVRIQQVL ADFLLNMVRY APSPDG----
osB  EFSLGSVMNA VVSQVMIQLR ERDLQLIRDI PDEIKEASAY GDQYRIQQVL CDFLLSMVRF APAENG----
atC  EFGLQESLEA VVKQVMELSI ERKVQISCDY PQEVSSMRLY GDNLRLQQIL SETLLSSIRF TPALRGL---
ANG  -F-------- ---Q------ ---------- ---------- GD----Q--- ---------- -P--------
CON  -F-------- ---Q------ ---------- ---------- GD----Q--- ---------- ----------

          1110        1120   INTRON 1130        1140        1150        1160        1170
             *           *       |   *           *           *           *           *
sm   WVGIKVATSR KRLGGVVHVM HLEFRITHPG VGLPEELVQE MFDRGRGM-T QEGLGLSMCR KLVKLMN-GE
cp   ---------- ---------- ---------- ---------- ---------- ---------- ----------
ac   WVTIKVSRTK TRLDDGVHLM HFESRISHSG QGISEALVEE MTNKSQKW-T PEGLAISISC TLIRLMN-GD
atA  QLTVSASLRK DQLGRSVHLA NLEIRLTHTG AGIPEFLLNQ MFGTEE-DVS EEGLSLMVSR KLVKLMN-GD
cpA  QLTISTDVTK NQLGKSVHLV HLEFRITYAG GGIPESLLNE MFGSEE-DAS EEGFSLLISR KLVKLMN-GD
psA  QVVIAASLTK EQLGKSVHLV NLELSITHGG SGVPEAALNQ MFGNNV-LES EEGISLHISR KLLKLMN-GD
stA  KLSISGKLTK DRIGESVQLA LLEFRIRHTG GGVPEELLSQ MFGSEA-DAS EEGISLLVSR KLVKLMN-GE
asA  SVEISSKLTK NSIGENLHLI DLELRIKHQG LGVPAELMAQ MFEEDNKEQS EEGLSLLVSR NLLRLMN-GD
osA  SVEISCSLTK NSIGENLHLI DLELRIKHQG KGVPADLLSQ MYEDDNKEQS DEGMSLAVSR NLLRLMN-GD
zmA  SVDISSKLTK NSIGENLHLI DFELRIKHRG AGVPAEILSQ MYEEDNKEQS EEGFSLAVSR NLLRLMN-GD
atB  WVEIHLSQLS KQMADGFAAI RTEFRMACPG EGLPPELVRD MFHSSR-WTS PEGLGLSVCR KILKLMN-GE
atD  SVELHLCPTL NQMADGFSAV RLEFRMACAG EGVPPEKVQD MFHSSR-WTS PEGLGLSVCR KILKLMN-GG
atE  WVGISISPGQ ELSRDNGSRI HLQFRMIHPG KGLPSEMLSD MFETRDGWVT PDGLGLKLSR KLLEQMN-GR
stB  WVEIQLRPSM MPISDGVTVV HIELGLYAPG -RLPPELVQD MFHSSR-WVT QEGLGLSMCR KMLKLMN-GE
osB  WVEIQVRPNI KQNSDGTDTM LFPFRFACPG EGLPPEIVQD MFSNSR-WTT QEGIGLSICR KILKLMG-GE
atC  CVSFKVIARI EAIGKRMKRV ELEFRIIHPA PGLPEDLVRE MFQPLRKGTS REGLGLHITQ KLVKLMERGT
ANG  ---------- ---------- ---------- ---P------ M--------- --G--L---- -----M----
CON  ---------- ---------- ---------- ---------- M--------- --G------- -----M----

          1180        1190        1200        1210
             *           *           *           *
sm   VEYIREAGKN YFLVSLELPL AQRDDAGSVK FQASS----- -
cp   ---------- ---------- ---------- ---------- -
ac   VKYTTDAGNK CFLVTIQFPL AHRDDATSVR ---------- -
atA  VQYLRQAGKS SFIITAELAA ANK------- ---------- -
cpA  VRYMREAGKS SFIITVELAA AHKSRTT--- ---------- -
psA  VRYLKEAGKS SFILSVELAA AHKLKG---- ---------- -
stA  VQYLREAGRS TFIISVELAV ATKSS----- ---------- -
asA  VRHLREAGVS TFIITAELAS APTAMGQ--- ---------- -
osA  VRHMREAGMS TFILSVELAS APAK------ ---------- -
zmA  IRHLREAGMS TFILTAELAA APSAVGR--- ---------- -
atB  VQYIRESERS YFLIILELPV PRKRPLSTAS GSGDMMLMMP Y
atD  VQYIREFERS YFLIVIELPV PLMMMMPSS- ---------- -
atE  VSYVREDERC FFQVDLQVKT MLGVESRGTE GSSSIK---- -
stB  IQYIRESERC YFLIILDLPM TRKGPKSVG- ---------- -
osB  VQYIRESERS FFHIVLELPQ PQQAASRGTS ---------- -
atC  LRYLRESEMS AFVILTEFPL I--------- ---------- -
ANG  ---------- -F-------- ---------- ---------- -
CON  ---------- -F-------- ---------- ----------
```

Appendix B

Alignment Of 3417 Homologous Nucleotides Used For Cladistic Analysis Of Full Length Phytochromes

```
Labels:
[Name: Physcomitrella    GENBANK ACCESSION X75025]
[Name: Ceratodon         GENBANK ACCESSION X17084]
[Name: Selaginella       GENBANK ACCESSION X61458]
[Name: Adiantum          GENBANK ACCESSION D13519]
[Name: Psilotum          GENBANK ACCESSION X74931]
[Name: Psilotum2         GENBANK ACCESSION X74930]
[Name: PHYB_Arabidopsis GENBANK ACCESSION X17342]
[Name: PHYB_Nicotiana    GENBANK ACCESSION L10114]
[Name: PHYB_Solanum      GENBANK ACCESSION S51538]
[Name: PHYD_Arabidopsis GENBANK ACCESSION X76609]
[Name: PHYE_Arabidopsis GENBANK ACCESSION X76610]
[Name: PHYB_Oryza        GENBANK ACCESSION X57563]
[Name: PHYA_Avena        GENBANK ACCESSION X03242]
[Name: PHYA_Oryza        GENBANK ACCESSION X14172]
[Name: PHYA_Zea          GENBANK ACCESSION NONE]
[Name: PHYA_Arabidopsis GENBANK ACCESSION X17341]
[Name: PHYA_Cucurbita    GENBANK ACCESSION M15265]
[Name: PHYA_Pisum        GENBANK ACCESSION M37217]
[Name: PHYA_Nicotiana    GENBANK ACCESSION X66784]
[Name: PHYA_Solanum      GENBANK ACCESSION S84872]
[Name: PHYC_Arabidopsis GENBANK ACCESSION X17343]

Physcomitrella     AGTGTGAGGGTCGCCCAAAC AACCGCAGATGCAGCACTTC AAGCTGTGTTCGAAAAGTCC
Ceratodon          AGCGTGCGGGTCGCTCAAAC TACGGCAGATGCAGCTCTTG AAGCTGTGTACGAAATGTCT
Selaginella        AGCGTCCGAGTTGCCCAGAC GACTGCCGATGCCAAGCTCC ACGCCGTGTACGAGGAGTCT
Adiantum           GGCAGGAGAATCGCACAAAC ATCAGCAAATGCTAAGCTTT ATGCCGCATATGAAGAGTCG
Psilotum           CGCCTCGTCGCCGCTCAGAC CGCAGCGGACGCTGAGCTTC ATGTCGTGTTTGAAGAATCT
Psilotum2          -------------------- -------------------- --------------------
PHYB_Arabidopsis   AGCAAAGCAATTCAACAGTA CACCGTCGACGCAAGACTCC ACGCCGTTTTCGAACAATCC
PHYB_Nicotiana     AGCAAAGCCATAGCACAGTA CACGGCTGATGCTAGGCTTC ATGCTGTGTTTGAACAATCT
PHYB_Solanum       AGCAAAGCTATAGCACAGTA CACAGCTGATGCTAGGCTTC ATGCTGTGTTTGAACAATCT
PHYD_Arabidopsis   AACAAAGCTATTCAACAGTA CACTGTCGACGCGAGACTCC ACGCCGTCTTCGAACAATCC
PHYE_Arabidopsis   AAATCCAACACTGCTCAATA CTCTGTTGATGCTGCTCTCT TTGCTGATTTCGCTCAATCC
PHYB_Oryza         TCCAAGGCCGTGGCGCAGTA CACCCTGGACGCGCNCCTCC ACGCCGTGTTCGAGCAGTCG
PHYA_Avena         GCAAGGGTGTTAGCACAGAC AACCCTTGATGCCGAGCTCA ATGCTGAATATGAAGAATCT
PHYA_Oryza         GCAAGGATATTAGCACAAAC AACTCTTGATGCTGAACTCA ATGCTGAATATGAAGAATAT
PHYA_Zea           GCAAGGATATTAGCACAAAC AACCCTTGATGCTGAACTCA ATGCAGAGTACGAAGAATCT
PHYA_Arabidopsis   GCTAGGATCATTGCGCAGAC CACTGTAGATGCGAAACTCC ATGCTGATTTTGAGGAGTCA
PHYA_Cucurbita     ACTAGAATTATTGCTCAGAC ATCTGTTGATGCGAACGTGC AAGCTGATTTTGAGGAATCT
PHYA_Pisum         GCTAGGATTATTGCTCAGAC GACTGTGGATGCAAAGCTTC ATGCAACTTTTGAGGAGTCC
PHYA_Nicotiana     GCTAGGATCATTGCACAGAC CACTATAGATGCAAAGCTTC ATGCAGATTTCGAGGAGTCG
PHYA_Solanum       GCTAGGATCATTGCACAGAC CTCTATTGATGCAAAGTTGC ATGCAGACTTTGAGGAGTCA
PHYC_Arabidopsis   TCTCGAGTTTCTTCACAAGT TCTCGTCGACGCAAAGCTAC ACGGAAACTTCGAAGAATCT

Physcomitrella     GGAGATTCTGGAGATTCATT CGATTATTCAAAATCTGTC- -----AGCAAATCAACAGCT
Ceratodon          GGGGACTCCGGGGACTCTTT TGACTACTCAAAATCTGTT- --------GGTCAATCTGCA
Selaginella        GGCGAGTCCGGGGACTCCTT CGACTACAGCAAGTCCATCA ATGCCACCAAGTCGACGGGC
Adiantum           TCCGAGTCTGGG---TCTTT CGACTACAGTCAATCTGTTA GTGCAGGGAAAGAAGGCATT
Psilotum           GATGATGCTGGTGAATCCTT CGACTACAGTCGTTGTATCC CTGTCTCTCTGTCTGCTGCT
Psilotum2          -------------------- -------------------- --------------------
PHYB_Arabidopsis   GGCGAATCAGGGAAATCATT CGACTACTCACAATCACTCA AAACGACGACGTACGGTTCC
PHYB_Nicotiana     GGTGAGTCTGGCAAGTCTTT TGATTATTCACAGTCTATTA AAACTACTACA---CAATCT
PHYB_Solanum       GGTGAGTCTGGAAAGTTTTT TGATTATTCACAGTCTGTTA AAACTACTACA---CAA---
PHYD_Arabidopsis   GGAGAGTCAGGTAAGTCGTT TGATTACTCACAGTCTCTTA AAACGGCTCCGTACGATTCC
PHYE_Arabidopsis   ATTTACACCGGCAAGTCTTT TAACTACTCCAAATCTGTTA TTTCACCTCCC---AATCAC
PHYB_Oryza         GGCGCGTCGGGCCGCAGCTT CGACTACACGCAGTCGCTGC GTGCGTCCCCCACCCCGTCC
PHYA_Avena         GGTGACTCC---------TT TGACTACTCCAAGCTGGTTG AAGCCCAGCGGGATGGTCCA
PHYA_Oryza         GGCGACTCC---------TT TGATTACTCCAAATTGGTTG AAGCACAGAGAACTACTGGA
PHYA_Zea           GGTGATTCC---------TT TGATTACTCCAAGTTGGTTG AAGCACAGCGGAGCACTCCA
PHYA_Arabidopsis   GGCAGCTCC---------TT TGATTACTCAACCTCAGTGC GTGTC------ACTGGCCCG
PHYA_Cucurbita     GGGAATTCG---------TT TGACTACTCAAGTTCAGTGC GTGTC------ACTAGTGAT
PHYA_Pisum         GGTAGTTCG---------TT TGACTACTCGAGTTCGGTGC GTGTT------TCTGGCTCG
PHYA_Nicotiana     GGTGATTCC---------TT TGACTATTCAAGCTCAGTGA GGGTT---------ACAAGT
PHYA_Solanum       GGTGATTCC---------TT TGACTATTCAAGCTCGGTGA GGGTT---------ACAAAT
PHYC_Arabidopsis   ---GAG------CGTTTATT TG--TATTCAGCTTCAATAA ACTTGAACATGCCAAGTTCT

Physcomitrella     GAGTCG---------CTTCC TTCAGGGGCA------GTGA CG---GCCTACCTTCAGCGT
Ceratodon          GAGTCG---------GTTCC TGCAGGGGCA------GTAA CA---GCCTACCTACAGCGT
Selaginella        GAGACG---------ATCCC AGCGCAGGCC------GTGA CG---GCGTACCTGCAGCGG
Adiantum           TCTTCC------CAACTG-- ----------------GTTA CT---GCCTACCTTCAGCGT
Psilotum           ACT------GAGAATGTTCC CCCCGAGACT------GTTA TT---GCTTACCTAATGCGT
Psilotum2          -------------------- -------------------- --------------------
```

```
PHYB_Arabidopsis   TCTGTACCTGAGCAACAG-- -------------------A TCACAGCTTATCTCTCTCGA
PHYB_Nicotiana     GTTGTTCCTGAACAGCAA-- -------------------A TTACTGCTTATTTGACTAAA
PHYB_Solanum       TCTGTGCCTGAAAGGCAA-- -------------------A TTACTGCTTATTTGACCAAA
PHYD_Arabidopsis   TCCGTACCAGAGCAGCAG-- -------------------A TCACAGCTTATCTCTCCCGG
PHYE_Arabidopsis   ---GTTCCTGATGAACAC-- -------------------A TCgcAGCTTACTTGTCTAAC
PHYB_Oryza         TCC------GAGCAGCAG-- -------------------A TCGCCGCCTACCTCTCCCGC
PHYA_Avena         CCTGTG------CAGCAAGG G---CGGTCGGAGAAGGTCA TA---GCCTACTTACAGCAC
PHYA_Oryza         CCT------GAGCAGCAAGC T---CGTTCTGAGAAGGTCA TA---GCTTACTTGCATCAC
PHYA_Zea           CCT------GAGCAGCAAGG G---CGATCGGGAAAGGTCA TA---GCCTACTTGCAGCAT
PHYA_Arabidopsis   GTTGTG---GAGAATCAGCC ACCAAGGTCTGACAAAGTTA CCACGACTTATCTTCATCAT
PHYA_Cucurbita     GTTAGCGGAGATCAACAGCC T---AGGTCAGACAAAGTTA CTACAGCTTATCTCCATCAT
PHYA_Pisum         GTGGATGGAGATCAACAACC G---AGGTCCAACAAAGTGA CGACGGCTTACCTCAATCAT
PHYA_Nicotiana     GTAGCTGGGGATGAGCGAAA GCCCAAGTCGGACAGAGTAA CTACCGCTTACCTCAATCAG
PHYA_Solanum       GTAGCTGAGGGTGAGCAAAG GCCAAAGTCAGACAAAGTTA CCACCGCTTACCTTCATCAG
PHYC_Arabidopsis   TCCTGT------GAGATTCC TTCTTCAGCT------GTCT CAACG---TACTTACAGAAG

Physcomitrella     ATGCAGAGGGGAGGTCTCAC TCAAAGTTTCGGGTGTATGA TAGCAGTT---GAAGGAACA
Ceratodon          ATGCAGAGGGAAGGTTTAAT CCAAAATTTTGGGTGTATGG TAGCAGTT---GAAGAGCCG
Selaginella        ATGCAGCGGGTGGTTTGGT  GCAGCCGTTTGGCTGCATGC TGGCGGTG---GAGGAGGGC
Adiantum           ATGCAACGCGGAGGTCTTGT TCAGCAATTTGGGTGCTTGA TTGCAGTG---GAGGAAGAG
Psilotum           ATGCAACGAGGTGGGTTAAT CCAGCCCTTTGGTTGCCTCT TGGCTGTG---GACGAGAAT
Psilotum2          -------------------- -------------------- --------------------
PHYB_Arabidopsis   ATCCAGCGAGGTGGTTACAT TCAGCCTTTCGGATGTATGA TCGCCGTC---GATGAATCC
PHYB_Nicotiana     ATCCAAAGAGGGGGTCATAT TCAGCCTTTTGGTTGTATGA TAGCTGTA---GATGAGGCT
PHYB_Solanum       ATTCAAAGAGGAGGTCATAT TCAGCCTTTTGGTTGTATGA TAGCTGTA---GATGAGGCT
PHYD_Arabidopsis   ATCCAACGCGGTGGCTATAC CCAGCCTTTTGGCTGCTTGA TCGCCGTC---GAAGAATCC
PHYE_Arabidopsis   ATCCAAAGAGGCGGTCTAGT TCAGCCCTTTGGTTGTTTGA TTGCTGTC---GAAGAACCT
PHYB_Oryza         ATCCAGCGCGGCGGGCACAT ACAGCCCTTCGGCTGCACGC TCGCCGTCGCCGACGACTCC
PHYA_Avena         ATTCAGAAAGGAAAGCTAAT CCAAACATTTGGTTGCCTGT TGGCCCTT---GATGAGAAG
PHYA_Oryza         ATTCAGAGAGCAAAGCTAAT CCAACCATTTGGTTGCTTGT TGGCCCTT---GATGAGAAG
PHYA_Zea           ATTCAAAGAGGAAAGCTTAT CCAACCATTTGGTTGCCTGT TGGCCCTT---GACGAGAAG
PHYA_Arabidopsis   ATACAGAAGGGAAAGCTGAT TCAGCCCTTCGGTTGTTTAC TTGCCTTG---GATGAGAAG
PHYA_Cucurbita     ATTCAGAAAGGCAAACTTAT TCAACCATTTGGTTGCTTGT TGGCCTTA---GATGACAAA
PHYA_Pisum         ATACAGAGAGGTAAGCAGAT CCAGCCTTTCGGGTGCTTGC TAGCTTTA---GATGAGAAA
PHYA_Nicotiana     ATCCAGAAGGGTAAGTTTAT CCAGCCATTTGGCTGTTTGT TAGCCCTT---GATGAGAAA
PHYA_Solanum       ATCCAAAAGGGCAAGTTTAT CCAGCCATTTGGTTGTCTGT TAGCCCTG---GATGAGAAA
PHYC_Arabidopsis   ATTCAGAGAGGGATGTTGAT TCAACCCTTTGGTTGTTTAA TCGTTGTT---GATGAGAAA

Physcomitrella     GGATTTCGTGTGATAGCGTA CAGTGAAAATGCACCCGAGA TTCTAGACCTGGTGCCACAG
Ceratodon          AATTTCTGTGTTATAGCGTA CAGTGAGAATGCGTCCGAGT TTCTAGATCTGATACCCCAG
Selaginella        AGCTTCCGTGTGATTGCCTT CAGTGACAATGCGGGGGAGA TGCTGGACCTCATGCCTCAG
Adiantum           ACCTTCAGAGTCTTGCATAT GTGCGAA---GCACCTGAGA TGCTAGATGTAGCCACTCAA
Psilotum           TCTTTTCGGGTGATTGCCTA CAGCGAAAATGCACCAGAAG TACTCGATCTTATACCACAG
Psilotum2          -------------------- -------------------- --------------------
PHYB_Arabidopsis   AGTTTCCGGATCATCGGTTA CAGTGAAAACGCCAGAGAAA TGTTAGGGATTATGCCTCAA
PHYB_Nicotiana     AGTTTTCGTGTTATTGCTTA CAGCGAAAATGCGTGCGAAA TGCTTAGTCTAACTCCACAA
PHYB_Solanum       AGTTTTCGTGTAATAGCTTA TAGTGAAAATGCCTGTGAAA TGCTTAGTTTAACTCCACAA
PHYD_Arabidopsis   ACTTTCACAATCATCGGTTA CAGTGAAAATGCGCGGGAAA TGCTAGGGCTCATGTCTCAA
PHYE_Arabidopsis   AGTTTTAGGATACTTGGTCT TAGTGACAACTCTTCTGACT TTCTTGGTTTGTTGTCTCTT
PHYB_Oryza         TCCTTCCGCCTCCTCGCCTA CTCCGAGAACACCGCCGACC TGCTCGACCTGTCGCCCCAC
PHYA_Avena         AGCTTCAATGTCATCGCGTT CAGTGAGAACGCGCCAGAAA TGCTTACAACGGTCAGCCAT
PHYA_Oryza         ACCTTCAATGTTATAGCGCT CAGCGAGAATGCACCAGAGA TGCTTACAACTGTCAGCCAT
PHYA_Zea           AGCTTCAGGGTCATTGCATT CAGTGAGAATGCACCTGAAA TGCTTACAACGGTCAGCCAT
PHYA_Arabidopsis   ACCTTCAAAGTTATTGCATA CAGCGAGAATGCATCTGAGC TGTTGACAATGGCCAGTCAT
PHYA_Cucurbita     ACATTCAAGGTTATTGCGTA TAGTGAAAATGCCCCTGAAA TGTTGACCATGGTGAGCCAT
PHYA_Pisum         ACGTGCAAGGTTGTTGCGTA TAGTGAGAATGCGCCTGAGA TGCTGACTATGGTGAGTCAT
PHYA_Nicotiana     ACATTTAAGGTCATAGCATT CAGCGAGAATGCCCCCGAAA TGCTGACCATGGTTAGCCAT
PHYA_Solanum       ACATTAAAGGTCATAGCATT CAGTGAAAATGCCCCTGAAA TGCTGACCATGGTTAGCCAT
PHYC_Arabidopsis   AACCTTAAAGTCATTGCCTT TAGTGAAAACACTCAAGAGA TGTTGGGTTTGATTCCACAT

Physcomitrella     ---GCCGTTCCAAGCGTG-- -GGG------GAGATGGAC- -----ACGCTACGAATCGGG
Ceratodon          ---GCCGTCCCAAGTATG-- -GGG------GAGATGGAC- -----GTGCTAGGAATCGGG
Selaginella        ---TCGGTTCCAAGCCTG-- -GGCAGCGGGCAGCAGGAC- -----GTGCTGACCATCGGG
Adiantum           ---GCCGTGCCAACAATG-- -GGG------CAATATAGC- -----CGTCTCTGTATAGGT
Psilotum           ---TTCGTTCCAACGGTG-- -GGA------CTGCAGGAG- -----ATACTGAGAATTGGC
Psilotum2          -------------------- -------------------- --------------------
PHYB_Arabidopsis   ---TCTGTTCCTACTCTT-- -------------GAGAAAC CTGAGATTCTAGCTATGGGA
PHYB_Nicotiana     ---TCAGTTCCAAGCCTT-- -------------GAGCGGC CTGAGATCCTCACTGTTGGA
PHYB_Solanum       ---TCTGTTCCAAGCCTT-- -------------GAGAAGT GTGAGATCCTCACTATTGGA
PHYD_Arabidopsis   ---TCTGTACCAAGCATC-- ----------GAGGACAAAT CAGAGGTTTTAACGATTGGT
PHYE_Arabidopsis   ---CCTTCCACCTCCCAT-- -TCTGGT---GAGTTTGATA AAGTCAAGGGTTTGATTGGA
PHYB_Oryzaa        CACTCCGTCCCCTCGCTCGA CTCCTCCGCGGTGCCTCCCC CCGTC------TCGCTCGGC
PHYA_Avena         ---GCGGTACCCAGTGTT-- ----------GATGATCCCC CAAGGCTG---GGGATTGGC
```

```
PHYA_Oryza          ---GCAGTGCCAAGTGTT-- ----------GATGATCCCC CAAAGCTA---CGCATTGGC
PHYA_Zea            ---GCTGTGCCGAACGTT-- ----------GATGATCCCC CGAAGCTA---GGAATTGGC
PHYA_Arabidopsis    ---GCAGTTCCTAGTGTT-- -GGC---------GAACACC CTGTTCTA---GGCATTGGG
PHYA_Cucurbita      ---GCTGTCCCAAGCATG-- -GGG---------GATTACC CTGTTCTT---GGCATTGGC
PHYA_Pisum          ---GCTGTTCCTAGCGTT-- -GGT---------GACCATC CTGCTCTT---GGCATTGGA
PHYA_Nicotiana      ---GCTGTTCCAAGTGTG-- -GGT---------GAGCTTC CAGCTCTT---GGCATTGGG
PHYA_Solanum        ---GCTGTTCCAAGTGTT-- -GGT---------GAGCATC CAGTTCTT---GGGATCGGG
PHYC_Arabidopsis    ---ACAGTACCAAGTATG-- ----------GAGCAGCGT- --GAAGCTTTGACTATAGGA

Physcomitrella      ACGGATGTGAGAACGCTGTT CACAGCTTCAAGTGTCGCCT CTCTTGAGAAGGCAGCTGAA
Ceratodon           ACGGATATAAGAACTTTATT CACACCGTCGAGTAGTGCCG CTCTTGAGAAGGCAGCTGCA
Selaginella         ACCGACGCCCGCACCCTCTT CACGGCAGCAGCCAGCGCGC TG---GAGAAGGCGGCAGGA
Adiantum            GCTGATGTAAGAACTCTGTT ATCTCCGGCGAGTGCATCAG CATTGGATCGAGTCATCGGT
Psilotum            ACCGATGCTCGTGACCTCTT CACTGCAGCTGGTGCAGCTC GGCTTGGAAGAGTTGTTGGG
Psilotum2           -------------------- -------------------- --------------------
PHYB_Arabidopsis    ACTGATGTGAGATCTTTGTT CACTTCTTCGAGCTCGATTC TACTCGAGCGTGCTTTCGTT
PHYB_Nicotiana      ACTGATGTTAGGACCCTTTT TACTCCTTCTAGCTCTGTTT TGCTAGAAAGAGCATTTGGG
PHYB_Solanum        ACTGATGTTAGGACCCTTTT TACCCCTTCTAGCTCTGTTT TGCTGGAAAGAGCATTTGGG
PHYD_Arabidopsis    ACGGATTTGCGATCTCTCTT CAAGTCATCGAGCTACCTTC TCCTCGAGCGCGCGTTCGTG
PHYE_Arabidopsis    ATCGATGCAAGGACGCTTTT TACTCCTTCCTCTGGAGCTT CTTTGTCTAAAGCTGCTTCC
PHYB_Oryzaa         GCAGACGCGCGCCTCCTTTT CGCTCCCTCGTCCGCCGTCC TCCTCGAGCGCGCCTTCGCC
PHYA_Avena          ACCAATGTACGGTCTCTTTT CAGTGACCAAGGTGCCACAG CACTGCATAAGGCACTAGGA
PHYA_Oryza          ACCAATGTACGGTCTCTTTT CACTGACCCAGGTACCACAG CACTGCAGAAGGCACTGGGA
PHYA_Zea            ACCAATGTGCGATCCCTTTT CACTGACCCTGGTGCTACAG CACTGCAGAAGGCACTTGGA
PHYA_Arabidopsis    ACAGATATAAGGAGTCTTTT CACTGCTCCTAGTGCGTCTG CATTGCAGAAAGCCCTTGGA
PHYA_Cucurbita      ACAGATGTAAGGACTATTTT CACCGCACCTAGTGCTTCTG CACTGTTGAAGGCCTTGGGC
PHYA_Pisum          ACTGACATAAGGACTGTTTT CACTGCGCCGAGTGCTTCTG CCTTGCAGAAGGCGCTAGGG
PHYA_Nicotiana      ACTGATATCAGAACGATCTT CACTGGTCCTAGCGCAGCTG CATTGCAGAAGGCTTTGGGG
PHYA_Solanum        ATTGATATCAGAACGATCTT CACTGGTCCTAGTGGCGCAG CATTGCAGAAAGCCTTGGGG
PHYC_Arabidopsis    ACTGATGTGAAATCATTGTT TCTGTCTCCAGGTTGTTCTG CTTTGGAGAAAGCTGTTGAC

Physcomitrella      GCACAAGAAATGAGCCTACT CAATCCAATCACCGTTAACT GCAGAAGATCTGGGAAGCAA
Ceratodon           ACTCAGGATATAAGCCTTCT TAACCCAATCACTGTTCATT GCAGACGCTCAGGGAAACCG
Selaginella         GCCGTCGACTTGAGCATGCT CAACCCGATATGGGTACAGT CCAAGACGTCGGCCAAGCCC
Adiantum            GTTGTCGATGTGAGCATGTT CAATCCCATAACCGTGCAAA GCCGAAGCTCTGGAAAGCCC
Psilotum            GCTGTCGATGTTTCATTGTT TAATCCCATAATTGTTCAGA CTAAAAGCACGGGCAAGCCT
Psilotum2           -------------------- -------------------- --------------------
PHYB_Arabidopsis    GCTCGAGAGATTACCTTGTT AAATCCGGTTTGGATCCATT CCAAGAATACTGGTAAACCG
PHYB_Nicotiana      GCGCGCGAGATCACTTTGCT GAATCCTATTTGGATTCATT CCAAGAATTCTGGCAAGCCA
PHYB_Solanum        GCACGTGAGATCACTTTACT CAACCCAATTTGGATACATT CCAAGAATTCTGGAAAGCCC
PHYD_Arabidopsis    GCTCGAGAGATCACGCTTCT GAATCCTATTTGGATTCACT CTAACAACACTGGTAAACCT
PHYE_Arabidopsis    TTTACTGAGATTTCCTTGTT GAATCCTGTTTTGGTCCATT CTAGGACGACCCAGAAGCCT
PHYB_Oryzaa         GCGCGCGAGATCTCGCTGCT CAACCCGCTCTGGATCCACT CCAGGGTCTCCTCTAACCCT
PHYA_Avena          TTTGCTGATGTGTCTTTGCT GAATCCTATCCTGGTTCAGT GCAAGACATCAGGCAAGCCT
PHYA_Oryza          TTTGCTGATGTTTCCTTGCT GAACCCTATCCTAGTTCAAT GCAAGACCTCAGGCAAGCCT
PHYA_Zea            TTTGCTGATGTTTCTTTGCT GAATCCTATCCTGGTTCAGT GCAAGACCTCAGGCAAGCCA
PHYA_Arabidopsis    TTTGGAGATGTCTCTCTTTT GAATCCCATTCTTGTGCACT GCAGGACTTCTGCAAAGCCC
PHYA_Cucurbita      TTTGGAGAGGTTACACTTCT TAATCCTATCCTGGTGCATT GCAAGACTTCTGGAAAACCC
PHYA_Pisum          TTTGCGGAGGTTTCGCTTCT TAACCCGATTCTTGTTCATT GCAAGACTTCTGGGAAGCCG
PHYA_Nicotiana      TTCGGAGAGGTTTCTCTGTT AAATCCGGTCCTCGTTCACT GCAAAACTTCTGGGAAGCCA
PHYA_Solanum        TTTGGGGAAGTTTCTCTGTT AAATCCTGTCCTTGTTCACT GCAAAAATTCTGGGAAGCCA
PHYC_Arabidopsis    TTTGGTGAGATTAGTATTTT GAATCCTATCACGCTTCATT GTAGGTCTTCAAGTAAGCCT

Physcomitrella      TTATATGCCATTGCTCATCG CATTGACATAGGTATAGTGA TCGACTTTGAGGCAGTAAAA
Ceratodon           TTATATGCCATTGCCCATCG CATAGACATTGGTATAGTCA TTGACTTTGAGGCGGTGAAA
Selaginella         TTCTACGCCATCGTGCACCG CATCGACGTGGGCCTGGTCA TGGACCTGGAGCCCGTCAAG
Adiantum            TTCTATGCAATTTTGCACAG AAATGATGTGGGGTTGGTAA TTGACCTTGAGCCTATCAGG
Psilotum            TTCGATGCCATTGTCCATCG CATAGACGTGGGATTGGTGA TTGACTTCGAGCCTTTGAGG
Psilotum2           -------------------- -------------------- --------------------
PHYB_Arabidopsis    TTTTACGCCATTCTTCATAG GATTGATGTTGGTGTTGTTA TTGATTTAGAGCCAGCTAGA
PHYB_Nicotiana      TTTTACGCAATTTTGCATAG GGTTGATGTCGGGATTGTAA TTGATTTGGAGCCTGCTAGA
PHYB_Solanum        TTTTATGCAATTTTGCACAG GGTTGATGTTGGTATTGTCA TTGATTTGGAGCCTGCTAGA
PHYD_Arabidopsis    TTCTACGCGATTCTCCACAG GGTTGATGTTGGAATTTTGA TCGATTTAGAGCCGGCTCGA
PHYE_Arabidopsis    TTTTATGCTATTCTTCACAG GATTGATGCAGGGATTGTCA TGGATTTGGAGCCTGCTAAA
PHYB_Oryza          TTCTACGCCATCCTCCACCG CANCGATGTCGGCGTCGTCA TCGACCTCGAGCCCGCCCGC
PHYA_Avena          TTCTATGCCATTGTTCATCG AGCAACTGGTTGTTTGGTGG TAGACTTTGAGCCTGTAAAG
PHYA_Oryza          TTCTATGCCATTGTTCATCG GGCAACTGGTTGTTTGGTGG TAGACTTTGAGCCTGTGAAA
PHYA_Zea            TTCTATGCCATTGTTCATAG GGCAACTGGTTGTCTGGTGG TAGATTTTGAGCCTGTGAAG
PHYA_Arabidopsis    TTTTATGCGATTATCCACAG GGTTACAGGGAGCATCATCA TCGACTTTGAACCCGTGAAG
PHYA_Cucurbita      TTCTATGCAATTGTTCATCG TGTTACTGGAAGCTTAATCA TTGACTTTGAGCCTGTGAAG
PHYA_Pisum          TTTTACGCGATCATTCATCG TGTTACTGGTAGTTTGATCA TTGACTTTGAGCCGGTGAAG
PHYA_Nicotiana      TATTATGCAATTGTTCATAG GGTTACTGGTAGCTTAATCA TTGATTTCGAGCCTGTGAAG
PHYA_Solanum        TTTTATGCAATTGTTCATAG GGTTACAGGTAGCTTAATCA TTGATTTTGAGCCTGTGAAG
```

```
PHYA_Arabidopsis    TTTTATGCGATTCTGCATCG GATTGAGGAAGGTCTTGTTA TAGATTTGGAGCCTGTGAGT

Physcomitrella      ACAGACGATCATTTAGTT-- -TCAGCTGCTGGTGCACTGC AATCTCATAAGCTTGCAGCC
Ceratodon           ATGATTGATGTTCCAGTTTC AGCTGCTGCCGGTGCACTGC AATCTCACAAACTTGCGGCC
Selaginella         GCCAGCGACACGAGGGTGGG ATCCGCTGCCGGGGCTCTGC AGTCCCACAAGCTGGCGGCC
Adiantum            CCTGATGAT---GCATCTAT CACT---GGCGGTGCCTTGC AGTCACATAAGCTGGCAGCG
Psilotum            TCTACTGATTTAGCTGCCTT CTCGGCTGCGGGAATATTGC AATCTCACAAACTTGCCGCC
Psilotum2           -------------------- -------------------- --------------------
PHYB_Arabidopsis    ACTGAAGAT---CCTGCGCT TTCTATTGCTGGTGCTGTTC AATCGCAGAAACTCGCGGTT
PHYB_Nicotiana      ACAGAGGAC---CCTGCTTT ATCCATTGCTGGCGCAGTGC AGTCACAAAAACTTGCTGTG
PHYB_Solanum        ACTGAGGAC---CCTGCTTT ATCTATTGCTGGAGCAGTGC AGTCACAGAAACTTCGCAGT
PHYD_Arabidopsis    ACCGAAGAT---CCGGCACT TTCAATCGCCGGAGCAGTCC AATCGCAGAAACTTGCGGTA
PHYE_Arabidopsis    TCAGGTGAT---CCGGCTTT GACCCTTGCAGGCGCAGTTC AGTCTCAGAAGCTAGCCGTT
PHYB_Oryza          ACCGAGGAT---CCTGCACT CTCCATCGCTGGCGCAGTCC AGTCTCAGAAGCTCGTGGTC
PHYA_Avena          CCTACAGAATTTCCTGCC-- -ACTGCTGCTGGGGCTTTGC AGTCCTACAAGCTTGCTGCC
PHYA_Oryza          CCTACAGAATTTCCAGCA-- -ACTGCCGCTGGGGCTTTGC AATCTTACAAACTTGCTGCC
PHYA_Zea            CCTACAGAATTTCCTGCC-- -ACTGCTGCTGGGGCTTTGC AGTCCTACAAGCTTGCTGCC
PHYA_Arabidopsis    CCTTATGAA---GTCCCCAT GACAGCTGCTGGTGCCTTAC AATCATACAAGCTCGCTGCC
PHYA_Cucurbita      CCTTATGAA---GGTCCAGT GACTGCAGCTGGAGCTCTAC AATCATATAAACTTGCCGCC
PHYA_Pisum          CCTTATGAA---GTTCCCAT GACTGCTGCGGGTGCCTTGC AATCTTACAAACTTGCTGCT
PHYA_Nicotiana      CCCTATGAA---GTACCCAT GACTGCTGCAGGGGCCCTGC AGTCGTATAAACTTGCAGCC
PHYA_Solanum        CCCTATGAA---GTTCCCAT GACTGCTGCAGGGGCCCTGC AGTCATATAAACTTGCAGCC
PHYA_Arabidopsis    CCTGATGAG---GTGCCTGT GACTGCTGCCGGGGCTTTAA GATCGTATAAGCTTGCGGCG

Physcomitrella      AAGGCAATC---ACACGGCT TCAAGCATTACCTGGAGGTA ACATAGGATTGCTCTGTGAC
Ceratodon           CGGGCTATT---ACACGACT TCAAGCATTACCTGGAGGCG ACATAGAGTTGCTTTGTGAT
Selaginella         AAGGCCATA---TCGCGCCT CCAGTCGCTGCCGGGCGGCG ACATCGGCCTGTTGTGCGAC
Adiantum            AAGGCCATC---GCTCGGTT GCAGTCTTTGCCAGGAGGTG ATATTGGCCTCTTATGTGAC
Psilotum            AAAGCAATT---TCGAGGTT GCAATCTTTACCGGTTGGGG ATATCGGTCTTCTCTGCGAT
Psilotum2           -------------------- -------------------- --------------------
PHYB_Arabidopsis    CGTGCGATT---TCTCAGTT ACAGGCTCTTCCTGGTGGAG ATATTAAGCTTTTGTGTGAC
PHYB_Nicotiana      AGGGCTATT---TCTCATTT GCAATCACTTCCTGGTGGGG ATGTTAAGCTTTTGTGTGAT
PHYB_Solanum        GAGGGCCTATTTCTCATTTG CAATCACTTCCTGGTGGGGA CATTAAAGCTTTTGTGTGAT
PHYD_Arabidopsis    CGTGCGATT---TCTCATTT ACAATCGTTGCCTAGCGGCG ACATTAAGCTTCTATGTGAC
PHYE_Arabidopsis    AGGGCCATT---TCTAGGCT GCAGTCACTTCCCGGAGGAG ATATTGGTGCCTTGTGTGAT
PHYB_Oryza          CGTGCCATC---TCCCGCCT CCAGGCCCTTCCCGGCGGTG ACGTCAAGCTCCTTTGCGAC
PHYA_Avena          AAGGCAATA---TCCAAGAT CCAGTCATTGCCAGGTGGAA GCATGGAGGTGCTATGCAAT
PHYA_Oryza          AAGGCAATC---TCTAAGAT CCAGTCACTGCCAGGTGGAA GCATGGAGGTGCTATGCAAT
PHYA_Zea            AAGGCAATC---TCCAAGAT CCAGTCACTACCAGGTGGAA GCATGGAGGCCTTATGCAAT
PHYA_Arabidopsis    AAAGCAATC---ACTAGGCT GCAATCTTTACCCAGCGGGA GTATGGAAAGGCTTTGTGAT
PHYA_Cucurbita      AAAGCGATT---ACTAGATT GCAGTCTTTGCCTAGTGGAA GCATGGCTAGGCTTTGTGAC
PHYA_Pisum          AAAGCAATT---ACAAGATT GCAATCTTTGGCTAGTGGCA GCATGGAAAGGCTTTGTGAT
PHYA_Nicotiana      AAAGCCATT---ACTCGCTT GCAGGCCTTGCCCAGCGGCA GTATGGAAAGACTTTGTGAC
PHYA_Solanum        AAAGCCATT---ACTCGCTT GCAGTCCTTGCCCAGTGGCA GTATGGAAAGACTTTGTGAC
PHYA_Arabidopsis    AAATCGATT---TCGAGGTT GCAGGCATTGCCTAGTGGGA ATATGTTGTTGTTGTGTGAT

Physcomitrella      ACTGTTGTCGAAGAGGTTCG GGAGCTCACTGGTTATGACA GGGTAATGGCTTATAGATTT
Ceratodon           ACTATTGTTGAGGAGGTGCG GGAACTTACTGGGTATGACA GGGTGATGGCTTTTAAATTT
Selaginella         ACGGTGGTGGAGGAGGTGCG GGACGTGACCGGTTACGACC TGGTCATGGCATACAAGTTT
Adiantum            TCAGTAGTGGAGGAGGTACA TGAGCTTACGGGTTTCGACA GAGTGATGGCATATAAGTTT
Psilotum            ACCGTTGTTGAGGAAGTTAG GCAGCTCACCGGTTATGATA GGGTGATGGCATATAAATTT
Psilotum2           -------------------- -------------------- --------------------
PHYB_Arabidopsis    ACTGTCGTGGAAAGTGTGAG GGACTTGACTGGTTATGATC GTGTTATGGTTTATAAGTTT
PHYB_Nicotiana      ACTGTGGTTGAGAGTGTGAG GGAGTTAACTGGGTATGATC GGGTTATGGTATATAAATTT
PHYB_Solanum        ACTGTTGTTGAGAGTGTCAG GGAGTTAACCGGGTATGACC GGGTTATGGTATATAAATTT
PHYD_Arabidopsis    ACTGTTGTGGAAAGCGTTAG AGATCTTACTGGCTACGACC GCGTTATGGTGTACAAGTTT
PHYE_Arabidopsis    ACTGTTGTGGAAGATGTTCA GAGACTTACCGGTTATGACC GTGTTATGGTCTATCAGTTT
PHYB_Oryza          ACCGTTGTTGAGCATGTTAG AGAGCTCACAGGTTATGACC GCGTTATGGTGTACAGGTTC
PHYA_Avena          ACTGTGGTGAAGGAAGTCTT TGACCTTACCGGGTATGACA GGGTTATGGCTTACAAGTTT
PHYA_Oryza          ACGGTGGTCAAGGAACTCTT TGACCTCACAGGATATGATA GAGTTATGGCTTATAAGTTC
PHYA_Zea            ACCGTGGTTAAGGAAGTCTT TGACCTGACAGGTTATGACA GGGTTATGGCTTACAAGTTC
PHYA_Arabidopsis    ACAATGGTTCAAGAGGTTTT TGAACTCACGGGGTATGACA GGGTGATGGCTTATAAGTTT
PHYA_Cucurbita      ACAATGGTTCAAGAAGTTTT TGAACTAACAGGTTATGATC GAGTTATGGCTTATAAATTC
PHYA_Pisum          ACCATGGTTCAAGAAGTTTT TGAACTAACGGGTTATGACA GGGTGATGGCTTATAAATTT
PHYA_Nicotiana      ACTATGGTTCAGGAGGTTTT CGAACTCACAGGCTATGACA GGGTGATGACGTATAAGTTT
PHYA_Solanum        ACAATGGTTCAGGAGGTTTT TGAACTCACAGGTTATGACA GGGTGATGGGATATAAGTTT
PHYA_Arabidopsis    GCTTTGGTTAAGGAAGTTAG TGAATTAACTGGTTATGATA GGGTGATGGTGTATAAGTTC

Physcomitrella      CATGAGGATGAGCATGGAGA GGTTGTGGCTGAGATACGTC GCGCGGATCTTGAGCCTTAC
Ceratodon           CATGAAGATGAGCATGGCGA AGTTGTGGCAGAAATACGTC GCATGGATCTTGAGCCCTAT
Selaginella         CACGAAGACGAGCATGGCGA GGTGGTGGCGGAGATTCGGC GCTCCGACCTGGAACCGTAC
Adiantum            CATGAAGATGAGCATGGAGA GGTGGTTGCTGAGATAAGGC GTACGGATCTTGAGCCTTAT
Psilotum            CATGAAGATGAACACGGTGA GGTGGTTGCGGAGATACGTC GTTCAGATCTGGAGCCTTTT
```

```
Psilotum2          -------------------- -------------------- --------------------
PHYB_Arabidopsis   CATGAAGATGAGCATGGAGA AGTTGTAGCTGAGAGTAAAC GAGACGATTTAGAGCCTTAT
PHYB_Nicotiana     CATGAGGATGAGCACGGGGA GGTAGTGGCTGAGAGCAAAA TACCAGATTTAGAGCCCTAT
PHYB_Solanum       CATGAGGATGAACATGGAGA GGTAGTGGCTGAGAGTAAAA GATCAGATTTAGAGCCCTAT
PHYD_Arabidopsis   CATGAAGATGAACATGGTGA AGTCGTAGCCGAGAGTAAAC GGAACGATTTAGAGCCTTAC
PHYE_Arabidopsis   CATGAAGATGATCATGGTGA AGTTGTTTCTGAGATTAGAA GGTCTGATTTGGAGCCTTAT
PHYB_Oryza         CATGAGGATGAGCATGGAGA AGTCGTTGCCGAGAGCCGGC GCAGTAACCTTGAGCCCTAC
PHYA_Avena         CATGAAGATGACCATGGTGA GGTATTCTCCGAAATCACAA AGCCTGGTCTTGAGCCTTAT
PHYA_Oryza         CATGAAGATGACCATGGTGA AGTCTTTGCTGAGATCACAA AGCCTGGTCTTGAACCTTAT
PHYA_Zea           CATGAAGATGAGCATGGGGA GGTCTTCGCTGAGATCACCA AACCTGGTATTGAGCCCTAT
PHYA_Arabidopsis   CATGAAGATGATCACGGTGA GGTTGTCTCCGAGGTTACAA AACCTGGGCTGGAGCCTTAT
PHYA_Cucurbita     CATGATGATGATCATGGGGA AGTGATCTCTGAAGTCGCAA AGCCCGGCCTTCAGCCATAT
PHYA_Pisum         CACGAGGATGATCACGGGGA GGTGATTGCTGAGATAGCAA AGCCAGGCCTAGAGCCATAT
PHYA_Nicotiana     CACGATGATGATCATGGAGA GGTGGTGGCCGAGATCACGA AGCCTGGCCTTGATCCTTAC
PHYA_Solanum       CACGATGATGATCATGGAGA GGTGGTGTCTGAGATCACAA AGCCTGGCCTCGAGCCTTAC
PHYA_Arabidopsis   CATGAGGATGGGCATGGGGA AGTGATTGCTGAATGCTGCC GGGAAGATATGGAACCTTAT

Physcomitrella     TTAGGTCTTCATTATCCGGG CACTGATATTCCCCAGGCGT CCCGGTTTCTGTTTATGAAG
Ceratodon          ATGGGTCTCCATTATCCGGC CACTGACATTCCCCAGGCGT CCCGTTTTCTGTTAATGAAG
Selaginella        CTCGGGCTGCACTACCCAGC CACGGACATACCCCAGGCTT CTCGATTCCTCTTCATGAAG
Adiantum           ATTGGGCTGCATTACCCAGC TACTGATATCCCACAAGCTG CTCGTTTTCTCTTCATGAAA
Psilotum           GTCGGTATACATTATCCAGC CACGGACATTCCCCAGGCTT GTCGTTTTCTGTTCTTGAAG
Psilotum2          -------------------- -------------------- --------------------
PHYB_Arabidopsis   ATTGGACTGCATTATCCTGC TACTGATATTCCTCAAGCGT CAAGGTTCTTGTTTAAGCAG
PHYB_Nicotiana     ATTGGTTTGCATTATCCTGC TACCGACATTCCTCAAGCTT CGCGGTTTTTGTTTAAGCAG
PHYB_Solanum       ATCGGTTTGCATTATCCTGC TACTGATATTCCTCAAGCTT CACGGTTTTTGTTTAAGCAG
PHYD_Arabidopsis   ATTGGTCTGCATTATCCCGC TACTGATATTCCTCAGGCAT CTCGGTTCTTGTTCAAGCAA
PHYE_Arabidopsis   TTGGGTTTACATTATCCTGC AACAGATATTCCTCAGGCTG CTCGGTTCTTGTTCAAACAG
PHYB_Oryza         ATCGGGTTGCATTATCCTGC TACAGATATCCCACAGGCAT CACGCTTCCTGTTCCGGCAG
PHYA_Avena         CTAGGCCTGCACTATCCAGC CACTGACATCCCTCAAGCAG CCAGGCTTCTTTTCATGAAG
PHYA_Oryza         CTTGGCCTGCATTATCCAGC TACTGATATCCCTCAGGCAG CCAGGTTTCTTTTCATGAAG
PHYA_Zea           ATAGGCCTGCACTATCCAGC CACTGATATCCCTCAAGCTG CCAGGTTTCTCTTCATGAAG
PHYA_Arabidopsis   CTTGGGCTGCATTATCCTGC CACCGACATCCCTCAAGCAG CCCGTTTTCTGTTTATGAAG
PHYA_Cucurbita     CTTGGTTTGCATTATCCAGC AACTGACATTCCTCAAGCTG CACGTTTTTTGTTCATGAAA
PHYA_Pisum         CTAGGTCTGCACTATCCGGC GACAGATATTCCCCAGGCTG CGCGGTTTCTATTTATGAAG
PHYA_Nicotiana     CTTGGTTTACACTATCCTGC TACGGATATCCCACAAGCTG CACGCTTTTTGTTCATGAAG
PHYA_Solanum       CTTGGTTTACATTATCCTGC TACGGATATTCCACAGGCTG CACGGTTTTTGTTTATGAAG
PHYA_Arabidopsis   CTTGGGTTGCATTACTCCGC TACTGATATACCGCAAGCTT CGAGATTTCTGTTTATGAGA

Physcomitrella     AACAAGGTGCGGATAATTGC TGATTGTTCCGCACCTCCAG TGAAAGTTATACAAGATCCA
Ceratodon          AACAGGGTGCGGTTGATAGC TGATTGCTATGCGTCTCCAG TGAAACTCATACAAGATCCA
Selaginella        AACCGGGTACGCATGATCTG TGACTGCTCCGCCCCGCCCG TGAAGATCACCCAGGACAAG
Adiantum           AACAGAGTGAGGATGATATG TGATTGCAGGTTGCCGCCAG TTAAGCTTATTCAGGACAAG
Psilotum           AACAGAGTCACGATGATTTG TGACTGTTATGCTCCCCCTA TTAGGATAATCCAAGACAGG
Psilotum2          -------------------- -------------------- --------------------
PHYB_Arabidopsis   AACCGTGTCCGAATGATAGT AGATTGCAATGCCACACCTG TTCTTGTGGTCCAGGACGAT
PHYB_Nicotiana     AACAGGGTAAGAATGATTGT GGACTGCCATGCCACCCCTG TGCGGGTTGTTCAGGATGAA
PHYB_Solanum       AACAGGGTGAGAATGATTGT GGACTGCCATGCTACCCCTG TGCGGGTTACTCAGGATGAA
PHYD_Arabidopsis   AACCGTGTTAGGATGATAGT AGATTGCTATGCGTCACCGG TTCGTGTGGTTCAAGACGAT
PHYE_Arabidopsis   AACCGTGTCCGAATGATTTG TGACTGCAATGCAACTCCGG TTAAGGTTGTTCAGAGTGAG
PHYB_Oryza         AACCGTGTGCGGATGATTGC TGATTGCCATGCTGCGCCGG TGAGGGTCATCCAGGATCCT
PHYA_Avena         AACAAAGTACGGATGATTTG TGATTGCCGTGCGAGATCCA TAAAGGTCATTGAGGCTGAG
PHYA_Oryza         AACAAAGTCCGGATGATTTG TGATTGCCGTGCAAGATCTA TCAAGATTATCGAAGATGAG
PHYA_Zea           AACAAAGTCAGAATGATCTG TGATTGCCGTGCAAGATCCG TGAAGATTATTGAAGATGAG
PHYA_Arabidopsis   AACAAGGTCCGGATGATAGT TGATTGCAATGCAAAACATG CTAGGGTGCTTCAAGATGAA
PHYA_Cucurbita     AATAAGGTCCGGATGATTGT TGATTGTCGTGCAAAACATT TAAAAGTACTCCAAGATGAG
PHYA_Pisum         AACAAGGTCCGTATGATAGT TGATTGTAATGCAAAACATG TGAAGGTTCTTCAAGACGAA
PHYA_Nicotiana     AATAAGGTCCGAATGATTTG TGATTGCCGAGCAAAACATG TGAAGGTAGTCCAAGATGAG
PHYA_Solanum       AATAAGGTCCGAATGATTTG TGATTGCCGAGCAAAACATG TGAAGGTAGTCCAAGATGAG
PHYA_Arabidopsis   AACAAGGTTAGGATGATTTG TGATTGTTCAGCGGTTCCGG TTAAAGTCGTTCAAGATAAG

Physcomitrella     ACCTTGAGGCAGCCGGTCAG CTTGGCAGGTTCGACTTTAC GTTCTCCTCACGGATGCCAT
Ceratodon          GACATTAGGCAGCCAGTCAG CTTGGCAGGTTCGACTTTAC GTGCCCCGCATGGATGTCAT
Selaginella        GAGCTGAGGCAACCGATCAG CCTGGCTGGCTCCACGCTGC GAGCACCGCACGGATGCCAT
Adiantum           ACGCTTAGCCAGCCCATGAG CTTGACAGGGTCAACGCTCC GCGCGCCGCATGGATGCCAC
Psilotum           CAACTAAAACAGCCTCTCAG CTTGGCTGGCTCTACACTGA GGGCCCCTCATGGTTGTCAT
Psilotum2          -------------------- -------------------- --------------------
PHYB_Arabidopsis   AGGCTAACTCAGTCTATGTG CTTGGTTGGTTCTACTCTTA GGGCTCCTCATGGTTGTCAC
PHYB_Nicotiana     TCACTGATGCAGCCTTTATG TTTAGTTGGTTCCACACTTA GAGCCCCTCATGGTTGCCAC
PHYB_Solanum       TCACTGATGCAGCCTTTATG TCTAGTTGGTTCCACACTTA GAGCACCTCATGGTTGCCAC
PHYD_Arabidopsis   AGGCTCACGCAGTTTATATG CTTGGTGGGTTCGACTTTGC GAGCTCCTCATGGCTGTCAT
PHYE_Arabidopsis   GAACTCAAGAGACCACTTTG TTTAGTTAATTCTACTCTAA GAGCTCCTCATGGCTGCCAT
PHYB_Oryza         GCACTAACACAGCCGCTGTG CTTGGTTGGGTCCACGCTGC GTTCGCCGCATGGTTGCCAT
```

```
PHYA_Avena         GCACTCCCCTTTGATATTAG CCTATGTGGTTCAGCACTCA GGGCACCACACAGTTGTCAC
PHYA_Oryza         TCGCTCCACTTGGATATTAG CTTATGTGGTTCAACACTGA GGGCACCACACAGTTGTCAT
PHYA_Zea           GCACTCTCCATTGATATTAG CTTGTGTGGTTCAACTCTTA GAGCACCACATAGCTGTCAC
PHYA_Arabidopsis   AAGCTTTCCTTTGACCTTAC CTTGTGTGGCTCCACCCTTA GAGCACCGCACAGCTGCCAT
PHYA_Cucurbita     AAATTACAGTTTGATCTAAC TTTATGTGGTTCAACTTTAA GAGCTCCACACAGTTGCCAT
PHYA_Pisum         AAACTCCCATTTGATTTGAC TCTGTGCGGTTCGACCTTGA GAGCTCCACATAGTTGCCAT
PHYA_Nicotiana     AAGCTTCCATTTGATTTAAC ATTGTGCGGCTCTACTCTTA GGGCTCCTCACTACTGCCAT
PHYA_Solanum       AAGCTTCCATTTGACTTAAC ATTGTGCGGCTCTACTCTTA GGGCCCCTCACTACTGCCAT
PHYA_Arabidopsis   AGTCTCTCACAGCCAATAAG TCTTTCTGGATCTACTTTGA GAGCTCCTCATGGTTGTCAC

Physcomitrella     GCCCAGTACATGGGCAATAT GGGGTCCATTGCGTCCCTAG TCATGGCTGTGATCATCAAC
Ceratodon          GCCCAGTACATGGGTAACAT GGGCTCCATTGCGTCGCTTG TCATGGCCGTAATCATCAAT
Selaginella        GCGCAGTACATGGGCAACAT GGGCTCTGTTGCGTCCCTGG TCATGGCCATGATCATCAAC
Adiantum           ACCCAATACATGGCCAACAT GAATTCAATCTCCTCCTTGG TGATGGCTGTAATTGTGAAT
Psilotum           GCGCACTATATGGGTAATAT GGGTTCTATTGCCTCTTTGG TGATGGCTGTTATAGTAAAG
Psilotum2          -------------------- -GGATCCATTGGCTCTTTGG TAATGGCTGTTATAGTGAAC
PHYB_Arabidopsis   TCTCAGTATATGGCTAACAT GGGATCTATTGCGTCTTTAG CAATGGCGGTTATAATCAAT
PHYB_Nicotiana     GCGCAGTACATGGCAAATAT GGGGTCTATTGCGTCATTAA CACTAGCAGTTATTATCAAT
PHYB_Solanum       GCACAGTACATGGCAAATAT GGGGTCTATTGCCTCATTAA CACTGGCAGTTATTATCAAC
PHYD_Arabidopsis   GCTCAATACATGACTAACAT GGGCTCTATTGCGTCGTTAG CTATGGCAGTTATAATAAAT
PHYE_Arabidopsis   ACGCAGTATATGGCGAATAT GGGCTCTGTAGCTTCTCTTG CACTCGCAATTGTAGTAAAA
PHYB_Oryza         GGCCAGTATATGGCGAACAT GGGTTCCATTGCATCTCTTG TTATGGCAGTGATCATTAGT
PHYA_Avena         CTTCAGTATATGGAGAACAT GAACTCGATTGCATCCCTTG TCATGGCTGTTGTGGTTAAT
PHYA_Oryza         CTTCAGTATATGGAGAACAT GAACTCGATTGCATCCCTTG TCATGGCTGTTGTGGTTAAT
PHYA_Zea           CTTAAGTATATGGAGAACAT GAACTCGATTGCATCCCTTG TCATGGCTGTTGTGGTCAAT
PHYA_Arabidopsis   TTGCAGTACATGGCCAACAT GGATTCAATTGCATCTCTGG TTATGGCGGTTGTAGTTAAC
PHYA_Cucurbita     TTACAGTATATGGAAAACAT GAACTCTATAGCCTCTTTGG TTATGGCAGTTGTGGTTAAT
PHYA_Pisum         TTGCAGTACATGGCTAACAT GGATTCAATTGCTTCGTTGG TTATGGCAGTTGTCGTCAAT
PHYA_Nicotiana     CTACAGTATATGGAGAACAT GAGTTCAATTGCATCCCTTG TAATGGCAGTTGTGGTCAAT
PHYA_Solanum       TTACAGTATATGGAGAACAT GAATTCAATTGCATCACTTG TAATGGCAGTTGTGGTCAAT
PHYA_Arabidopsis   GCACAGTATATGAGTAATAT GGGATCAGTGGCGTCTCTTG TCATGTCTGTAACTATCAAT

Physcomitrella     GATAACGAGGAAGAT----- -------TCG------CATG GC---TCAGTCCAAAGA---
Ceratodon          GATAACGAGGAATAT----- -------TCA------CGTG GG---GCAATTCAAAGA---
Selaginella        GACAACGATGAACCGTCTGG C------GGC------GGCG GCGGCGGCGGGCAGCACAAG
Adiantum           GATAGTGATGACGAT----- -------AGT------GCCG GCCATTCTTCGCAA------
Psilotum           CGTCATGGCGAGGAGGATAG G------AGC------CTGG GCTTTCAGTCACAAAAC---
Psilotum2          GACAACGATGCAGAGCCGAG C------ggg------CGCG GAAATCAGCCGAAGAAC---
PHYB_Arabidopsis   GGAAATGAAGATGATGGGAG C------AAT------GTAG CT------AGTGGAAGAAGC
PHYB_Nicotiana     GGAAACGATGAGGAA----- -------GCT------GTT- --------GGGGGCCGAAGT
PHYB_Solanum       GGAAATGATGAGGAA----- -------GCT------GTGG GT------GGCGGTCGAAAT
PHYD_Arabidopsis   GGAAACGAAGAAGATGGTAA T------GGG------GTT- --AATACTGGAGGAAGAAAC
PHYE_Arabidopsis   -------------------- -------------------- -----------GGCAAAGAT
PHYB_Oryza         AGTGGTGGGGATGATGATCA T------AACATTGCACGGG GC------AGCATCCCGTCG
PHYA_Avena         GAGAATGAAGAGGATGATGA AGCTGAGTCTGAACAACCAG CA------CAGCAGCAGAAA
PHYA_Oryza         GAGAATGAGGATGATGATGA AGTTGGGGCTGATCAACCTG CA------CAACAGCAGAAG
PHYA_Zea           GAAAATGAAGAGGATGATGA ACCCGAGCCTGAACAACCAC CA------CAACAGCAGAAG
PHYA_Arabidopsis   GAGGAAGATGGAGAAGGGGA T------GCTCCTGATGCTA CTACA---CAGCCTCAAAAG
PHYA_Cucurbita     GAAGGGGATGAAGAAAATGA A------GGC------CCTG CT---TTGCAGCAGCAAAAG
PHYA_Pisum         GACAGCGATGAAGATGGAGA T------AGCGCTGACGCA- -----GTTCTCCCACAAAAG
PHYA_Nicotiana     GACGGGGACGAAGAGGGAGA A------AGCTCTGATTCGA CA------CAATCTCAAAAG
PHYA_Solanum       GACGGGGATGAAGAAGGAGA A------AGCTCTGATTCTT CA------CAATCTCAAAAA
PHYA_Arabidopsis   GGTAGTGATAGTGATGAGAT G------AACAGAGATTTA- --------------CAGACT

Physcomitrella     GGTAGAAAGCTATGGGGACT TGTAGTGTGTCATCATACAT CTCCACGAACTGTACCGTTT
Ceratodon          GGTAGAAAGCTGTGGGGACT CGTAGTGTGTCAGCATACAT CTCCACGAACTGTACCGTTT
Selaginella        GGAAGAAGGCTGTGGGGGCT GGTGGTGTGTCACCACACGT CCCCGCGATCGGTGCCCTTC
Adiantum           GGCATTAAACTTTGGGGCTT GGTCGTATGCCACCATACGT CACCACGCTATGTACCTTTC
Psilotum           GGA---AGGCTATGGGGGAT GGTGGTTTGTCACCACACGA CTCCCAGAGCTGTCCCTTTC
Psilotum2          ---AGGAGGCTCTGGGGAAT GGTTGTTTGCCACCACACCA CTCCCAGGGCGGTCCCGTTC
PHYB_Arabidopsis   TCGATGAGGCTTTGGGGTTT GGTTGTTTGCCATCACACTT CTTCTCGCTGCATACCGTTT
PHYB_Nicotiana     TCAATGAGGCTGTGGGGCTT GGTTGTTGGACACCATACTT CTGCTAGGTGCATTCCATTC
PHYB_Solanum       TCAATGAGGCTATGGGGCTT GGTTGTTGGACACCACACTT CTGTTCGGTCCATTCCTTTC
PHYD_Arabidopsis   TCGATGAGGCTTTGGGGTTT AGTTGTTTGCCATCACACAT CAGCTCGTTGCATACCTTTT
PHYE_Arabidopsis   TCGAGCAAGCTTTGGGGATT AGTTGTTGGTCATCATTGTT CTCCTAGATACGTTCCATTC
PHYB_Oryza         GCGATGAAGTTGTGGGGGTT GGTAGTATGCCACCACACAT CTCCACGGTGCATCCCTTTC
PHYA_Avena         AAGAAGAAACTATGGGGCCT CCTTGTTTGCCACCATGAGA GCCCCAGATATGTCCCTTTT
PHYA_Oryza         AGGAAGAAACTATGGGGACT CCTTGTTTGCCACCATGAGA GCCCCAGATATGTTCCTTTC
PHYA_Zea           AAGAAGAGGCTGTGGGGTCT CATTGTTTGCCACCATGAGA GCCCCAGATATGTCCCTTTC
PHYA_Arabidopsis   AGAAAGAGACTATGGGGTTT AGTGGTTTGTCACAATACGA CTCCGAGGTTTGTTCCATTT
PHYA_Cucurbita     AGAAAGAGATTATGGGGCTT GGTAGTATGTCATAATTCAA GTCCCAGATTTGTTCCGTTT
PHYA_Pisum         AAAAAGAGACTTTGGGGTTT GGTAGTTTGTCATAACACTA CTCCAAGGTTTGTTCCTTTT
PHYA_Nicotiana     AGAAAAAGGCTTTGGGGCCT GGTGGTTTGCCACAACACCA CCCCGAGGTTTGTTCCCTTC
```

```
PHYA_Solanum        AGAAAAAGGCTATGGGGCTT GGTTGTTTCCCACAACACGA CCCCAAGGTTCGCGCCCTTC
PHYA_Arabidopsis    GGCAGACACTTATGGGGCTT GGTGGTTTGTCATCACGCAA GTCCTAGATTTGTTCCGTTT

Physcomitrella      CCCCTTCGGTCCGCGTGTGG GTTTTTGATGCAGGTGTTCG GATTACAGCTTAACATGGAG
Ceratodon           CCACTTCGGTCTGTGTGCGA GTTTTTGATGCAGGTATTTG GTATGCAGCTCAACCTCCAT
Selaginella         ---CTGCGCTCGGCGTGCGA GTTCCTCATGCAGGTGTTTG GCCTGCAGCTCAACATGGAG
Adiantum            CCCGTAAGGTCGGCTTGTGA GTTCTTGATGCAGGTCTTTA GTCTTCAACTAAATATGGAG
Psilotum            GCTCTCAGGTGTGCTTGCGA GTTTTTCGCACAGGTCTTTG CCCTTCAACTAAACATGGAG
Psilotum2           TCTCTGAGGTCAGCTTGTGA GTTTTTGATGCAGGTCTTTG GCCTTCAACTAAACATGGAA
PHYB_Arabidopsis    CCGCTAAGGTATGCTTGTGA GTTTTTGATGCAGGCTTTCG GTTTACAGTTAAACATGGAA
PHYB_Nicotiana      CCTCTTCGGTATGCCTGTGA ATTCCTTATGCAGGCCTTTG GACTCCAATTGAACATGGAG
PHYB_Solanum        CCTCTTAGGTATGCATGTGA ATTCCTTATGCAGGCCTTTG GACTCCAATTGAACATGGAG
PHYD_Arabidopsis    CCTTTGAGGTACGCTTGTGA GTTCTTTATGCAGGCCTTTG GCTTACAGCTAAACATGGAG
PHYE_Arabidopsis    CCGTTGCGGTATGCTTGTGA GTTTCTGATGCAAGCATTTG GGCTTCAGCTTCAAATGGAA
PHYB_Oryza          CCACTACGGTATGCATGCGA GTTCCTCATGCAAGCCTTTG GGTTGCAGCTCAACATGGAG
PHYA_Avena          CCGCTGCGTTATGCTTGTGA GTTCTTAGCACAGGTGTTTG CTGTCCATGTCAACAGGGAG
PHYA_Oryza          CCATTGCGGTATGCCTGTGA GTTCTTAGCACAAGTGTTTG CTGTCCATGTTAACAAGGAG
PHYA_Zea            CCACTGCGGTATGCCTGTGA ATTCTTGGCCCAAGTGTTTG CTGTCCATGTAAACAAGGAG
PHYA_Arabidopsis    CCTCTCAGGTATGCCTGTGA GTTTCTAGCTCAAGTGTTTG CCATACACGTCAATAAGGAG
PHYA_Cucurbita      CCTCTTAGGTATGCTTGTGA GTTCCTAGCTCAAGTATTTG CTATTCATGTGAACAAGGAA
PHYA_Pisum          CCTCTAAGGTATGCTTGTGA GTTTCTGGCTCAAGTGTTTG CCATCCATGTGAACAAAGAA
PHYA_Nicotiana      CCTCTGAGGTATGCATGTGA ATTTCTTGCGCAAGTCTTTG CCATACACGTCAACAAGGAA
PHYA_Solanum        CCACTGAGGTATGCATGTGA GTTTCTTGCACAAGTCTTTG CCATACTCGTTAACAAGGAA
PHYA_Arabidopsis    CCATTACGATATGCTTGTGA ATTCTTGACTCAAGTATTTG GCGTGCAGATCAACAAAGAA

Physcomitrella      GTCGAGTCAGCCGCTCAGCT AAGGGAAAAACATATTCTCA GAACTCAAACTCTACTTTGC
Ceratodon           GTTGAGCTGGCCGCTCAACT AAGGGAAAAACATATTCTCA GAACTCAAACTCTTCTTTGT
Selaginella         GCGGCAGTGGCGGCGCACGT GCGGGAGAAGCACATACTGC GCACCCAGACGCTCTTGTGC
Adiantum            GTGGGGATGGCTGCTCAGGT GAGAGAAAAGCACATCCTTC GCACACAAACACTTCTTTGT
Psilotum            CTAGAACTAGCAGCTCAAAT GAGAGAAAAGGATATTCTTC GCACTCAAAGTCTTCTCTGT
Psilotum2           ATCGAGTTGGCTGCTCAAAT GCGAGAAAAACACATTCTTC GCACTCAGACACTTCTATGC
PHYB_Arabidopsis    TTGCAGTTAGCTTTGCAAAT GTCAGAGAAACGCGTTTTGA GAACGCAGACACTGTTATGT
PHYB_Nicotiana      TTGCAACTGGCATCACAGTT GTCTGAGAAACATGTTTTGA GGACACAAACACTGTTATGT
PHYB_Solanum        TTGCAATTGGCGTCACAGTT GTCTGAGAAACATGTTTTAA GGACACAAACACTGTTATGT
PHYD_Arabidopsis    TTGCAGTTAGCCTTGCAGGT GTCTGAAAAACGCGTTCTGA GAATGCAGACACTATTATGT
PHYE_Arabidopsis    CTTCAGTTAGCATCACAGTT AGCCGAGAAGAAGGCTATGC GGACGCAGACCTTGTTGTGC
PHYB_Oryza          TTGCAGCTTGCACACCAACT GTCAGAGAAACACATTCTGC GGACCGGAACACTGCTGTGT
PHYA_Avena          TTTGAATTAGAGAAACAGTT GCGTGAGAAGAACATACTGA AGATGCAAACAATGCTCTCT
PHYA_Oryza          TTTGAATTAGAGAGGCAAGT ACGCGAGAAAAGCATATTGA GGATGCAAACAATGCTCTCT
PHYA_Zea            TTTGAATTGGAGAAGCAGAT ACGAGAGAAAAACATTCTGC GAATGCAAACAATGCTCTCT
PHYA_Arabidopsis    GTGGAACTCGATAACCAGAT GGTGGAGAAGAACATTTTGC GCACGCAGACACTCTTGTGC
PHYA_Cucurbita      TTAGAGTTGGAAAATCAAAT TATAGAAAAGAATATTCTGC GTACGCAGACACTCTTGTGC
PHYA_Pisum          ATAGAGTTAGAATATCAGAT TCTTGAGAAGAATATCCTGC GCACGCAGACGCTGTTGTGT
PHYA_Nicotiana      CTGGAATTGGAAAGTCAGAT TCTTGAGAAAAATATCCTGC GTACTCAGACTCTCTTGTGT
PHYA_Solanum        CTGGAATTGGAAAATCAATT CCTTGAGAAAAATATTCTGC GTACTCAGACTCTCTTGTGT
PHYA_Arabidopsis    GCGGAATCAGCTGTTCTGTT GAAAGAGAAGCGTATTTTGC AAACTCAGAGTGTGCTATGT

Physcomitrella      GACATGCTTCTTCGAGAT-- -GCTCCTATTGGTATTGTTT CTCAGATTCCAAATATTATG
Ceratodon           GACATGCTTCTTCGAGAT-- -GCTCCTATTGGAATTGTAT CTCAAACTCCAAATATTATG
Selaginella         GACATGCTGCTGCGCGAC-- -GCTCCCATCGGCATTGTTT CTCAGTCGCCAAACATCATG
Adiantum            GATATGCTTCTCAGAGAT-- -GCTCCCATTGGTATCGTCT CGCAATCACCCAACATCATG
Psilotum            GATATGCTTCTTCGTGAC-- -GCTCCTATTGGCATTGTCA CTCGTTCCCCGAACGTAATG
Psilotum2           GATATGCTTCTTAGAGAT-- -GCTCCTATCGGCATCGTCA CGCAGTCCCCTAATATAATG
PHYB_Arabidopsis    GATATGCTTCTGCGTGAC-- -TCGCCTGCTGGAATTGTTA CACAGAGTCCCAGTATCATG
PHYB_Nicotiana      GACATGCTCCTTCGAGAC-- -TCACCTACGGGGATTGTTA TCCAGAGCCCCAGTATTATG
PHYB_Solanum        GACATGCTCCTTCGAGAC-- -TCTCCACCGGGGATTGTTA CCCAAAGCCCCAGTATTATG
PHYD_Arabidopsis    GATATGCTTCTACGTGAC-- -TCACCAGCGGGGATTGTCA CGCAGAGGCCTAGTATCATG
PHYE_Arabidopsis    GATATGCTTCTCCGTGAT-- -ACTGTTTCCGCTATTGTTA CACAATCTCCGGGTATTATG
PHYB_Oryza          GATATGCTACTCCGGGAT-- -TCACCAACTGGCATTGTCA CACAAAGCCCCAGCATCATG
PHYA_Avena          GATATGTTGTTCCGAGAAGC CTCTCCCCTGACTATCGTAT CAGGGACCCCCAATATCATG
PHYA_Oryza          GACATGCTTCTCAGGGAATC CTCTCCTCTGAGTATAGTAT CAGGGACTCCCAACATCATG
PHYA_Zea            GACATGCTGTTCAAGGAATC ATCTCCCTTGAGTATCGTGT CTGGGAGTCCAAATATCATG
PHYA_Arabidopsis    GATATGCTGATGCGTGAT-- -GCTCCACTGGGTATTGTGT CGCAAAGCCCCAACATAATG
PHYA_Cucurbita      GACATGCTAATGCGTGAT-- -GCTCCTTTAGGTATTGTGT CGAGGAGTCCTAACATAATG
PHYA_Pisum          GATATGTTGATGCGAGAT-- -GCACCCTTAGGTATTGTAT CACAAAGCCCTAATATAATG
PHYA_Nicotiana      GATATGCTGATGCGAGTA-- -GCTCCCTTAGGTATAGTGT CACAGAGTCCAAATATTATG
PHYA_Solanum        GATATGCTGATGCGAGAT-- -GCTCCCCTAGGTATCGTGT CACAGAGCCCCAACATTATG
PHYA_Arabidopsis    GACATGCTTTTCCGCAAT-- -GCACCAATAGGTATAGTCA CTCAATCACCAAATATAATG

Physcomitrella      AGCCATTCTGGCTTCTGGGT GACTTGGTGAAATGCGATGG AGCTGCTCTTTATTATGGGA
Ceratodon           AGCGAGTGTGGCTTCTTGGC GATCTTGTGAAATGTGATGG AGCAGCTCTTTACTATGGGA
Selaginella         AACGCTTCTGGTTACTGGGC GACCTGGTCAAGTGCGACGG TGCCGCACTCTACTACGGCA
Adiantum            AGAAATGCTGGCTTTTAGGT GATTTGGTGACATGTGATGG TGCAGCATTGTATTATGGGA
```

```
Psilotum            AAAGGCTTTGGTTGCTTGGA GATCTTGTGAAGTGTGATGG GGCTGCGTTATATTACGGTG
Psilotum2           AAAAGTTTTGGTTGCTTGGA GATCTTGTGAAGTGTGATGG AGCTGCTTTGTATTACGGGA
PHYB_Arabidopsis    GGAAGTATTACCCGTTGGGT GACTTAGTGAAATGTGACGG TGCAGCATTTCTTTACCACG
PHYB_Nicotiana      GGAAGTACTATCCATTAGGC GACCTTGTGAAGTGCGATGG CGCTGCTCTGTACTGCCAGG
PHYB_Solanum        GGAAGTACTATCCATTAGGC GACCTTGTGAAGTGCGATGG TGCTGCTCTATACTACCAGG
PHYD_Arabidopsis    GGAAGTATTATCCGTTGGGT GATTTAGTAAAATGTAATGG TGCGGCATTTCTTTACCAAG
PHYE_Arabidopsis    GGAAATGTTGGTTGGTTGGT GACCTTGTGAAATGTGATGG AGCTGCGTTATATTACAAGG
PHYB_Oryza          GGAAGTACTACCCTCTTGGT GACCTTGTGAAGTGTGATGG TGCTGCTCTGTATTACCATG
PHYA_Avena          GAAAAGTATGGCGTCTGCGT GACCTAGTCAAATGTGATGG TGCTGCTCTTCTGTATGGGG
PHYA_Oryza          GAAAAGTGTGGCGGCTACAG GACCTTGTGAAATGTGATGG TGCTGCTCTTTTGTATGGGG
PHYA_Zea            ACAAAGTATGGCGGCTTCAA GACCTAGTTAAGTGTGATGG CGCTGCTCTTTTGTATGGGG
PHYA_Arabidopsis    ACAAGATATGGAAACTGGGA GACCTTGTGAAATGTGATGG AGCAGCTCTCTTGTATAAAG
PHYA_Cucurbita      AAAAAATTTGGCGATTAGGA GATCTTGTCAAATCTGATGG GGCTGCCTTGTTATATAAGA
PHYA_Pisum          ACAAGTTATGGTTATTAGGA GATCTAGTGAAATGTGATGG GGCTGCACTCTTTTATAGAA
PHYA_Nicotiana      ATAAGATACATCGACTTGGA GATCTTGTCAAATGTGATGG TGCTGCTTTGCTCTATAAGA
PHYA_Solanum        ATAAGATACATCGATTAGGA GATCTCATCAAATGTGATGG TGCTGCTTTGCTCTATAAGA
PHYA_Arabidopsis    ACAACCTCTGGTCTCTAGGA GATCTTGTTAAATGTGATGG AGCAGCATTATATTACAGAG

Physcomitrella      ACTACTCCTACGGAGAGTCA GATTAAAGATATTGCAGAAT GGTTGCTAGAATACCACAAG
Ceratodon           ACGACACCGACTGAGAATCA GATCAAAGAGATTGCAGACT GGTTGCTAGAGCATCACAAC
Selaginella         ATCACGCCCAGCGAGGCACA GATCAAAGACATCGCCGAGT GGCTGTTGGAGCATCACAAA
Adiantum            ACTACACCAACAGAGGCGCA GATTGTGGACATTGCCGCGT GGTTGTTAGATTGCCACAAG
Psilotum            ACAACTCCCACAAAGGCTCA AATAAAAGACATTGTTGACT GGCTTTCAGAGGTTCACAGG
Psilotum2           ACAACTCCCACGGAGGCTCA GATAAAAGATCTTGCTGATT GGCTTCTAGACGTTCACAGG
PHYB_Arabidopsis    GTTGCTCCTAGTGAAGTTCA GATAAAAGATGTTGTGGAGT GGTTGCTTGCGAATCATGCG
PHYB_Nicotiana      GTTACACCAACTGAAGCTCA GATAAAGGACATTGTGGAGT GGTTATTGACTTACCATGGG
PHYB_Solanum        GTAACACCAACTGAAGCTCA GATAAAAGACATTGTGGAGT GGTTATTGGCTTACCATGGA
PHYD_Arabidopsis    GTGACTCCAACTGATTCTCA GATTAATGACATTGTGGAGT GGTTGGTTGCTAACCATTCT
PHYE_Arabidopsis    GTTACTCCTAATGAGTCACA AGTTAAAGACTTGGTGAATT GGCTGGTGGAGAATCACGGT
PHYB_Oryza          GTCACTCCCACAGAAGTTCA GATTAAGGACATCATCGAGT GGTTGACTATGTGCCATGGA
PHYA_Avena          AATGCTCCAACGGAGTCTCA GATACATGATATCGCCTTCT GGCTATCAGATGTTCACAGG
PHYA_Oryza          AATGCTCCAACTGAGTCTCA GATACGTGATATTGCCTTCT GGCTGTCAGATGTCCACAGG
PHYA_Zea            ACGGCTCCAACCGAGTCTCA GATTCGTGATATTGCCTTCT GGCTTTCAGAAGTTCATGGG
PHYA_Arabidopsis    ACAACTCCAAGTGAGTTCCA CCTGCAGGAGATAGCTTCAT GGTTGTGTGAATACCACATG
PHYA_Cucurbita      TTGACACCTAATGACTTCCA GTTGCTGGACATAGCTTCGT GGCTTTCCGAGTATCATATG
PHYA_Pisum          GCGACACCGACTGAATCTCA ATTAAGAGAGATAGCTTTAT GGATGTCTGAGTATCATACA
PHYA_Nicotiana      ATGACCCCAAGCGACTTTCA GCTGCACGATATTGTCTCGT GGCTTTCTGAGTATCATACA
PHYA_Solanum        ATGAACCCAAGTGACTTTCA GCTGCATGATATAGTATCAT GGCTTTGTGAGTATCATACA
PHYA_Arabidopsis    GTTACTCCCACAGAGACACA AATTAGAGATCTAATTGACT GGGTTCTCAAAAGTCATGGA

Physcomitrella      GAC---TCAACAGGTCTGAG TACTGATAGTTTAGCGGATG CCAATTATCCAGCCGCACAC
Ceratodon           GAC---TCAACAGGTCTTAG TACGGATAGTTTAGCGGATG CGAATTATCCAGGTGCACAC
Selaginella         GAC---TCAACCGGCCTGAG CACCGACAGCCTGGCTGACG CAGGCTACCCGGGTGCTGCG
Adiantum            GAC---TCCACTGGCCTTAG TACCGACAGTCTTGCGAAAA CCGGGTATCCAGAGGCTAGC
Psilotum            GAC---TCCACGGGACTCAG TACAGAGAGTCTTATTGAGG CAGGTTACCCTGGTGCAGCC
Psilotum2           GAC---TCCACGGGTCTGAG CACAGAGAGCCTTGCAGATG CAGGTTTCCCCGGTGCAGCT
PHYB_Arabidopsis    GAT---TCAACCGGATTAAG CACTGATAGTTTAGGCGATG CGGGGTATCCCGGTGCAGCT
PHYB_Nicotiana      GAC---TCAACAGGTTTAAG TACTGACAGTTTGGCTGATG CAGGATATCCTGGGGCAGCT
PHYB_Solanum        GAC---TCAACAGGTTTAAG TACTGACAGTTTGCCTGATG CTGGGTATCCTGGAGCAGCT
PHYD_Arabidopsis    GAT---TCTACCGGGTTAAG CACAGATAGTTTAGGCGATG CGGGTTATCCTCGGGCAGCT
PHYE_Arabidopsis    GATGATTCGACGGGTTTAAC CACTGATAGTTTGGTGGATG CGGGATACCCTGGTGCCATC
PHYB_Oryza          GAC---TCCACAGGGCTCAG CACAGATAGCCTTGCTGATG CAGGCTACTCTGGTGCTGCT
PHYA_Avena          GAT---TCCACTGGCCTGAG TACTGACAGCCTCCATGATG CTGGCTATCCAGGAGCTGCT
PHYA_Oryza          GAT---TCCACTGGCCTGAG TACTGATAGCCTACATGATG CTGGATATCCAGGAGCTGCT
PHYA_Zea            GAT---TCCACTGGCTTGAG TACTGATAGCCTCCAGGATG CTGGATATCCAGGAGCTGCT
PHYA_Arabidopsis    GAT---TCAACGGGTTTGAG CACTGATAGTTTGCATGACG CCGGGTTTCCTAGGGCTCTA
PHYA_Cucurbita      GAT---TCAACGGGGTTGAG TACTGACAGTTTGTATGATG CAGGATACCCTGGAGCTATT
PHYA_Pisum          GAT---TCAACAGGTTTGAG TACAGACAGCTTGTCGGATG CAGGGTTTCCAGGGGCTCTT
PHYA_Nicotiana      GAT---TCCACAGGTTTGAG TACAGACAGCTTGTATGATG CTGGTTTCCCTGGGGCTCTT
PHYA_Solanum        GAT---TCCACAGGCTTGAG TACGGATAGCTTGTATGATG CTGGGTTTCCTGGGGCTCTT
PHYA_Arabidopsis    GGA---AACACTGGCTTTAC CACTGAAAGTCTAATGGAGT CTGGCTATCCGGATGCTTCT

Physcomitrella      CTGCTCGGGGATGCTGTGTG TGGTATGGCAGCTGCGAAGA TCACTGCGAAGGATTTTCTT
Ceratodon           CTGCTTGGCGACGCTGTTTG TGGTATGGCAGCTGCAAAAA TCACTGCGAAGGATTTCCTT
Selaginella         TCGCTAGGGGACGAGGTGTG CGGCATGGCTGCCGCCAAGA TCACCGCGAAGGATTTCCTG
Adiantum            TGCTTGGGTGATGCTGTCTG TGGATTGGCAGCTGCAAAGA TAACAGCGACTGACTTCTTG
Psilotum            ACTCTTGGTGATGCGGTTTG TGGAATGGTTGCGGTTAAGA TTACACAAAGAGACTACTTG
Psilotum2           GCTCTGGGCGATGCAGTTTG TGGCATGGCGGCGGTAAAGA TTACACCCAGAGACTACTTG
PHYB_Arabidopsis    GCGTTAGGGGATGCTGTGTG CGGTATGGCAGTTGCATATA TCACAAAAAGAGACTTTCTT
PHYB_Nicotiana      TTGCTTGGTGATGCAGTTTG TGGTATGGCTGTTGCTTATA TAACTTCTAAAGATTTCTTA
PHYB_Solanum        TCACTTGGTGATGCAGTTTG TGGTATGGCTGTCGCTTATA TAACTTCTAAAGATTTCTTG
PHYD_Arabidopsis    GCTTTGGGAGATGCTGTGTG CGGTATGGCAGTCGCGTGTA TCACAAAAAGGGACTTCCTT
PHYE_Arabidopsis    TCACTTGGAGATGCCGTTTG TGGTGTGGCTGCCGCAGAAT TTTCTTCAAAAGATTACTTA
```

PHYB_Oryza	GACCTAGGAGATGCAGTGAG	CGGAATGGCGGTAGCATATA	TCACGCCAAGTGATTATTTG
PHYA_Avena	GCTCTTGGTGATATGATTTG	TGGAATGGCAGTGGCTAAGA	TCAACTCCAAGGATATTCTT
PHYA_Oryza	GCTCTTGGTGATATGATTTG	TGGAATGGCAGTAGCTAAAA	TAAATTCCAAGGATATCCTG
PHYA_Zea	TCCCTTGGTGACATGATTTG	TGGAATGGCAGTGGCCAAGA	TCACGTCCAAGGACATTCTT
PHYA_Arabidopsis	TCTCTCGGGGATTCGGTATG	TGGGATGGCAGCTGTGAGGA	TATCATCGAAAGACATGATT
PHYA_Cucurbita	GCTTTAGGTGATGAAGTGTG	TGGGATGGCAGCTGTGAGGA	TAACTAATAATGACATGATT
PHYA_Pisum	TCTCTTAGTGATACTGTATG	TGGAATGGCAGCTGTTAGAA	TAACTTCAAAAGACATAGTT
PHYA_Nicotiana	GCTCTTGGTGATGTAGTGTG	TGGTATGGCAGCTGTTAGAA	TATCTGATAAGGGCTGGCTG
PHYA_Solanum	GCTCTTGGTGATGCAGTCTG	TGGTATGGCGGCTGTAAGAA	TATCTGATAAGGACTGGCTG
PHYA_Arabidopsis	GTTCTTGGGGAGTCAATATG	TGGAATGGCTGCCGTATATA	TTTCCGAAAAAGATTTCCTT

Physcomitrella	TTTTGGTTCAGGTCTCATAC	TGCTAAAGAGATTAAGTGGG	GTGGTGCTAAACACGATCCA
Ceratodon	TTCTGGTTCAGGTCTCACAC	TGCTACAGAGGTCAAATGGG	GTGGTGCTAAACACGATCCA
Selaginella	TTTTGGTTCCGGTCGCACAC	TGCCAAGGAGGTGAAGTGGG	GTGGTGCCAAGCACGACCCG
Adiantum	TTTTGGTTTAGATCTCATAC	TGCGAAAGAGGTTAGGTGGG	GAGGAGCCCGGCATGACCCC
Psilotum	TTTTGGTTCAGGTCTCACAC	TGCAAAAGAGATTCAATGGG	GTGGTGCAAAGCACCATCCA
Psilotum2	TTTTGGTTCAGGTCTCATAC	CGCGAAAGAGATTAAGTGGG	GAGGAGCCAAACACAATCCC
PHYB_Arabidopsis	TTTTGGTTTCGATCTCACAC	TGCGAAAGAAATCAAATGGG	GAGGCGCTAAGCATCATCCG
PHYB_Nicotiana	TTTTGGTTTCGCTCCCATAC	AGCGAAAGAGATAAAGTGGG	GTGGTGCAAAGCATCATCCT
PHYB_Solanum	TTTTGGTTTCGCTCCCACAC	AGCGAAAGAAATAAAGTGGG	GTGGTGCAAAGCATCATCCT
PHYD_Arabidopsis	TTCTGGTTTCGGTCTCATAC	TGAGAAAGAAATCAAATGGG	GAGGGGCTAAGCACCATCCT
PHYE_Arabidopsis	CTTTGGTTCAGGTCCAATAC	TGCAAGTGCAATCAAATGGG	GAGGAGCTAAACATCATCCA
PHYB_Oryza	TTTTGGTTCCGGTCACACAC	AGCTAAGGAGATAAAGTGGG	GTGGTGCAAAGCATCATCCA
PHYA_Avena	TTTTGGTTCAGGTCACATAC	AGCTGCTGAAATCAGATGGG	GAGGTGCAAAGAATGATCCA
PHYA_Oryza	TTCTGGTTCAGGTCACATAC	AGCTGCTGAAATCAGATGGG	GAGGTGCAAAACATGATCCA
PHYA_Zea	TTCTGGTTCAGGTCACATAC	AGCTGCTGAAATCAAATGGG	GAGGTGCAAAGCATGATCCA
PHYA_Arabidopsis	TTCTGGTTCCGTTCTCATAC	CGCTGGTGAAGTGAGATGGG	GAGGTGCGAAGCATGATCCA
PHYA_Cucurbita	TTTTGGTTTCGATCTCACAC	TGCTTCAGAGATTCGATGGG	GTGGAGCAAAGCATGAACAT
PHYA_Pisum	TTCTGGTTTAGGTCGCACAC	TGCTGCAGAAATCCGATGGG	GTGGTGCAAAGCATGAACCG
PHYA_Nicotiana	TTCTGGTATAGGTCACACAC	TGCTGCTGAAGTTAGATGGG	GTGGAGCAAAGCATGAACCT
PHYA_Solanum	TTCTGGTACAGGTCACACAC	TGCTGCTGAAGTTAGATGGG	GTGGTGCAAAGCATGAACCT
PHYA_Arabidopsis	TTCTGGTTCCGGTCTAGCAC	TGCAAAACAGATCAAGTGGG	GTGGTGCAAGACACGATCCT

Physcomitrella	GGCGAAAACCACGATGGTCG	AAAAATGCACCCTAGAAGCT	CTTTTAAAGCCTTTTTGGAG
Ceratodon	GATGAAAAAGATGATGGCCG	AAAAATGCATCCCAGAAGCT	CTTTCAAAGCCTTTTTAGAG
Selaginella	GACGACAAAGATGATGGGCG	CAAGATGCATCCCCGCTCCT	CCTTCAAGGCCTTTCTGGAG
Adiantum	GAGGAGAGGGATGATGGCCG	CCGCATGCATCCCCGGTCTT	CGTTCAAGGCTTTCTTAGAA
Psilotum	GATGAAAGGGATGATGGGAG	GAAAATGCATCCTCGTTCTT	CTTTCAAAGCATTCTTGGAA
Psilotum2	GATGATAAAGATGATGGGAA	GAAAATGCACCCTAGATCTT	CCTTCAAGGCATTCTTGGAA
PHYB_Arabidopsis	GAGGATAAAGATGATGGGCA	ACGAATGCATCCTCGTTCGT	CCTTTCAGGCTTTTCTTGAA
PHYB_Nicotiana	GAAGACAAGGATGACGGGCA	GAGAATGCATCCACGTTCTT	CCTTCAAGGCATTTCTGGAA
PHYB_Solanum	GAAGACAAGGATGACGGACA	GAGAATGCATCCACGTTCTT	CCTTCAAGGCATTTCTGGAA
PHYD_Arabidopsis	GAGGACAAAGATGATGGTCA	GCGGATGAATCCGCGTTCTT	CGTTCCAGACTTTTCTCGAA
PHYE_Arabidopsis	AAGGATAAAGATGATGCCGG	AAGAATGCATCCGAGGTCAT	CGTTTACAGCCTTTCTTGAA
PHYB_Oryza	GAGGATAAGGATGATGGACA	ACGAATGCATCCACGATCAT	CGTTCAAGGCATTTCTTGAA
PHYA_Avena	TCGGACATGGATGACAGCAG	AAGGATGCACCCTAGGTTGT	CTTTCAAAGCTTTCCTTGAA
PHYA_Oryza	TCAGACAAGGATGACAGCAG	AAGAATGCACCCTAGGCTGT	CCTTCAAGGCATTCCTTGAG
PHYA_Zea	TCTGATAAGGATGACAACAG	AAGGATGCACCCTAGGTTAT	CCTTTAAGGCTTTCCTTGAG
PHYA_Arabidopsis	GATGATAGGGATGATGCAAG	GAGAATGCACCCAAGGTCAT	CGTTCAAGGCTTTCCTTGAA
PHYA_Cucurbita	GGTCAAAAGGATGATGCCAG	AAAAATGCATCCAAGATCAT	CTTTCAAGGCTTTCCTTGAA
PHYA_Pisum	GGCGACCAAGACGATGGTAG	GAAGATGCATCCAAGATCAT	CATTCAAGGCTTTCCTTGAA
PHYA_Nicotiana	GGTGAAAAGGATGATGGCAG	GAAAATGCATCCTAGGTCAT	CATTCAAAGCATTCCTGGAA
PHYA_Solanum	GGTGAGAAGGATGATGGCAG	GAAAATGCATCCTAGGTCAT	CGTTCAAAGGATTCCTGGAA
PHYA_Arabidopsis	AATGACAGAGAT---GGTAA	GAGAATGCATCCTAGATCCT	CATTCAAGGCTTTTATGGAA

Physcomitrella	GTCGTGAAGCGACGAAGTCT	ACCCTGGGAAGATGTAGAAA	TGGATGCTATACATTCTCTT
Ceratodon	GTTGTGAACAAAAGAAGTCC	ACCCTGGGAAGACGTAGAAA	TGGATGCTATACATTCCCTT
Selaginella	GTGGTGAAGCGCCGAAGCCT	GCCGTGGGAGGACGTGGAGA	TGGATGCCATACACTCGCTG
Adiantum	GTGGTTAAGCAGCAGAGCTT	ACCATGGGAGGATGTGGAGA	TGGATGCCATTCACTCATTG
Psilotum	GTAGTAAAGCTAAGGAGTGT	GCCTTGGGAAGATGTGGAGC	TTGGTGCTGTTCATTCGTTA
Psilotum2	GTCGTAAAGTGGAGGAGTTC	TCCCTGGGAAGATGTGGAGA	TGGATGCTGTTCACTCGCTG
PHYB_Arabidopsis	GTTGTTAAGAGCCGGAGTCA	GCCATGGGAAACTGCGGAAA	TGGATGCGATTCACTCGCTC
PHYB_Nicotiana	GTTGTTAAAAGCCGGAGCTT	ACCATGGGAAAATGCAGAAA	TGGATGCAATTCACTCTCTG
PHYB_Solanum	GTTGTTAAAAGTCGGAGCTC	ACCATGGGAAAATGCCGAAA	TGGATGCAATCCACTCTTTG
PHYD_Arabidopsis	GTTGTTAAGAGCCGATGTCA	GCCATGGGAAACTGCTGAAA	TGGACGCCATTCACTCGCTC
PHYE_Arabidopsis	GTTGCGAAGAGCAGGAGCTT	GCCGTGGGAAATCTCAGAAA	TTGATGCTATCCATTCCCTG
PHYB_Oryza	GTTGTGAAGAGTAGGAGCTT	ACCATGGGAGAATGCAGAGA	TGGATGCAATACATTCCTTG
PHYA_Avena	GTTGTCAAGATGAAGAGCTT	GCCTTGGAGTGACTATGAAA	TGGATGCTATTCATTCATTG
PHYA_Oryza	GTTGTCAAGATGAAGAGCTT	GCCTTGGAATGACTATGAGA	TGGATGCTATTCACTCATTA
PHYA_Zea	GTTGTCAAGACGAAGAGTTT	GCCATGGAGTGACTACGAGA	TGGATGCTATTCACTCATTG
PHYA_Arabidopsis	GTGGTCAAGACAAGGAGTTT	ACCTTGGAAGGACTATGAGA	TGGATGCCATACACTCCTTG
PHYA_Cucurbita	GTAGTCAAGACAAGAAGTTT	GCCCTGGAAGGACTATGAGA	TGGATGCAATCCACTCTTTA
PHYA_Pisum	GTTGTGAAAGCCAGAAGCGT	GCCGTGGAAAGACTTTGAAA	TGGATGCTATTCATTCGTTG

```
PHYA_Nicotiana      GTTGTCAAGACTAGAAGTGT ACCATGGAAGGACTATGAAA TGGACGCAATCCATTCTTTG
PHYA_Solanum        GTTGTCAAGACAAGGAGTAT ACCCTGGAAGGACTATGAGA TGGATCGAATCCACTCCTTG
PHYA_Arabidopsis    ATAGTCAGGTGGAAAAGTGT GCCCTGGGATGACATGGAAA TGGATGCAATTAATTCTCTG

Physcomitrella      CAGCTCATTCTGCGTGGATC CTTTCAAGATATTGCTGAC- -----AGC---GATACGAAG
Ceratodon           CAGCTCATTCTACGTGGCTC CTTTCGAGATATTGCTGAC- -----AGC---GACACAAAG
Selaginella         CAGCTCATCCTCCGTGGCTC CTTCCAGGACATTGACGAC- -----AGC---GACACCAAG
Adiantum            CAGCTAATCCTTCGGGGCTC GTTTCAAGACATTGATGAT- -----AGT---AATACTAAG
Psilotum            CAGATTATACTACGTGATTC ACTTCTGGATATCGATGAC- -----ACG---GGGGCAAAG
Psilotum2           CAGCTCATATTACGTGGCTC GTTTCAAGATATCGATGAC- -----AGT---GAATCAAAA
PHYB_Arabidopsis    CAGCTTATTCTGAGAGACTC TTTTAAAGAATCTGAGGCG- -----GCTATGAACTCTAAA
PHYB_Nicotiana      ---CTTATTCTGCGAGATTC ATTTAAGGATGCCGAGGCA- -----AGT---AATTCTAAG
PHYB_Solanum        CAGCTAATTCTGCGAGATTC ATTTAAGGATGCTGAGGCA- -----AGT---AATTCTAAG
PHYD_Arabidopsis    CAGCTTATTCTAAGAGACTC TTTCAAAGAGTCTGAAGCGA TG---------GACTCTAAA
PHYE_Arabidopsis    AGACTTATAATGAGAGAGTC ATTTACCAGCTCTAGGCCAG TTTTGTCTGGTAATGGCGTA
PHYB_Oryza          CAGCTCATATTGCGGGACTC TTTCAGAGATTCTGCAGAGG GCACAAGT---AACTCAAAA
PHYA_Avena          CAACTTATACTGCGAGGGAC ACTAAATGATGCC------- --------------AGCAAG
PHYA_Oryza          CAACTTATACTTAGAGGGAC ACTGAATGATGACATC---- -----------------AAG
PHYA_Zea            CAGCTTATTCTTAGAGGTAC ACTGAATGATGCC------- --------------TCGAAG
PHYA_Arabidopsis    CAACTTATTTTGAGGAATGC TTTCAAGGATAGTGAAACTA CTGAT---------------
PHYA_Cucurbita      CAACTTATCCTTAGAAATAC TTTTAAGGATACAGATGCAA CTGAA---------------
PHYA_Pisum          CAGTTAATACTGAGAAATGC GTCCAAAGATACAGATATTA TAGAT---------------
PHYA_Nicotiana      CAGCTTATACTAAGAAATGC ATCCAAGGATGCCGATGCCA TGGAT---------------
PHYA_Solanum        CAGCTCATACTAAGAAATGC TTTCAAGGATGCTGATGCTG TGAAT---------------
PHYA_Arabidopsis    CAGCTAATAATAAAAGGCTC ATTGCAAGAGGAG------- --------------------

Physcomitrella      ACT------------ATGAT TCATGCACGCCTGAATGACT TGAAGCTTCATGACATGGAT
Ceratodon           ACC------------ATGAT CCACGCACGTCTGAATGACT TGAAGCTTCAGGGCGTGGAA
Selaginella         ACC------------ATGAT CCATGCGCGGCTCAACGACC TCAAGCTCCAGGGCATGGAC
Adiantum            ACC------------ATGAT TCATGCGCGTCTAAATGACC TGAAGCTTCAAGGTTTGGAT
Psilotum            TCT------------ACTAT CCAACATCAGATGAATGCAT TGGACCTGCATGGTGTGGAA
Psilotum2           ACT------------ATGAT CCACCATCGGCTGAACGATT TGAATCTGCATGGCATGGAT
PHYB_Arabidopsis    GTTGTGGATGGT---GTGGT TCAGCCATGTAGGGATATGG CGGGGGAACAGGGGATTGAT
PHYB_Nicotiana      GCT------------GTTGT GCATGCTCAGCTTGGGGAAA TGGAGTTGCAAGGGATAGAT
PHYB_Solanum        GCT------------ATTGT GCATGCTCATCTTGGGGAAA TGGAGTTGCAAGGGATAGAT
PHYD_Arabidopsis    GCTGCTGCAGCTGGGGCGGT TCAGCCACATGGAGATGATA TGGTACAGCAAGGGATGCAG
PHYE_Arabidopsis    GCAAGAGATGCTAAT----- -------------------- --------------------
PHYB_Oryza          GCCATAGTGAATGGCCAGGT TCAGCTTGGGGAGCTAGAAT TACGG------GGAATAGAT
PHYA_Avena          CCAAAGCGGGAAGCTAGTTT AGATAACCAGATTGGTGATC TAAAACTTGATGGGCTTGCT
PHYA_Oryza          CCAACAAGGGCCGCTAGTTT AGATAATCAGGTTGGTGATC TCAAGCTTGATGGGCTTGCT
PHYA_Zea            CCGGCCCAGGCATCTGGTTT AGATAACCAGATCGGTGATC TAAAACTTGATGGGCTTGCT
PHYA_Arabidopsis    ---GTGAATACAAAGGTCAT TTACTCGAAGCTAAATGATC TCAAAATTGATGGTATACAA
PHYA_Cucurbita      ---ATAAATAGAAAATCAAT TCAAACAACACTTGGTGACC TAAAAATTGAAGGGAGGCAA
PHYA_Pisum          ---TTGAATACGAAGGCAAT CAATACAAGACTAAATGATT TGAAGATTGAAGGAATGCAG
PHYA_Nicotiana      ---TCAAATACCAATATTAT CCACACGAAACTTAACGATC TTAAGATTGATGGGTTGCAG
PHYA_Solanum        ---TCAAATACAATTTCCAT CCATACGAAGCTTAATGATC TAAAGATTGATGGAATGCAG
PHYA_Arabidopsis    ---CATTCAAAGACTGTTGT GGATGTCCCACTTGTGGATA ATAGGGTTCAGAAGGTAGAT

Physcomitrella      GAGCTCAGCGTAGTGGCGAA TGAAATGGTACGCTTGATAG AAACTGCAACTGCTCCGATA
Ceratodon           GAACGAAACGCACTCGCTAA TGAGATGTCGCGCGTATTAG AAACCGCGGCTGCCCCAATC
Selaginella         GAGCTCAGCACCGTGGCCAA CGAGATGGTGCGCCTCATTG AGACCGCCACGGCACCCATC
Adiantum            GAGCTAAGCACTGTTGCAAG TGAAATGGTGCGACTTATTG AGACCGCGACCGCACCTATT
Psilotum            GAATTAAGCACAGTTGCCAA TGAGATGATGAGGCTTATTG AAACAGCCACAGTGCCTGTC
Psilotum2           GAATTGAGTACTGTTGCTAA TGAGATGGTGAGGCTCATGG AAACTGCAACGGCACCCATC
PHYB_Arabidopsis    GAGTTAGGTGCAGTTGCAAG AGAGATGGTTAGGCTCATTG AGACTGCAACTGTTCCTATA
PHYB_Nicotiana      GAACTGAGTTCTGTTGCCAG AGAAATGGTTAGATTGATAG AGACTGCAACTGCTCCCATA
PHYB_Solanum        GAACTGAGTTCTGTTGCCAG AGAAATGGTTAGATTGATCG AAACTGCAACAGCTCCCATA
PHYD_Arabidopsis    GAGATAGGTGCAGTTGCAAG AGAGATGGTTAGGCTCATTG AGACTGCGACGGTTCCTATA
PHYE_Arabidopsis    GAGCTTACTTCTTTTGTGTG TGAAATGGTTAGGGTGATTG AAACCGCAACTGCACCTATT
PHYB_Oryza          GAGCTTAGCTCGGTAGCGAG GGAGATGGTTCGGTTGATCG AGACAGCAACAGTACCCATC
PHYA_Avena          GAACTGCAGGCCGTGACCAG TGAAATGGTTCGTCTAATGG AAACAGCAACTGTTCCAATC
PHYA_Oryza          GAATTGCAGGCAGTTACAAG TGAAATGGTTCGTCTCATGG AAACAGCAACTGTCCCAATC
PHYA_Zea            GAATTGCAAGCAGTGACAAG TGAAATGGTCCGCCTGATGG AAACAGCAACTGTTCCGATC
PHYA_Arabidopsis    GAACTAGAAGCTGTGACCAG TGAGATGGTTCGTTTAATTG AGACTGCTACGGTGCCAATA
PHYA_Cucurbita      GAATTGGAATCAGTAACAAG TGAGATGGTCCGATTAATCG AGACAGCTACTGTGCCGATT
PHYA_Pisum          GAATTGGAAGCAGTGACAAG TGAGATGGTTAGATTAATTG AAACAGCAACAGTGCCTATT
PHYA_Nicotiana      GAGCTAGAAGCAGTGACTGC TGAAATGGTCCGCTTGATTG AAACAGCTTCAGTTCCTATC
PHYA_Solanum        GAACTTGAAGCAGTGACAGC TGAAATGGTCCGCTTAATTG AAACAGCTTCAGTTCCTATA
PHYA_Arabidopsis    GAATTGTGTGTTATCGTGAA TGAAATGGTGCGGTTGATTG ATACAGCAGCTGTTCCCATC

Physcomitrella      CTGGCAGTTGATTCAAACGG GATGATCAATGGCTGGAATG CAAAAATAGCCCAGGTAACA
Ceratodon           CTGGCGGTTGATTCAAGGGG AATGATTAATGCTTGGAATG CAAAAATAGCACAGGTCACA
Selaginella         CTTGCCGTAGACTCCAGTGG CTTCATAAACGGCTGGAACG CAAAGGTGGCGGACGTCACC
```

```
Adiantum            TTGGCTGTCGATGGGCAAGG GCTTATCAATGGATGGAATG GGAAGGTTGCTGAATTGACT
Psilotum            CTTGCTGTTAATTCCAGTGG GTTTATAAATGGGTGGAACG GCAAGGCAGTGGAATTGACA
Psilotum2           TTGGCTGTTGATTCCAGTGG TTTTATAAATGGATGGAATG GTAAGGCAGCAGAGTTGACA
PHYB_Arabidopsis    TTCGCTGTGGATGCCGGAGG CTGCATCAATGGATGGAACG CTAAGATTGCAGAGTTGACA
PHYB_Nicotiana      TTTGCTGTTGATGTCGAAGG TCGCATTAATGGGTGGAATG CAAAGGTCGCTGAATTGACA
PHYB_Solanum        TTTGCTGTTGATGTTGAAGG TCGCATAAATGGGTGGAATG CAAAGGTCGCTGAATTGACA
PHYD_Arabidopsis    TTTGCTGTGGACATAGACGG TTGCATCAATGGGTGGAACG CCAAGATCGCAGAGCTGACC
PHYE_Arabidopsis    TTTGGGGTTGATTCATCCGG ATGTATTAATGGTTGGAACA AGAAAACCGCTGAAATGACG
PHYB_Oryza          TTTGCAGTAGATACTGATGG ATGTATAAATGGTTGGAATG CAAAGGTTGCTGAGCTGACA
PHYA_Avena          TTGGCAGTAGATGGCAATGG ACTGGTCAACGGGTGGAATC AGAAAGCAGCGGAGTTGACT
PHYA_Oryza          TTGGCTGTAGATAGCAATGG ATTGGTCAATGGATGGAATC AGAAGGTTGCTGAGTTGACA
PHYA_Zea            TTGGCAGTAGATGGCAATGG ATTGGTCAATGGATGGAACC AAAAGGTAGCGGAGTTGTCA
PHYA_Arabidopsis    TTGGCGGTTGATTCTGATGG ACTGGTTAATGGTTGGAACA CGAAAATTGCTGAGCTGACT
PHYA_Cucurbita      TTAGCTGTTGATTTAGATGG GTTAATTAATGGGTGGAATA CAAAAATTGCTGAATTGACT
PHYA_Pisum          TTGGCAGTGGATGTTGATGG GACGGTCAACGGATGGAATA TAAAAATCGCCGAGTTGACA
PHYA_Nicotiana      TTCGCAGTTGATGTTGATGG GCAGCTTAATGGCTGGAACA CAAAAATTGCTGAATTAACC
PHYA_Solanum        TTCGCAGTTGATGTTGATGG GCAGGTTAATGGATGGAACA CTAAAGTTGCTGAATTAACT
PHYA_Arabidopsis    TTTGCGGTTGATGCCTCTGG TGTTATAAATGGTTGGAATT CTAAAGCGGCTGAGGTAACA

Physcomitrella      GGACTTCCCGTTAGTGAGGC CCATGGTAGATCCTTGGTGA AAGATCTCGTTACGGATGAG
Ceratodon           GGGCTTCCAGTCGAAGAGGC TATGCATTGTTCGTTGACGA AAGATCTCGTGTTGGATGAG
Selaginella         GGCCTGCCCGTGACCGAGGC CATGGGACGGTCCCTGGCCA AGGAACTGGTGCTCCACGAG
Adiantum            GGGCTTTCTTTTGAAACGGC GATGGGAAAATCACTAGCCA AAGAGCTTGTACACGAGGAG
Psilotum            GGTCTCAGTCTTGAAGATGC CATAAGCAAGTCCTTGCTGA AAGATCTTGTTGTGGATGAA
Psilotum2           GGTCTCCCACTCGAGGATGC TATGAACAAGTCTTTAGTGA GGGATCTTGTTGTGGATGAA
PHYB_Arabidopsis    GGTCTCTCAGTTGAAGAAGC TATGGGGAAGTCTCTGGTTT CTGATTTAATATACAAAGAG
PHYB_Nicotiana      GATTTATCTGTTGAAGAAGC AATGGGGAAGTCCTTGGTTC ATGATCTTGTGCATAAAGAG
PHYB_Solanum        GGTGTATCAGTTGAAGAAGC AATGGGGAAGTCCTTGGTTC ATGATCTTGTGTACAAAGAA
PHYD_Arabidopsis    GGTCTTTCTGTTGAAGACGC TATGGGAAAGTCGCTGGTTC GCGAATTGATATACAAAGAG
PHYE_Arabidopsis    GGATTACTAGCTAGTGAAGC AATGGGGAAGTCACTTGCCG ATGAGATTGTTCAAGAGGAA
PHYB_Oryza          GGCCTCTCTGTTGAGGAAGC AATGGGCAAATCATTGGTAA ATGATCTCATCTTCAAGGAA
PHYA_Avena          GGGCTAAGAGTTGATGATGC AATTGGAAGGCACATACTTA CC---CTTGTGGAGGACTCC
PHYA_Oryza          GGGTTGAGAGTAGATGAGGC TATTGGAAGACACATACTTA CC---GTTGTAGAGGAATCT
PHYA_Zea            GGGCTGAGAGTTGATGAGGC TATAGGAAGACACATACTTA CA---CTTGTGGAGGATTCT
PHYA_Arabidopsis    GGTCTTTCGGTTGATGAAGC AATCGGGAAGCATTTCCTCA CA---CTTGTTGAAGATTCT
PHYA_Cucurbita      GGACTGCCTGTGGATAAAGC TATCGGCAAGCATTTACTTA CG---TTAGTGGAAGATTCA
PHYA_Pisum          GGTCTTCCGGTTGGCGAAGC TATTGGAAAACATCTACTCA CC---CTGGTTGAGGATTCT
PHYA_Nicotiana      GGTCTTCCTGTTGATGAAGC AATTGGGAATCATTTGCTCA CA---CTCGTGGAGGATTCC
PHYA_Solanum        GGTCTTCCTGTTGATGAAGC AATTGGGAAACATCTGCTGA CT---CTCGTGGAGGATTCA
PHYA_Arabidopsis    GGATTGGCAGTTGAACAAGC AATAGGCAAACCTGTATCAG AT---CTCGTTGAGGACGAT

Physcomitrella      TCAGTGGCGGTTGTCGAGAG ATTACTTTATCTGGCACTGC GAGGTGAGGAAGAACAGAAT
Ceratodon           TCAGTGGTGGTTGTTGAGAG ATTACTTTCTCTGGCGTTGC AAGGTGAGGAGGAGCAGAAT
Selaginella         TCCGCCGACATGGTGGAGAG GCTGCTGTACCTGGCGCTTC AAGGTGACGAGGAGCAGAAC
Adiantum            TCAAAAACGATTGTGGAAAG AGTTCTCCATCTTGCATTGG AAGGTGAAGAAGAACAAGAT
Psilotum            TCAGTTGAGACAGTAAATAG TCTGTTGTGCTTAGCCTTGC AAGGGCAGGAGGAACAGAAT
Psilotum2           TCAGTCTCAACTGTGGAGAG GCTGTTGTACTTAGCATTGC AAGGTGAAGAAGAACGAAAT
PHYB_Arabidopsis    AATGAAGCAACTGTCAATAA GCTTCTTTCTCGTGCTTTGA GAGGGGACGAGGAAAAGAAT
PHYB_Nicotiana      TCACAGGAGACTGCTGAGAA GCTTCTCTTCAATGCTCTGA GAGGCGAAGAAGATAAAAAT
PHYB_Solanum        TCACAGGAGACTGCTGAGAA GCTTCTGTACAATGCTCTAA GAGGCGAGGAAGATAAAAAT
PHYD_Arabidopsis    TACAAAGAAACAGTTGATAG GCTTCTTTCTTGTGCTCTCA AAGGGGATGAAGGCAAGAAT
PHYE_Arabidopsis    TCACGCGCAGCTCTTGAAAG TCTCTTGTGCAAAGCCCTAC AAGGTGAAGAGGAGAAAAGT
PHYB_Oryza          TCTGAGGAAACAGTAAACAA GCTACTCTCACGAGCTTTAA GAGGTGATGAAGACAAAAAT
PHYA_Avena          TCTGTACCAGTTGTCCAGAG GATGCTATATCTAGCTCTGC AGGGTAAAGAAGAGAAGGAA
PHYA_Oryza          TCTGTACCAGTTGTCCAGAG GATGCTGTATTTAGCTTTGC AAGGCAAAGAAGAGAAGGAA
PHYA_Zea            TCTGTATCACTTGTTCAGAG GATGCTATATTTAGCTCTGC AAGGCAGAGAAGAGAAGGAA
PHYA_Arabidopsis    TCAGTGGAAATCGTTAAAAG GATGCTAGAGAACGCATTAG AAGGAACTGAGGAGCAGAAT
PHYA_Cucurbita      TCTGTAGAAGTTGTCAGGAA GATGTTGTTCTTGGCATTGC AAGGACAAGAAGAGCAAAAC
PHYA_Pisum          TCAACTGATATAGTCAAGAA GATGCTCAACTTGGCACTGC AGGGTGAAGAAGAGAAGAAT
PHYA_Nicotiana      TCAGTTGATACCGTGAGTAA AATGTTGGAATTAGCATTGC AAGGGAAAGAGGAAAGAAAT
PHYA_Solanum        TCAGTTGATACCGTGAATAA GATGTTGGAATTAGCATTGC AGGGGCAAGAGGAAAGAAAT
PHYA_Arabidopsis    TCTGTAGAAACCGTGAAGAA CATGTTAGCCTTGGCTCTCG AAGGTAGTGAAGAACGTGGT

Physcomitrella      GTGGAAATCAAGTTGAAGAC CTTTGGCACTCAGACTGAAA AAGGAGTG---------GTT
Ceratodon           GTGGAAATCAAGTTGAAGAC CTTTGGCACTCAGACCACTG AAAGAGCA---------GTT
Selaginella         GTTGAGCTCAAGCTCAAGAC TTTTGGTGGCCAAAAGGACA AGGAGGCG---------GTC
Adiantum            ATAGAAATACACTTGAGAAC TTATGACCAGCATAAGCAGA AAGGGGTA---------GTC
Psilotum            GTTGAAATAAGGTTAAAAAC TTTTGGTACGCATGCAGAAG AAGGGTCA---------GTG
Psilotum2           GTTGAAATCAAATTGAAAAC TTTTGGCACCCAGGCAGATA AAGGGCCA---------GTG
PHYB_Arabidopsis    GTGGAGGTTAAGCTGAAAAC TTTCAGCCCCGAACTACAAG GGAAAGCA---------GTT
PHYB_Nicotiana      GTAGAAATAAAGTTAAGGAC ATTTGGACCCGAGCAACTGA AGAAGGCT---------GTT
PHYB_Solanum        GTAGAAATAAAGTTGAGGAC ATTTGGAGCTGAACAACTGG AGAAAGCT---------GTT
PHYD_Arabidopsis    GTGGAGGTCAAGCTGAAAAC TTTTGGTTCCGAGCTACAAG GAAAAGCA---------ATG
```

```
PHYE_Arabidopsis   GTAATGCTGAAACTGAGAAA GTTTGGTCAAAACAATCATC CGGATTATTCTTCTGATGTG
PHYB_Oryza         GTAGAGATAAAGTTGAAGAC ATTCGGGCCAGAACAATCTA AAGGACCA---------ATA
PHYA_Avena         GTTCGATTTGAGGTAAAGAC TCATGGCCCGAAGAGGGATG ATGGTCCA---------GTT
PHYA_Oryza         GTGAAATTTGAGGTGAAAAC TCATGGCTCCAAGAGAGATG ATGGCCCT---------GTT
PHYA_Zea           GTTCGATTTGAGCTGAAAAC ACATGGCTCCAAGAGGGATG ATGGCCCT---------GTT
PHYA_Arabidopsis   GTCCAGTTTGAGATCAAGAC ACATCTGTCCAGGGCTGATG CTGGGCCA---------ATA
PHYA_Cucurbita     GTTCAATTCGAGATCAAGAC ACACGGTTCTCACATTGAGG TTGGCTCC---------ATC
PHYA_Pisum         GTTCAATTCGAGATAAAGAC ACATGGGGATCAGGTGGAAT CCGGTCCT---------ATT
PHYA_Nicotiana     GTAGAGTTTGAAATAAAAAC ACATGGGCCGTCAGGAGATT CTAGTCCA---------ATC
PHYA_Solanum       GTAGAGTTTGAAATAAAAAC ACATGGGCCGTCGAGGGATT CTAGTCCA---------ATC
PHYA_Arabidopsis   GCTGAGATCAGGATCAGAGC ATTTGGTCCTAAAAGGAAAA GCAGTCCG---------GTT

Physcomitrella     ATTCTGATTGTTGACGCCTG CTCGAGTATACACGTTTCGG AGAATGTTGTTGGTGTTTGC
Ceratodon          ATTCTGATTGTTAACGCCTG CTGCTCTAGAGATGCGTCAG ACTTTGTTGTTGGCGTTTTC
Selaginella        ATTCTGGTTGTCAACGCGTG CGCCAGCAGGGACGTGTCGG ACAACGTCGTTGGCGTGTGC
Adiantum           ATTCTAATTGTTAACACTTG CTGCAGCAGGGACGTTTCCA ACAATGTTGTGGGCGTTTGT
Psilotum           ATTTTGGTTGTCAATGCATG TGCAAGTAGAGATTCTACTG GTAATGTTGTTGGGGTTTGC
Psilotum2          ATTTTGGTTGTCAATGCATG TGCAAGTAGGAATATTACTG CAAATGTGGTTGGAGTTTGC
PHYB_Arabidopsis   TTTGTGGTTGTGAATGCTTG TTCCAGCAAGGACTACTTGA ACAACATTGTCGGCGTTTGT
PHYB_Nicotiana     TTTGTGGTGGTTAATGCTTG CTCTAGCAAAGATTACACAA ACAACATTGTTGGTGTTTGT
PHYB_Solanum       TTTGTGGTGGTTAATGCTTG CGCT---AGAGATTACACAA ACAACATTGTTGGTGTTTGC
PHYD_Arabidopsis   TTTGTGGTTGTCAACGCATG TTCAAGCAAGGACTACTTAA ACAACATCGTTGGAGTCTGC
PHYE_Arabidopsis   TGTGTTCTCGTTAACTCCTG CACGAGTCGGGATTATACTG AAAATATCATCGGTGTCTGC
PHYB_Oryza         TTCGTTATTGTGAATGCTTG TTCTACGAGGGATTACACTA AAAATATTGTTGGTGTTTGT
PHYA_Avena         ATCTTGGTTGTGAATGCTTG TGCCAGTCGGGACCTTCATG ATCATGTTGTTGGAGTGTGC
PHYA_Oryza         ATCTTGGTTGTGAATGCCTG TGCCAGCCGGGACCTTCACG ACCATGTTGTTGGTGTGTGC
PHYA_Zea           ATCTTGGTCGTAAATGCTTG TGCCAGTCGTGATCTTCATG ACCATGTTGTTGGGGTGTGC
PHYA_Arabidopsis   AGTTTAGTTGTAAATGCATG CGCAAGTAGAGATCTCCATG AAAACGTGGTTGGGGTGTGT
PHYA_Cucurbita     AGCCTAGTTGTAAATGCTTG TGCAAGTAGGGACTTGCGTG AAAATGTCGTGGGTGTGTTT
PHYA_Pisum         AGTTTGATAGTTAACGCATG TGCAAGCAAGGATCTTCGTG AAAATGTAGTGGGGGTTTGT
PHYA_Nicotiana     AGCTTAATTGTGAATGCATG TGCAAGCAGGGATGTTGGAG ATAGTGTTGTGGGTGTATGT
PHYA_Solanum       AGCTTAATCGTGAATGCATG TGCAAGCAAAGATGTTCGGG ATAGTGTTGTGGGTGTGTGT
PHYA_Arabidopsis   GAGTTAGTTGTCAACACTTG TTGTAGCAGAGATATGACGA ATAATGTTCTTGGTGTATGC

Physcomitrella     TTTGTGGGTCAGGATGTGAC TGGGCAAAAAATGTTCATGG ACAAGTTTACTCGTATACAG
Ceratodon          TTTGTGGGCCAAGATGTAAC TGAGCAGAGAATGTTCATGG ACAGGTTTACTCGTATACAA
Selaginella        TTCGTTGGGCAGGACGTGAC TGGGCAGAAGGTGGTCATGG ACAAGTTCACGCGCATCCAG
Adiantum           TTCGTGGGGCAGGATGTGAC GGGTCAAAAGTTGGTGCTGG ACAGGTTTATACGCATCCAA
Psilotum           TTTGTTGGGCAAGATGTAAC TCGTGAGAAGGTAGTAATGG ATAAGTTTACACGAATAATG
Psilotum2          TTTGTAGGTCAAGATGTAAC TGGTGAGAAGGTGGTAATGG ATAAGTTTACACGCATTCAG
PHYB_Arabidopsis   TTTGTTGGACAAGACGTTAC TAGTCAGAAAATCGTAATGG ATAAGTTCATCAACATACAA
PHYB_Nicotiana.    TTTGTTGGCCAGGATGTTAC TGGGCAAAAAGTTGTAATGG ACAAGTTTATTCACATCCAA
PHYB_Solanum       TTTGTTGGGCAGGATGTTAC TGGGGAAAAAGTTGTTATGG ACAAGTTTATTAACATCCAA
PHYD_Arabidopsis   TTTGTTGGACAAGATGTAAC TGGTCATAAAATTGTTATGG ACAAGTTCATCAACATACAA
PHYE_Arabidopsis   TTCGTTGGTCAAGACATCAC TAGTGAGAAAGCAATAACAG ATAGATTCATCAGATTGCAA
PHYB_Oryza         TTTGTTGGCCAAGATGTCAC AGGACAAAAGGTGGTCATGG ATAAATTTATCAACATACAA
PHYA_Avena         TTTGTTGCCCAAGATATGAC TGTCCATAAGTTGGTGATGG ACAAGTTTACTCGGGTTGAG
PHYA_Oryza         TTTGTTGCACAAGATATGAC TGTTCATAAGTTGGTCATGG ACAAATTTACTCGGGTTGAG
PHYA_Zea           TTTGTAGCCCAGGACATGAC TGTTCACAAGTTGGTCATGG ACAAATTTACTCGGGTCGAG
PHYA_Arabidopsis   TTTGTAGCCCATGATCTTAC TGGCCAGAAGACTGTGATGG ACAAGTTTACGCGGATTGAA
PHYA_Cucurbita     TTTGTAGCACAAGATATCAC TGGTCAGAAGATGGTAATGG ACAAGTTCACTCGATTAGAA
PHYA_Pisum         TTTGTGGCCCAAGATATAAC TGCTCAGAAGACTGTCATGG ACAAATTCACCCGAATCGAA
PHYA_Nicotiana     TTTATTGCTCAGGATATAAC AGGACAAAAAAATATTATGG ACAAGTTCACCCGGATTGAA
PHYA_Solanum       TTTATTGCTCAGGATATAAC TGGACAAAAAAGTATCATGG ACAAGTTCACCCGAATCGAA
PHYA_Arabidopsis   TTCATTGGACAAGATGTTAC AGGCCAGAAAACGCTTACTG AAAACTATAGCCGCGTGAAA

Physcomitrella     GGTGATTACAAGACTATAGT CCAGAATCCTCACCCCCTCA TTCCTCCTATTTTTGGTGCT
Ceratodon          GGTGGTGAAAAGACCACTGT CCAGGACCCCCACCCTCTCA TGCGTCCAAGTTTTGATGGT
Selaginella        GGCGATTACAAGGCTATTGT GCAGAATCCGAATCCGCTCA TACCGCCGATCTTCGGGGCG
Adiantum           GGGGACTATAAAGCAATTGT GCAGAGTCTGAATCCGCTTA TTCCCCCCATTTTTGGTGCA
Psilotum           GGTGATTATAGAACTATAGT ACAAAGTCCCCACCAACTTA TTCCCCCAATTTTTGGTTCA
Psilotum2          GGTGATTATAGGACTATAGT ACAAAGTCCTAACCCACTTA TTCCACCTATTTTTGGTTCA
PHYB_Arabidopsis   GGAGATTACAAGGCTATTGT ACATAGCCCAAACCCTCTAA TCCCGCCAATTTTTGCTGCT
PHYB_Nicotiana     GGTGATTACAAGGCCATTGT GCACAGCCCCAATCCTCTGA TCCCACCCATATTTGCGTCA
PHYB_Solanum       GGTGATTACAAGGCCATTGT GCACAGCCCCAATCCTCTGA TCCCTCCAATATTTGCATCA
PHYD_Arabidopsis   GGTGATTACAAGGCCATCAT CCATAGCCCGAACCCTCTGA TCCCTCCAATCTTTGCAGCG
PHYE_Arabidopsis   GGAGATTACAAGACTATTGT TCAAAGCTTAAATCCTTTGA TTCCACCGATATTTGCTTCA
PHYB_Oryza         GGGGATTACAAGGCTATCGT ACACAACCCTAATCCTCTCA TACCCCCAATATTTGCTTCA
PHYA_Avena         GGTGACTACAAGGCGATCAT TCACAACCCGAACCCACTCA TTCCTCCTATATTTGGTGCT
PHYA_Oryza         GGAGACTACAAAGCAATTAT TCACAATCCAAGCCCGCTTA TTCCTCCCATATTTGGTGCT
PHYA_Zea           GGGGACTACAAGGCAATCAT CCACAACCCGAACCCACTCA TTCCTCCTATATTTGGTGCT
PHYA_Arabidopsis   GGTGATTACAAGGCAATCAT CCAAAATCCAAACCCGCTGA TCCCGCCAATATTTGGTACC
PHYA_Cucurbita     GGCGATTACAAAGCCATTGT ACAAAATCCCAATCCATTGA TCCCTCCAATATTTGGATCA
```

```
PHYA_Pisum          GGCGATTACAAAGCAATTGT GCAGAACCCAAATCAGTTGA TCCCTCCTATATTCGGTACA
PHYA_Nicotiana      GGTGACTATAGAGCTATTAT CCAAAATCCTCACCCATTGA TCCCACCAATATTTGGCACT
PHYA_Solanum        GGTGACTATAGAGCGATTAT CCAAAATCCTCACCCATTGA TCCCACCAATATTTGGCACT
PHYA_Arabidopsis    GGAGATTATGCCCGAATCAT GTGGAGCCCTTCCACACTCA TTCCACCAATTTTTATAACC

Physcomitrella      GACGAATTTGGGTACTGTTT CGAGTGGAATCCAGCCATGG AGGGCTTGACTGGTTGGAAA
Ceratodon           GACGAATTTGGACGCACTTT CAAGAGGAACTCAGCATTGG GGGGTTTGAAAGATCATGCT
Selaginella         GACGAGTTTGGATACTGCTC TGAGTGGAATCCGGCCATGG AGAAGCTGTCTGGGTGGAGA
Adiantum            GATGAATATGGGTTTTGCTC AGAGTGGAATGCGGCCATGG AAAAGTTATCCAATTGGAGA
Psilotum            GATGGGTTTGGTTACTGCTC TGAGTGGAACGCAGCAATGG AAAAGTTATCTGGGTGGAAA
Psilotum2           GATGAATTTGGATACTGTTC TGAGTGGAATCCGGCAATGG TGAAGTTGTCTGGGTGGAAA
PHYB_Arabidopsis    GACGAGAACACGTGCTGCCT GGAATGGAACATGGCGATGG AAAAGCTTACGGGTTGGTCT
PHYB_Nicotiana      GATGAGAACACTTGTTGCTC TGAGTGGAACACTGCCATGG AAAAGCTCACTGGTTGGTCC
PHYB_Solanum        GATGAGAACACTTGTTGCTC CGAGTGGAACACTGCCATGG AAAAACTCACTGGTTGGTCT
PHYD_Arabidopsis    GATGAGAATACGTGCTGCCT TGAGTGGAACACTGCAATGG AAAAGCTCACAGGCTGGCCT
PHYE_Arabidopsis    GATGAAAATGCTTGTTGTTC TGAGTGGAACGCAGCAATGG AAAAGCTTACGGGATGGTCA
PHYB_Oryza          GATGAGAATACTTGTTGTTC GGAGTGGAACACAGCAATGG AAAAACTCACAGGATGGTCA
PHYA_Avena          GACGAATTTGGATGGTGTTC GGAGTGGAATGCTGCAATGA CCAAGTTGACTGGGTGGAAT
PHYA_Oryza          GACGAATTTGGATGGTGCTC TGAGTGGAATGCTGCCATGA CGAAATTGACCGGGTGGCAT
PHYA_Zea            GACCAGTTTGGATGGTGCTC TGAGTGGAATGCAGCCATGA CCAAGCTTACCGGGTGGCAC
PHYA_Arabidopsis    GATGAGTTTGGATGGTGCAC AGAGTGGAATCCAGCAATGT CAAAGTTAACCGGTTTGAAG
PHYA_Cucurbita      GATGAATTTGGATGGTGCTC AGAGTGGAATCCTGCAATGG CGAAACTAACTGGGTGGTCA
PHYA_Pisum          GATGAATTTGGCTGGTGTTG TGAGTGGAATGCAGCTATGA TTAAGTTAACCGGATGGAAG
PHYA_Nicotiana      GATCAATTCGGCTGGTGTTC TGAGTGGAACTCGGCAATGA CAAAGTTAACTGGATGGCGG
PHYA_Solanum        GATCAATTTGGCTGGTGTTC TGAATGGAACTCTGCGATGA CCATGTTAACTGGATGGCGG
PHYA_Arabidopsis    AATGAAAATGGGGTATGCTC AGAGTGGAACAACGCAATGC AGAAGCTCTCTGGGATAAAG

Physcomitrella      AAAGACGAGGTGGTCGGAAA ACTCCTTGTGGGAGAAATCT TCGGCATGCAAATGATGTGT
Ceratodon           ACCGGATCAGTGGAAAGAC- -------------------- --------------------
Selaginella         CGTGAGGAAGTACTAGGTAA GATGCTGGTGGGAGAAATAT TTGGCATCCAGATGATGTAC
Adiantum            AGGGAGGAGGTCTTGGGGAA AATGCTCGTGGGTGAGATAT TTGGATTGCAAATGGTTTGC
Psilotum            AGGGAAGAGGTACTTGGCAA AATGCTTGTAGGTGAGGTAT TTGGGACAGGCATTGCAATT
Psilotum2           ATGGAAGAGGTGCTTGGGAA AATGAATGTAGGGGAAATAT TTGGATCAGAGATGGCATGC
PHYB_Arabidopsis    CGCAGTGAAGTGATTGGGAA AATGATTGTCGGGGAAGTGT TTGGGAGC---------TGT
PHYB_Nicotiana      AGAGGGGAGATCATTGGAAA AATGTTAGTTGGTGAGATTT TTGGAAGT---------TGC
PHYB_Solanum        AGAGGGGAGATTGTTGGAAA AATGTTAGTTGGTGAGATTT TTGGAAGT---------TGT
PHYD_Arabidopsis    CGCAGCGAAGTGATTGGAAA ATTACTTGTTAGGGAAGTAT TTGGGAGC---------TAT
PHYE_Arabidopsis    AAACACGAGGTGATTGGGAA AATGCTACCCGGTGAAGTCT TTGGAGTC---------TTT
PHYB_Oryza          AGAGGGGAAGTTGTTGGTAA GCTTCTGGTCGGTGAGGTCT TTGGTAAT---------TGT
PHYA_Avena          AGAGATGAAGTGCTCGATAA GATGCTTCTTGGTGAAGTGT TTGACAGTAGCAATGCTTCC
PHYA_Oryza          AGAGATGAGGTGATCAATAA GATGCTTCTTGGTGAGGTGT TTGATAGCACCAACGCCTCC
PHYA_Zea            AGAGATGAAGTGGTTGATAA GATGCTCCTTGGCGAGGTTT TTAACAGCAGCAATGCTTCC
PHYA_Arabidopsis    CGAGAGGAAGTGATTGACAA AATGCTCTTAGGAGAAGTAT TTGGGACGCAGAAGTCATGT
PHYA_Cucurbita      CGTGAAGAAGTAATCGATAA GATGCTTTTGGGAGAGGTTT TTGGTGTTCACAAATCATGT
PHYA_Pisum          CGCGAGGAGGTGATGGATAA AATGCTTCTCGGAGAGGTTT TCGGAACTCAAATGTCTTGT
PHYA_Nicotiana      CGTGATGATGTTATTGATAA AATGCTATTAGGGGAGGTTT TTGGGACACAGGCAGCTTGT
PHYA_Solanum        CGTGATGATGTTATGGATAA AATGCTGCTAGGGGAGGTTT TTGGGACACAGGCAGCTTGC
PHYA_Arabidopsis    AGAGAAGAAGTTGTCAATAA AATTCTTCTCGGGGAGGTTT TTACCACAGATGATTATGGT

Physcomitrella      TGTCGGATGAAGAGCCAAGA TGCCATGACGAAATTCATGA TAGCATTGAACACAGCTATG
Ceratodon           -------------------- -------------------- --------------------
Selaginella         TGCCGCCTTAAGGGCCAGGA TGCCGTCACCAAGTTCATGA TCGTCCTCAACAGTGCAGCA
Adiantum            TGCCGCTTGCAGGGGCAAGA TGTCGTTACCAAATTGATGA TAGTGCTGAATGATGCTGTC
Psilotum            TGCCGCCTCCGTGGTCAGGA TTCTTTGACAAAATTCATGA TAGTTCTCAATAGTGCAATG
Psilotum2           TGCCGCCTTCGGGGCCAGGA TGCCATGACAAAGTTTATGA TAGTTCTCAACAGTGCAATG
PHYB_Arabidopsis    TGCATGCTAAAGGGTCCTGA TGCTTTAACCAAGTTCATGA TTGTATTGCATAATGCGATT
PHYB_Nicotiana      TGTCGGCTCAAGGGTCCAGA CGCCATGACAAAGTTCATGA TCGTGTTGCATAATGCGATT
PHYB_Solanum        TGTCGCCTCAAGGGCCCAGA TGCCATGACAAAGTTCATGA TCGTGTTGCATAATGCAATT
PHYD_Arabidopsis    TGCAGACTAAAGGGTCCTGA TGCGTTAACTAAGTTCATGA TCGTCTTGCATAACGCGATC
PHYE_Arabidopsis    TGTAAAGTGAAATGCCAAGA TTCGCTCACAAAGTTCTTGA TCTCTCTGTACCAAGGAATT
PHYB_Oryza          TGTCGACTCAAGGGCCCAGA TGCATTAACGAAATTCATGA TTGTCCTACACAACGCTATA
PHYA_Avena          TGCCCTTTGAAGAACAGAGA TGCATTTGTAAGTCTTTGTG TTCTTATCAACAGTGCATTA
PHYA_Oryza          TGTCTTGTGAAGAATAAAGA TGCATTTGTAAGTCTCTGCA TTCTTATCAACAGTGCATTA
PHYA_Zea            TGCCTTCTGAAGAGCAAAGA TGCCTTTGTACGTCTTTGCA TTGTCATCAACAGCGCATTA
PHYA_Arabidopsis    TGTCGTCTAAAGAATCAAGA AGCCTTTGTAAACCTTGGGA TTGTGCTGAACAATGCTGTG
PHYA_Cucurbita      TGTCGTTTAAAGAATCAAGA AGCTTTTGTTAATCTTGGGA TTGTCTTGAACAATGCCATG
PHYA_Pisum          TGTCGTCTAAAGAATCAAGA AGCTTTTGTTAATTTCGGCA TTGTACTTAATAAAGCCATG
PHYA_Nicotiana      TGCCGTCTCAAGAATCAAGA AGCTTTTGTAAATTTTGGAG TTGTACTAAACAATGCTATG
PHYA_Solanum        TGCCGTCTCAAGAATCAAGA AGCTTTTGTGAATTTCGGAG TTATACTAAACAATGCTATC
PHYA_Arabidopsis    TGCTGCCTTAAAGACCATGA CACTTTAACGAAGCTGAGAA TAGGTTTCAATGCTGTGATT

Physcomitrella      GACGGCCAA---AGTACAGA TAAATTTACTTTCTCCTTTT TCGATCGAGAAGGGAAATAC
Ceratodon           -------------------- -------------------- --------------------
```

```
Selaginella        GATGGCCAG---GACACTGA AAAGTTCCCGTTTGCGTTCT TCGATCGACAGGGAAAGTAT
Adiantum           AATGGCCAA---GAATCGGA AAAATTTCCTCTAGTGTTTT ATGACCGCAATGGGCGCCGT
Psilotum           GGTGGACAG---GATTCTGA TGGATTTCCGTTTGGATTTT TTGATAGAGAAGGGAAGTAT
Psilotum2          GGTGGTCAG---GACTCAGA TAGGTTCCCATTGGGCTTTT TTGATCGACAAGGGAAGTAT
PHYB_Arabidopsis   GGTGGCCAA---GATACGGA TAAGTTCCCTTTCCCATTCT TTGACCGCAATGGGAAGTTT
PHYB_Nicotiana     GGAGTCCAG---GATACGGA CAAGTTTCCATTTTCCTTTT TTGACCGAAATGGGAAATAT
PHYB_Solanum       GGAGGCCAG---GATACAGA CAAGTTTCCATTTTCCTTTT TTGACCGAAATGGGAAATAT
PHYD_Arabidopsis   GGTGGCCAA---GATACTGA TAAATTCCCATTCCCGTTCT TTGATCGCAAAGGGGAATTC
PHYE_Arabidopsis   GCTGGTGATAATGTTCCCGA GAGTTCACTGGTTGAGTTCT TTAATAAGGAAGGGAAGTAC
PHYB_Oryza         GGAGGACAG---GATTGTGA AAAGTTCCCCTTTTCATTTT TTGACAAGAATGGGAAATAC
PHYA_Avena         GCCGGGGAA---GAAACAGA AAAGGCTCCATTTGGCTTCT TCGACAGAAGTGGAAAGTAC
PHYA_Oryza         GCTGGTGAT---GAAACAGA AAAGGCTCCATTCAGCTTCT TCGACCGGAACGGGAAGTAT
PHYA_Zea           GCTGGTGAA---GAGGCAGA AAAGGCTTCATTCGGCTTCT TTGACCGCAATGAGAAATAT
PHYA_Arabidopsis   ACCAGTCAA---GATCCAGA TAAAGTATCGTTTGCTTTCT TTACAAGAGGTGGCAAGTAT
PHYA_Cucurbita     TGTGGTCAA---GATCCTGA AAAGGCTTCCTTTGGTTTCT TAGCTCGGAACGGGATGTAC
PHYA_Pisum         ACCGGTTTG---GAAACAGA AAAGGTCCCTTTTGGCTTCT TCTCTCGAAAAGGCAAGTAT
PHYA_Nicotiana     ACTGGTCAA---GAGTGTGC GAAGATATCTTTTGGTTTCT TTGCACGTAATGGGAAATAC
PHYA_Solanum       ACTGGTCAA---GAATCTGA GAAGATTCCTTTTGGTTTCT TTGCACGTTACGGGAAATAT
PHYA_Arabidopsis   TCTGGCCAAAAGAACATnGA GAAGCTTTTATTTGGCTTTT ACCATCGTGATGGTAGCTTC

Physcomitrella     GTTGACGTTCTACTCTCCAC CAATAAACGTACCAATGCTG ATGGAGTCATAACTGGCGTG
Ceratodon          -------------------- -------------------- --------------------
Selaginella        GTTGAGGCGTTACTCACTGC CACCAAGCGGGCAGATGCCG AGGGCTCCATCACTGGAGTG
Adiantum           GTCGAAGCTCTACTCATCGC AAGCAAGAGAACTGATGCTG ACGGGCGAATCACCGGTGTC
Psilotum           GTGCAGGCCTTGCTTATTGC AAACAAGAGAACACATGGTG ATGGAGCTATCACAGGTGTT
Psilotum2          GTAGAGGCCCTTCTAATTGC AAACAAGAGAACAGACGGTG CAGGGGCTATCACAGGTGTT
PHYB_Arabidopsis   GTTCAGGCTCTATTGACTGC AAACAAGCGGGTTAGCCTCG AGGGAAAGGTTATTGGGGCT
PHYB_Nicotiana     GTGCAAGCTCTTTTGACTGC GAACAAGAGAGTCAATATGG AGGGCCAGATTATCGGGGCT
PHYB_Solanum       GTGCAAGCTCTTTTGACTCG AAACAAGAGAGTTAATATGG AGGGCGATACTATTGGGGCT
PHYD_Arabidopsis   ATTCAGGCTCTCCTGACTTT GAACAAACGGGTCAGCATCG ATGGCAAAATCATTGGGGCT
PHYE_Arabidopsis   ATAGAAGCATCCTTAACCGC GAACAAGAGTACAAACATCG AAGGAAAAGTTATAAGATGT
PHYB_Oryza         GTGCAGGCCTTATTGACTGC AAACACGAGGAGCAGAATGG ATGGTGAGGCCATAGGAGCC
PHYA_Avena         ATTGAGTGTCTTCTATCAGC AAACAGAAAAGAAAATGAGG GTGGTCTCATCACTGGAGTA
PHYA_Oryza         ATCGAGTGCCTTCTTTCTGT TAACAGAAAAGTAAATGCAG ATGGTGTCATCACTGGAGTA
PHYA_Zea           GTCGAATGCCTTCTGTCGGT GAACAGGAAAGTAAATGCAG ATGGTGTTGTCACTGGAGTG
PHYA_Arabidopsis   GTGGAGTGTCTGTTGTGTGT GAGTAAGAAACTGGACAGGA AAGGTGTAGTGACAGGTGTC
PHYA_Cucurbita     GTGGAATGTCTTCTATGTGT CAATAAGATCTTGGATAAAG ATGGCGCGGTTACAGGGTTC
PHYA_Pisum         GTAGAGTGCCTACTCTCGGT GAGTAAGAAAATCGACGCAG AGGGCCTAGTTACCGGAGTC
PHYA_Nicotiana     GTAGAGTGCTTACTTTGTGT GAGCAAAAGGTTGGATAGAG AGGGTGCAGTCACGGGACTC
PHYA_Solanum       GTGGAGTGCTTACTCTGTGT GAGCAAAAGGTTAGATAAAG AGGGTGCAGTCACGGGACTC
PHYA_Arabidopsis   ATCGAGGCATTGCTTTCTGC AAACAAAAGGACTGATATTG AGGGAAAGGTTACCGGGGTT

Physcomitrella     TTTTGCTTCTTACAAATTGC AAGTTCGGAGCTACAACAGG CTCTCAAAGTACAACGGGCT
Ceratodon          -------------------- -------------------- --------------------
Selaginella        TTTTGCTTCCTACACATTGC CAGTGCGGAGTTGCAACAGG CCTTGACAGTCCAGAGAGCC
Adiantum           TTCTGCTTTCTTCACACTGC AAGCCCCGAGCTGCTTCAGG CGTTGATCATAAAGAGAGCC
Psilotum           TTCTGTTTTCTGCACATTGC CAGTGCAGAAGTGCAGCAAG CTTTACAGGTGCAAAAACAG
Psilotum2          TTCTGTTTTTTGCACATTGC CAGTGCAGAAGTGCAACAAG CATTGCAGGTACAAAAAAGG
PHYB_Arabidopsis   TTCTGTTTCTTGCAAATCCC GAGCCCTGAGCTGCAGCAAG CTTTAGCAGTCCAACGGAGG
PHYB_Nicotiana     TTCTGTTTCATACAGATAGC CAGTCCTGAATTGCAGCAAG CTCTAAGAGTTCAAAGGCAA
PHYB_Solanum       TTCTGTTTCATACAGATAGC CAGTCCCGAACTGCAGCAAG CTCTAAGAGTTCAAAGGCAA
PHYD_Arabidopsis   TTCTGTTTTTTGCAGATACC GAGTCCCGAGCTGCAGCAAG CTCTAGAAGTTCAGAGGAGG
PHYE_Arabidopsis   TTCTTCTTCTTGCAGATT-- ----ATCAATAAAGAATCGG GGTTGAGCTGCCCAGAACTG
PHYB_Oryza         TTCTGTTTCTTGCAGATTGC AAGTCCTGAATTACAGCAAG CCTTTGAGATTCAGAGACAC
PHYA_Avena         TTCTGTTTTATTCATGTTGC TAGTCATGAGCTGCAACATG CACTACAGGTGCAGCAAGCC
PHYA_Oryza         TTTTGTTTCATTCAAGTTCC TAGTCATGAGCTGCAACATG CACTACATGTGCAGCAAGCC
PHYA_Zea           TTCTGTTTCATCCATGTTCC TAGTGATGACCTGCAGCATG CGCTACATGTGCAGCAAGCC
PHYA_Arabidopsis   TTCTGTTTCCTGCAACTTGC CAGCCATGAGCTGCAGCAAG CGCTCCATGTTCAACGTTTA
PHYA_Cucurbita     TTTTGCTTTTTGCAGCTTCC TAGTCATGAGTTGCAACAAG CACTAAATATCCAACGCTTA
PHYA_Pisum         TTCTGTTTCTTGCAGTTAGC TAGCCCTGAGCTGCAACAAG CATTACATATTCAGCGCCTG
PHYA_Nicotiana     TTTTGTTTCCTGCAGCTTGC AAGCCATGAGCTGCAACAAG CTCTTCACATTCAACGATTA
PHYA_Solanum       TTTTGTTTCCTGCAGCTTGC AAGCCATGAGTTGCAACAAG CTCTTCATGTTCAACGATTA
PHYA_Arabidopsis   TTATGCTTTTTGCAAGTACC TAGTCCAGAACTCCAATATG CTCTACAGGTTCAGCAAATA

Physcomitrella     ACGGAGAAGGTCGCAGTAGC TAAGCTGAAGGAACTTGCAT ATATCGTCAGGGAGATTAAA
Ceratodon          -------------------- -------------------- --------------------
Selaginella        ACTGAGAAGGTTGCATTATC CAAGCTCAAGGAGCTGGCTT ACATCCGCCAGGAGATAAAG
Adiantum           AAGGAAAAGGTT-------- ----GACAAAGAGTTGAGTT ATGTGAAGGAGGAGTTGAAG
Psilotum           TCAGCAAGGGCAGCTTTGGA TAGACTCAAGGAGGTAGCGT ACATGCGCCAGGAGATTGGG
Psilotum2          TCAGCAAGGACAGCTTTAGA CAAACTCAAGGAGGTAGCAT ACATGAGGCAGGAAATTAGG
PHYB_Arabidopsis   CAGGACACAGAGTGTTTCAC GAAGGCAAAAGAGTTGGCTT ATATTTGTCAGGTGATAAAG
PHYB_Nicotiana     CAGGAAAAGAAGTGTTATTC TCAGATGAAAGAGTTGGCAT ACCTTTGTCAGGAAATAAAG
PHYB_Solanum       CAGGAAAAGAAGTGCTATTC TCAGATGAAAGAGCTGGCAT ACATTTGTCAGGAAATAAAA
```

```
PHYD_Arabidopsis   CAGGAGAGTGAATATTTCTC AAGGAGGAAAGAGTTGGCTT ACATTTTCCAAGTTATAAAG
PHYE_Arabidopsis   AAAGAGAGC---GCTCAAAG C---CTCAACGAATTAACTT ACGTAAGACAAGAAATCAAG
PHYB_Oryza         CATGAAAAGAAGTGTTATGC AAGGATGAAGGAATTGGCTT ACATTTACCAGGAAATAAAG
PHYA_Avena         TCGGAGCAAACGTCGCTAAA AAGGCTCAAGGCTTTCTCCT ACATGAGACATGCGATCAAC
PHYA_Oryza         TCACAGCAGAATGCACTAAC AAAGTTGAAAGCTTACTCCT ACATGAGACATGCAATCAAC
PHYA_Zea           TCTGAGCAGACAGCACAAAG AAAGTTGAAGGCTTTCTCGT ACATGCGACATGCCATCAAC
PHYA_Arabidopsis   GCTGAGCGAACCGCAGTGAA GAGACTAAAGGCTCTAGCAT ACATAAAAAGACAGATCAGG
PHYA_Cucurbita     TGTGAGCAAACTGCATTGAA GAGATTGAGAGCATTGGGAT ACATAAAAAGACAGATACAA
PHYA_Pisum         TCCGAACAAACCGCTCTCAA GAGACTGAAAGTACTAACTT ACATGAAAAGGCAGATCAGG
PHYA_Nicotiana     TCAGAACAAACTGCATTGAA GAGGTTGAAAGTATTAGCTT ACATAAGGAGGCAGATTAGA
PHYA_Solanum       TCAGAACAAACTGCATTGAA AAGGTTGAAAGTATTAGCTT ACATAAGGAGGCAAATTAGG
PHYA_Arabidopsis   TCAGAGCATGCAATTGCCTG TGCCCTCAACAAATTGGCAT ATCTCCGCCATGAAGTGAAG

Physcomitrella     AATCCACTATGTGGTCTTAC ATTCACACGACAATTGTTAG AAGATACTGATCTATCAGAC
Ceratodon          -------------------- -------------------- --------------------
Selaginella        AATCCGCTCTATGGTATTAT GTTCACCCGGACTCTCATGG AAACTACCGACCTGTCCGAG
Adiantum           AAGCCTTTGGAAGGACTTGC ATTCACACGGACGGTGCTAG AAGGTACAAATCTAACTATA
Psilotum           AACCCCTTGTCTGGTATTGC CTTTACCCGTAACTTGTTGG GGGCTACAAATCTTTCTGAG
Psilotum2          AACCCTTTGTATGGCATTAT GTTTACTCGCAGATTGCTGG AAGGTACAAATCTTTCAGAG
PHYB_Arabidopsis   AATCCTTTGAGCGGTATGCG TTTCGCAAACTCATTGTTGG AGGCCACAGACTTGAACGAG
PHYB_Nicotiana     AGTCCTTTGAATGGTATACG CTTTACAAATTCATTGTTGG AAGCGACAGATTTGACAGAA
PHYB_Solanum       AGTCCTCTTAATGGTATACG CTTTACAAATTCATTGTTGG AGGCCACAAATTTGACAGAA
PHYD_Arabidopsis   AATCCATTGAGTGGATTGCG TTTCACAAATTCATTGCTGG AAGACATGGATTTAAACGAG
PHYE_Arabidopsis   AATCCTCTCAACGGTATCCG ATTTGCACATAAGCTTCTTG AATCCTCAGAGATTTCAGCT
PHYB_Oryza         AATCCTCTCAACGGTATCCG ATTTACAAACTCGTTATTGG AGATGACTGATCTAAAGGAT
PHYA_Avena         AACCCTCTCTCAGGCATGCT CTACTCTAGAAAAGCATTGA AGAACACAGATTTGAATGAA
PHYA_Oryza         AACCCTCTCTCAGGTATGCT TTACTCTAGGAAAGCACTGA AGAACACAGGTCTGAATGAA
PHYA_Zea           AAACCTCTCTCAGGTATGCT TTATTCTAGGGAAACACTCA AGAGCACAGGTCTGAATGAA
PHYA_Arabidopsis   AATCCGCTATCTGGGATCAT GTTTACAAGGAAAATGATAG AGGGTACTGAATTAGGACCA
PHYA_Cucurbita     AATCCTCTTTCTGGGATAAT CTTTTCACGAAGATTATTGG AGCGCACCGAGTTGGGAGTA
PHYA_Pisum         AATCCGTTGGCTGGGATCGT GTTTTCCAGTAAAATGCTGG AGGGTACTGACTTGGAAACC
PHYA_Nicotiana     AACCCTCTTTCTGGAATTAT ATTCTCTCGGAAAATGCTGG AGGGGACTAACTTGGGCGAA
PHYA_Solanum       AATCCTCTTTCTGGGATTAT ATTCTCTCGGAAGATGCTAG AGGGGACTAGCTTGGGCGAA
PHYA_Arabidopsis   GACCCCGAAAAGGCAATATC CTTCCTTCAAGATTTGCTCC ATTCATCTGGATTAAGTGAA

Physcomitrella     GATCAGCAGCAGTTTCTAGA CACGAGTGCCGTATGTGAGC AGCAGTTACAAAAAAGTCTG
Ceratodon          -------------------- -------------------- --------------------
Selaginella        GACCAGAAGCAGTATGTTGA GACTGGAGCCGTGTGCGAGA AGCAAATCCGCAAGATCCTG
Adiantum           GAGCAGAGGCAGCTCATCAA GACAAATGCTTGGTGCGAAA GGCAGCTGCGAAAGATATTG
Psilotum           GAGCAGAAGAGGATTGTTGA AACTAGTGTTGCTTGTGAAA GGCAGCTGGAACAGATCTTG
Psilotum2          GAACAGAAACAGATCATTGA TACAAGTGCTGTCTGTGAAA AGCAGCTGCATCAGATCTTG
PHYB_Arabidopsis   GACCAGAAGCAGTTACTTGA AACAAGTGTTTCTTGCGAGA AACAGATCTCAAGGATCGTC
PHYB_Nicotiana     AACCAGAAGCAGTATCTGGA GACAAGTGCTGCTTGTGAGA GGCAGATGTCTAAGATCATA
PHYB_Solanum       AATCAGAAGCAGTATCTAGA GACAAGTGCTGCTTGTGAGA GGCAGATGTCTAAGATCATT
PHYD_Arabidopsis   GATCAGAAGCAGCTTCTTGA AACGAGTGTTTCATGTGAGA AGCAGATCTCAAAGATTGTA
PHYE_Arabidopsis   AGCCAAAGGCAGTTTCTGGA GACTAGTGATGCTTGTGAGA AGCAAATCACGACAATAATC
PHYB_Oryza         GACCAGAGGCAGTTTCTTGA AACCAGCACTGCTTGTGAGA AACAGATGTCCAAAATTGTT
PHYA_Avena         GAACAGATGAAGCAGATTCA TGTTGGAGATAATTGTCACC ACCAGATAAACAAGATACTT
PHYA_Oryza         GAGCAGATGAAGGAGGTCAA TGTTGCAGATAGTTGTCACC GCCAGCTGAATAAAATACTT
PHYA_Zea           GAGCAGATGAGGCAGGTTCG CGTCGGAGACAATTGCCATC GCCAGCTAAACAAGATACTT
PHYA_Arabidopsis   GAGCAAAGACGGATTTTGCA AACTAGCGCGTTATGTCAGA AGCAACTAAGCAAGATCCTC
PHYA_Cucurbita     GAACAAAAAGAACTTCTGCG TACTAGCGGACTCTGTCAAA AGCAGATCTCCAAGGTTCTC
PHYA_Pisum         GAACAAAAGCGAATCGTGAA CACTAGTTCTCAGTGCCAAC GCCAGCTTAGCAAAATTCTT
PHYA_Nicotiana     GAGCAGAAAAATATACTGCG TACTAGTTCCCAGTGTCAGC GTCAGCTCAACAAAATTCTT
PHYA_Solanum       GAGCAGAAAAATATACTGCA TACAAGTGCACAGTGTCAGC GTCAGCTCGACAAAATTCTT
PHYA_Arabidopsis   GACCAAAAGCGGCTCCTGAG GACAAGCGTTTTATGCAGGG AGCAGTTAGCCAAAGTCATA

Physcomitrella     AATGACATGGATTTGGAGAG TATCGAAGACGGG------T ACTTGGAGTTGGATACTGCC
Ceratodon          -------------------- -------------------- --------------------
Selaginella        GACGACATGGATCTGGAGAG CATCGAAGACGGG------T ACCTGGAGCTGGATACAACC
Adiantum           ---GAAGATGACCTGAACAA TATAGAAGAAGGG------T ACATGGATTTGGAGATGTCC
Psilotum           AATGAAGATAACTTTGGAAA TCTTTAC------------- --------------------
Psilotum2          GATGAAGATAACTTCGAAAA TCTTGACCATGGA------A ACATTGATTTGGATACCATT
PHYB_Arabidopsis   GGGGACATGGATCTTGAAAG CATTGAAGACGGT------T CATTTGTGCTAAAGAGGGAA
PHYB_Nicotiana     AGGGATGTTGACCTGGAAAA CATTGAGGATGGC------T CACTGACCCTTGAGAAAGAA
PHYB_Solanum       AGGGATATTGATCTGGAAAA CATTGAGGACGGG------T CACTGACCCTTGAGAAAGAA
PHYD_Arabidopsis   GGAGACATGGACGTCAAAAG CATAGATGACGGT------T CATTTCTGCTAGAGAGAACA
PHYE_Arabidopsis   GAAAGCACGGACTTGAAAAG CATTGAGGAAGGC------A AGTTGCAATTGGAAACAGAA
PHYB_Oryza         AAGGATGCTAGCCTCCAAAG TATTGAGGATGGC------T CTTTGGTGCTTGAGAAAGGT
PHYA_Avena         GCAGACTTGGATCAAGATAG CATCACCGAAAAATCTAGCT GCTTGGATTTGGAGATGGCT
PHYA_Oryza         TCTGACTTGGATCAAGATAG CGTCATGAACAAGTCTAGTT GCTTGGATTTGGAGATGGTT
PHYA_Zea           GCCGACTTGGATCAAGATAA CATCACTGACAAGTCAAGCT GCTTGGATTTGGATATGGCT
PHYA_Arabidopsis   GATGATTCGGATCTTGAAAG CATCATTGAAGGA------T GCTTGGATTTGGAAATGAAA
```

```
PHYA_Cucurbita      GATGAGTCAGACATCGATAA AATTATTGATGGG------T TTATTGATTTAGAAATGGAC
PHYA_Pisum          GACGACTCTGATCTCGACGG CATCATTGATGGG------T ACTTGGATCTTGAGATGGCT
PHYA_Nicotiana      GATGATACAGATCTTGATAG CATCATTGATGGT------T ATTTGGATCTGGAGATGCTC
PHYA_Solanum        GATGATACAGATCTTGATAG CATCATAGAGGGT------T ATTTGGATCTGGAGATGCTC
PHYA_Arabidopsis    AGCGACTCAGACATAGAGGG AATCGAAGAAGGC------T ATGTGGAACTGGATTGCAGC

Physcomitrella      GAGTTCGAGATGGGCACAGT GATGAACGCCGTCATCAGTC AAGGTATGACAACATCCCGT
Ceratodon           -------------------- -------------------- --------------------
Selaginella         GAGTTCATGATGGGAACGGT TATGGACGCTGTGATAAGTC AGGGCATGATCACGTCCAAA
Adiantum            GAATTTTTCATGGGATCAGT TATTGATGCTGTGATAAGTC AAGGAATGGCTGCCTCAAGA
Psilotum            -------------------- -------------------- --------------------
Psilotum2           GAGTTTACCATGGGAACGGT GATGGATGCTGTCATCAGTC AAGGGATGATCAGATCCAGA
PHYB_Arabidopsis    GAGTTTTTCCTTGGAAGTGT CATAAACGCGATTGTAAGTC AAGCGATGTTCTTATTAAGG
PHYB_Nicotiana      GAATTTTTCCTTGGGAGTGT AATAGATGCTGTTGTTAGCC AAGTGATGTTATTGCTGAGG
PHYB_Solanum        GATTTTTTTCTTGGGAGTGT AATAGATGCTGTTGTTAGCC AAGTGATGTTATTGCTGAGG
PHYD_Arabidopsis    GAGTTCTTCATTGGCAATGT CACAAATGCAGTGGTAAGCC AAGTCATGTTGGTGGTGAGA
PHYE_Arabidopsis    GAGTTTCGACTTGAAAACAT CTTGGACACAATCATTAGCC AAGTGATGATTATATTGAGA
PHYB_Oryza          GAATTTTCACTAGGTAGTGT TATGAATGCTGTTGTCAGCC AAGTGATGATACAGTTGAGA
PHYA_Avena          GAATTTCTGTTGCAAGATGT GGTGGTGGCTGCTGTAAGTC AAGTACTGATAACCTGCCAG
PHYA_Oryza          GAGTTTGTATTGCAAGATGT GTTTGTGGCTGCTGTAAGTC AAGTACTCATAACTTGCCAG
PHYA_Zea            GAATTTGTGTTGCAAGATGT GGTGGTGTCTGCTGTAAGTC AAGTGCTGATAGGTTGCCAG
PHYA_Arabidopsis    GAATTCACCTTAAATGAAGT GTTGACTGCTTCCACAAGTC AAGTAATGATGAAGAGTAAC
PHYA_Cucurbita      GAGTTTACATTGCATGAAGT ATTGATGGTATCAATTAGTC AAGTGATGCTAAAGATTAAA
PHYA_Pisum          GAATTCACTTTACATGAGGT ACTGGTTACCTCTCTTAGTC AAGTCATGAATAGGAGTAAC
PHYA_Nicotiana      GAGTTCAAGCTGCACGAAGT ATTAGTGGCATCTATTAGTC AAATCATGATGAAGAGCAAT
PHYA_Solanum        GAGTTTAAGCTACATGAGGT ATTGGTAGCATCCATAAGTC AAGTCATGATGAAGAGCAAT
PHYA_Arabidopsis    GAATTCGGCCTGCAGGAATC CCTGGAAGCAGTTGTAAAAC AAGTGATGGAGCTGAGCATA

Physcomitrella      GAGAAGGGGCTTCAGATTTT TCGAGAGACACCTCGAGAAA TAAATACAATGCGTCTGCTT
Ceratodon           -------------------- -------------------- --------------------
Selaginella         GAGAAGAACTTGCAGCTCAT CCGGGAAACTCCAAAGGAGA TCAAGGCAATGTTCCTTTAT
Adiantum            GGGAAGGGAGTGCAGATTCT AACTGAGATTCCAAATGATG TGAAGTTGATGTGTTTGTTT
Psilotum            -------------------- -------------------- --------------------
Psilotum2           GAAAAGGGTCTGCAACTCAT TCGGGAAACTCCTGTGGATA TCAAAAACATGCGCCTTTAT
PHYB_Arabidopsis    GACAGAGGTCTTCAGCTGAT CCGTGACATTCCCGAAGAGA TCAAATCAATAGAGGTTTTT
PHYB_Nicotiana      GAAAGAAGTGTGCAATTAAT CAGGGATATTCCAGAGGAAA TTAAGACCTTAACAGTACAT
PHYB_Solanum        GAAAAAGGCGTGCAGTTAAT CCGTGATATACCAGAGGAAA TTAAGACATTAACAGTACAT
PHYD_Arabidopsis    GAGAGAAATCTCCAGCTGAT CCGTAACATTCCCACGGAGG TCAAATCCATGGCTGTCTAC
PHYE_Arabidopsis    GAGAGGAACTCACAACTAAG GGTTGAAGTCGCCGAGGAGA TCAAAACTCTGCCTCTCAAT
PHYB_Oryza          GAAAGAGATTTACAACTTAT TCGAGATATCCCTGATGAAA TTAAAGAAGCCTCAGCATAT
PHYA_Avena          GGAAAAGGGATCAGAATCTC TTGCAACCTGCCAGAGAGAT TTATGAAGCAGTCAGTCTAT
PHYA_Oryza          GGAAAAGGGATTAGAGTCTC TTGCAACCTACCGGAGAGAT ATATGAAGCAAACAGTCTAC
PHYA_Zea            GCTAAAGGTATCAGAGTTGC TTGCAACCTGCCAGAGAGAT CCATGAAGCAAAAGGTTTAC
PHYA_Arabidopsis    GGAAAGAGTGTTCGGATAAC AAATGAGACCGGAGAAGAAG TAATGTCTGACACTTTGTAT
PHYA_Cucurbita      GGAAAGGGTATCCAGATAGT TAATGAGACTCCAGAAGAGG CTATGTCCGAGACCTTATAT
PHYA_Pisum          ACAAAGGGTATCCGTATAGC CAACGATGTTGCGGAGCATA TCGCGAGGGAAACCTTGTAT
PHYA_Nicotiana      GGAAAGAATATAATGATTGT GAATGACATGGTTGAAGATC TTCTCAATGAAACTTTATAC
PHYA_Solanum        GGAAAGAATATAATGATCTC GAATGACATGGTTGAAGATC TTCTTAATGAAACTTTATAT
PHYA_Arabidopsis    GAACGTAAAGTACAAATCAG CTGCGATTATCCTCAAGAAG TTTCATCAATGAGATTGTAT

Physcomitrella      GGAGATCAGATTCGGCTCCA GCAAGTTCTTTCTGATTTTC TATTAAACACAGTGCGATTT
Ceratodon           -------------------- -------------------- --------------------
Selaginella         GGAGATCAGGTGCGCCTCCA GCAGGTTCTGGCAGATTTTC TGCTCAATGCAATCCGCTTC
Adiantum            GGCGACCAGGCACGGCTGCA GCAAGTGCTGGCAGATTTGT TGTTTTGTGCCATAAATCAT
Psilotum            -------------------- -------------------- --------------------
Psilotum2           GGGGATCAGTTGAGGTTGCA GCAAGTGTTGGCTGATTTTT TGACAACTGCTGTGCGTTTT
PHYB_Arabidopsis    GGAGACCAGATAAGGATTCA ACAGCTCCTGGCTGAGTTTC TGCTGAGTATAATCCGGTAT
PHYB_Nicotiana      GGTGATCAAGTGAGAATTCA ACAGGTCTTGGCAGATTTCT TGCTTAACATGGTACGGTAT
PHYB_Solanum        GGTGATCAAGTGAGAATTCA ACAGGTCTTGGCAGATTTCT TGTTGAACATGGTACGGTAT
PHYD_Arabidopsis    GGTGACCAGATAAGGCTCCA ACAGGTTCTCGCAGAATTTC TGCTAAGTATTGTCCGTTAT
PHYE_Arabidopsis    GGTGACAGAGTCAAGCTCCA GCTTATTCTTGCTGATCTTC TACGCAACATTGTGAATCAT
PHYB_Oryza          GGTGACCAATATAGAATTCA ACAAGTTTTATGTGACTTTT TGCTAAGCATGGTGAGGTTT
PHYA_Avena          GGAGATGGTGTTCGACTCCA GCAGATCCTCTCTGACTTCC TGTTTATTTCAGTGAAGTTC
PHYA_Oryza          GGGGATGGTGTTCGACTACA GCAGATTCTCTCTGACTTCC TATTCGTCTCAGTGAAGTTC
PHYA_Zea            GGGGATGGTATCCGACTCCA GCAGATCGTCTCTGACTTCC TATTTGTTTCGGTGAAGTTC
PHYA_Arabidopsis    GGAGACAGTATTAGGCTTCA ACAAGTCTTGGCAGATTTCA TGCTGATGGCTGTAAACTTT
PHYA_Cucurbita      GGAGATAGTTTAAGGCTTCA ACAGGTCTTGGCGGACTTTC TATTGATATCAGTTAGTTAT
PHYA_Pisum          GGTGATAGTCTTAGGCTTCA GCAGGTCCTAGCTGACTTTT TACTAATTTCCATCAATTCC
PHYA_Nicotiana      GGAGATAGTCCGAGGCTTCA ACAGGTCTTAGCTAACTTTT TGTTAGTATGCGTGAATTCT
PHYA_Solanum        GGAGATAGTCCGAGACTTCA ACAAGTCCTAGCTAACTTTT TGTTAGTATCTGTGAACTCT
PHYA_Arabidopsis    GGAGACAACTTAAGGCTTCA GCAAATCCTTTCAGAGACAC TATTAAGCAGCATACGCTTC

Physcomitrella      ACGCCTTCACCTGAAGGT-- ----------TGGGTCAAAA TCAAGGTGGTCCCCACTAGG
```

```
Ceratodon            -------------------- -------------------- --------------------
Selaginella          ACACCGTCGTCAGAGAAC-- ----------TGGGTGGGGA TAAAGGTGGCCACATCTAGA
Adiantum             GCAACAACGACGAATGAGGA TGAGAAGGATTGGGTCACAA TCAAGGTGTCACGTACCAAA
Psilotum             -------------------- -------------------- --------------------
Psilotum2            ACTTCATCATCTGATGGG-- ----------TGGGTGGGCA TCAAGGTTGTTCCCACAATG
PHYB_Arabidopsis     GCACCATCTCAAGAG----- ----------TGGGTGGAGA TCCATTTAAGCCAACTTTCA
PHYB_Nicotiana       GCACCATCACCTGATGGT-- ----------TGGGTAGAGA TCCAACTTCAGCCAAATATG
PHYB_Solanum         GCACCATCACCTGATGGG-- ----------TGGGTAGAAA TCCAACTTCGACCAAGTATG
PHYD_Arabidopsis     GCACCCATGGAAGGC----- ----------TCGGTAGAGC TCCATCTATGCCCGACTCTG
PHYE_Arabidopsis     GCGCCGTTTCCAAATAGT-- ----------TGGGTAGGTA TCAGTATCTCACCAGGGCAG
PHYB_Oryza           GCTCCAGCTGAAAATGGC-- ----------TGGGTGGAGA TACAGGTCAGACCAAATATA
PHYA_Avena           TCTCCTGTTGGAGGT----- ----------TCTGTTGAGA TTTCTTCCAAGCTGACAAAG
PHYA_Oryza           TCTCCTGTTGGGGGT----- ----------TCTGTTGAGA TCTCTTGTAGCCTGACCAAG
PHYA_Zea             TCTCCTGCTGGTGGC----- ----------TCTGTTGACA TCTCTTCCAAGCTGACTAAG
PHYA_Arabidopsis     ACACCATCCGGAGGT----- ----------CAGCTAACTG TTTCAGCTTCCCTGAGGAAG
PHYA_Cucurbita       GCGCCTTCAGGAGGC----- ----------CAACTAACAA TTTCAACCGATGTGACCAAG
PHYA_Pisum           ACACCTAATGGAGGC----- ----------CAGGTTGTTA TAGCAGCCTCCTTAACTAAA
PHYA_Nicotiana       ACACCAAGTGGTGGT----- ----------CAGCTTAGTA TTTCAGGCACATTAACAAAA
PHYA_Solanum         ACACCAAGTGGTGGC----- ----------AAGCTTAGTA TTTCAGGCAAGCTAACTAAA
PHYA_Arabidopsis     ACGCCTGCATTGAGAGGATT GTGT---------GTCTCAT TCAAGGTAATTGCACGGATA

Physcomitrella       AAACGTCTTGGCGGGAGTGT GCATGTCGTGCATTTAGAAT TCAGGGTAAGTCATCCTGGA
Ceratodon            -------------------- -------------------- --------------------
Selaginella          AAGCGCTTGGGTGGTGTGGT GCACGTGATGCACCTCGAAT TCAGGATAACTCATCCGGGG
Adiantum             ACCAGACTTGATGATGGTGT ACATTTGATGCACTTCGAGA GTCGGATTTCACATTCAGGG
Psilotum             -------------------- -------------------- --------------------
Psilotum2            AAAGGCTTAGGAGGTGGCTT ACATGTGATGCGCTTTGATT ACAGAATTTCGCATCCAGGC
PHYB_Arabidopsis     AAGCAAATGGCTGATGGATT CGCCGCCATCCGCACAGAAT TCAGAATGGCGTGTCCAGGT
PHYB_Nicotiana       AAGCAAATATCTGATGAAGT AACTGTTGTGCATATTGAAT TCAGGATTGTATGCCCTGGT
PHYB_Solanum         ATGCCAATATCTGATGGAGT AACTGTTGTGCATATTGAAC TAGGATTATATGCGCCGGGA
PHYD_Arabidopsis     AACCAAATGGCTGACGGATT CTCCGCTGTACGTTTGGAGT TCAGAATGGCGTGTGCAGGG
PHYE_Arabidopsis     GAGCTTTCACGTGACAATGG TCGCTATATCCATCTACAGT TCAGGATGATACATCCGGGG
PHYB_Oryza           AAACAAAATTCTGATGGAAC AGACACAATGCTTTTCCCCT TCAGGTTTGCCTGTCCTGGC
PHYA_Avena           AACAGCATCGGAGAAAACCT TCATCTTATTGACCTTGAAC TTAGGATCAAGCACCAAGGA
PHYA_Oryza           AACAGCATTGGGGAAAACCT TCATCTCATAGACCTAGAAC TTAGGATCAAGCACCAGGGC
PHYA_Zea             AACAGCATTGGGGAAAACCT TCACCTCATAGACTTCGAAC TTAGGATCAAGCACCGAGGA
PHYA_Arabidopsis     GATCAGCTCGGGCGTTCTGT GCATCTTGCTAATCTAGAGA TCAGGTTAACGCATACCGGA
PHYA_Cucurbita       AATCAATTAGGAAAGTCGGT CCATCTGGTGCATTTGGAGT TCAGGATAACATATGCTGGA
PHYA_Pisum           GAACAGTTGGGAAAATCTGT CCATCTTGTTAACTTGGAGC TCAGCATAACACACGGTGGT
PHYA_Nicotiana       GATCGCATAGGAGAATCTGT TCAGCTTGCTCTCTTGGAAG TCAGGATAAGCCACACAGGG
PHYA_Solanum         GATCGCATAGGAGAATCTGT TCAGCTTGCTCTCTTGGAAT TCAGGATAAGACATACAGGG
PHYA_Arabidopsis     GAAGCTATAGGAAAAAGAAT GAAAAGAGTCGAACTTGAGT TCAGGATAATACACCCGGCA

Physcomitrella       GCTGGGCTTCCTGAAGAGCT TGTGTTGGAGATGTAT---G ATAGAGGCAAAGGCATGACT
Ceratodon            -------------------- -------------------- --------------------
Selaginella          GTTGGCCTTCCGGAAGAGCT TGTGCAAGAGATGTTT---G ACAGGGGCCGAGGCATGACA
Adiantum             CAAGGGATTTCTGAAGCCTT GGTGGAGGAAATGACC---A ACAAGTCGCAGAAATGGACA
Psilotum             -------------------- -------------------- --------------------
Psilotum2            AAAGGGATCCCAGAGGATTT GGTTCAACAGATGTTT---G ATTGCTCACGTGAAATAACA
PHYB_Arabidopsis     GAAGGTCTGCCTCCAGAGCT AGTCCGAGACATGTTC---C ATAGCAGCAGGTGGACAAGC
PHYB_Nicotiana       GAAGGGCTTCCTCCTGAATT GGTTCAAGACATGTTC---C ACAGCAGTCGGTGGGTAACC
PHYB_Solanum         ---AGGCTTCCTCCTGAATT GGTTCAAGACATGTTC---C ACAGCAGTCGGTGGGTAACT
PHYD_Arabidopsis     GAAGGTGTGCCGCCAGAGAA AGTGCAAGACATGTTC---C ATAGTAGCCGATGGACAAGT
PHYE_Arabidopsis     AAAGGACTTCCTTCAGAGAT GCTAAGTGATATGTTTGAGA CTCGAGATGGATGGGTCACC
PHYB_Oryza           GAAGGCCTTCCCCCAGAGAT TGTTCAAGACATGTTT---A GTAACTCCCGCTGGACAACC
PHYA_Avena           TTAGGAGTCCCAGCAGAGCT CATGGCACAAATGTTTGAGG AGGACAACAAGGAGCAGTCA
PHYA_Oryza           AAAGGAGTCCCAGCAGATCT GCTGTCACAAATGTACGAGG ATGACAATAAGGAGCAGTCG
PHYA_Zea             GCAGGAGTCCCAGCGGAAAT ATTGTCGCAAATGTATGAGG AGGACAATAAAGAGCAGTCA
PHYA_Arabidopsis     GCTGGGATACCTGAGTTTTT ACTAAACCAAATGTTT---G GGACTGAGGAAGATGTGTCA
PHYA_Cucurbita       GGAGGTATACCGGAGTCGTT GCTGAACGAGATGTTT---G GAAGCGAGGAGGACGCGTCC
PHYA_Pisum           AGTGGCGTGCCAGAAGCGGC GCTGAACCAGATGTTT---G GAAATAATGTGCTAGAATCT
PHYA_Nicotiana       GGAGGAGTGCCAGAAGAACT GCTAAGCCAAATGTTC---G GTACTGAGGCCGAAGCATCT
PHYA_Solanum         GGTGGAGTGCCAGAAGAATT GCTTAGCCAAATGTTT---G GTAGTGAAGCCGATGCATCT
PHYA_Arabidopsis     CCAGGACTGCCTGAGGATCT GGTAAGAGAGATGTTTCAGC CTTTGAGAAAGGGAACATCA

Physcomitrella       CAAGAAGGACTAGGTTTGAA CATGTGTCGGAAGCTCGTTA GGTTGATGAAT---GGGGAT
Ceratodon            -------------------- -------------------- --------------------
Selaginella          CAGGAGGGCCTGGGGCTTAG TATGTGCCGAAAACTCGTCA AGCTCATGAAC---GGAGAG
Adiantum             CCAGAAGGTTTAGCCATATC AATAAGCTGCACGCTAATAC GCCTCATGAAT---GGAGAT
Psilotum             -------------------- -------------------- --------------------
Psilotum2            CAGGAAGGCATGGGGCTTAG TGTTTCTCGGAAACTTGTCA GGCTTATGAAT---GGAGAT
PHYB_Arabidopsis     CCTGAAGGTTTAGGTCTAAG CGTATGTCGAAAGATTTTAA AGCTAATGAAC---GGTGAG
PHYB_Nicotiana       AAGGAAGGCCTAGGACTGAG CATGTGCAGAAAAATCTTAA AGCTTATGAAT---GGAGAT
```

```
PHYB_Solanum        CAGGAAGGCCTAGGACTGAG CATGTGCAGAAAAATGTTAA AGCTTATGAAT---GGAGAA
PHYD_Arabidopsis    CCAGAAGGATTAGGACTAAG CGTTTGCAGAAAGATTTTGA AGCTGATGAAC---GGAGGG
PHYE_Arabidopsis    CCTGATGGTTTAGGGCTTAA GCTTTCGCGGAAACTATTGG AGCAGATGAAT---GGCCGT
PHYB_Oryza          CAAGAGGGTATTGGCCTAAG CATATGCAGGAAGATCCTAA AATTGATGGGT---GGCGAG
PHYA_Avena          GAGGAGGGCTTGAGCCTCCT AGTTTCTAGAAACCTGCTGA GGCTCATGAAT---GGTGAT
PHYA_Oryza          GATGAAGGCATGAGTCTTGC GGTTTCTAGAAACCTGCTGA GGCTCATGAAT---GGCGAT
PHYA_Zea            GAGGAGGGCTTCAGCCTTGC TGTTTCTAGAAACCTTCTGA GGCTCATGAAT---ggngAC
PHYA_Arabidopsis    GAAGAAGGATTGAGCTTAAT GGTTAGCCGGAAACTGGTGA AGCTGATGAAT---GGAGAT
PHYA_Cucurbita      GAAGAGGGTTTCAGTCTGCT CATCAGTAGAAAGCTGGTGA AGCTGATGAAT---GGAGAC
PHYA_Pisum          GAGGAGGGTATTAGCCTACA CATCAGTCGGAAGTTGTTAA AGCTTATGAAT---GGAGAT
PHYA_Nicotiana      GAAGAAGGGATCAGCTTACT CATCAGCAGAAAGCTGGTGA AGCTGATGAAT---GGGGAA
PHYA_Solanum        GAAGAAGGGATCAGCTTACT TGTCAGCAGAAAACTGGTCA AGCTGATGAAT---GGGGAA
PHYA_Arabidopsis    AGGGAAGGTTTGGGATTACA CATTACCCAGAAGCTGGTGA AACTCATGGAGAGAGGAACA

Physcomitrella      GTTCATTACGTACGGGAAGC TATGCAGTGTTATTTTGTTG TAAATGTGGAGCTCCCA
Ceratodon           -------------------- -------------------- -----------------
Selaginella         GTTGAATATATCAGAGAAGC TGGCAAAAACTACTTCTTAG TCAGCCTGGAGCTCCCT
Adiantum            GTCAAATACACCACCGATGC CGGAAACAAATGCTTTCTTG TGACTATCCAGTTCCCT
Psilotum            -------------------- -------------------- -----------------
Psilotum2           GTGAGTTATATTCGGGAAGC TGGAGTGTGCTACTTCCTTG TCAATGTGGAATTCCCG
PHYB_Arabidopsis    GTTCAATACATCCGAGAATC AGAACGGTCCTATTTCCTCA TCATTCTGGAACTCCCT
PHYB_Nicotiana      ATCCAGTATATCAGAGAATC AGAAAGATGTTATTTCCTGA TCATCCTTGACCTACCA
PHYB_Solanum        ATCCAGTATATCAGAGAATC AGAAAGATGCTATTTCCTGA TTATCCTTGACCTGCCA
PHYD_Arabidopsis    GTTCAGTACATAAGAGAATT CGAACGCTCTTATTTCCTAA TCGTTATCGAACTCCCG
PHYE_Arabidopsis    GTGAGTTATGTCCGAGAAGA CGAACGGTGTTTCTTTCAGG TnGATCTTCAAGTGAAG
PHYB_Oryza          GTCCAATATATAAGGGAGTC GGAGCGGAGTTTCTTCCATA TCGTACTTGAGCTGCCC
PHYA_Avena          GTTCGGCATCTAAGGGAAGC TGGTGTGTCAACCTTCATCA TCACCGCTGAACTTGCT
PHYA_Oryza          GTCCGACATATGAGGGAAGC TGGCATGTCAACCTTCATCC TCAGCGTCGAACTTGCT
PHYA_Zea            ATTCGTCACCTCAGGGAAGC TGGCATGTCAACCTTCATTC TCACTGCTGAACTTGCT
PHYA_Arabidopsis    GTTCAGTACTTGAGACAAGC TGGGAAATCAAGTTTCATTA TCACTGCGGAACTCGCT
PHYA_Cucurbita      GTACGATATATGAGGGAAGC CGGGAAGTCGAGCTTCATCA TAACTGTTGAGCTTGCT
PHYA_Pisum          GTTCGTTATTTAAAAGAAGC AGGAAAATCATCGTTTATTC TATCTGTTGAACTTGCA
PHYA_Nicotiana      GTTCAGTACCTAAGAGAGGC GGGGCGATCAACTTTCATTA TATCCGTTGAACTTGCA
PHYA_Solanum        GTTCAGTACCTAAGAGAGGC TGGGCGATCAACTTTCATTA TATCCGTTGAACTCGCA
PHYA_Arabidopsis    TTGAGATACCTCAGAGAGTC TGAAATGTCAGCCTTTGTGA TCCTCACAGAATTTCCC
```

Appendix C

Alignment Of Nucleotides And Amino Acids For Comparison Of Phytochrome Genes From Nonangiosperms And Angiosperms

```
LABELS:
[MOUGEOTIA          GENBANK ACCESSION S52048]
[CHARA              GENBANK ACCESSION X80291]
[MARCHANTIA         GENBANK ACCESSION X80296]
[FUNARIA            GENBANK ACCESSION X80294]
[PHYSCOMITRELLA     GENBANK ACCESSION X75025]
[CERATODON          GENBANK ACCESSION X17084]
[SELAGINELLA        GENBANK ACCESSION X61458]
[EQUISETUM          GENBANK ACCESSION X80299]
[MARSILEA           GENBANK ACCESSION X80300]
[ADIANTUM           GENBANK ACCESSION D13519]
[PSILOTUM           GENBANK ACCESSION X74931]
[PSILOTUM2          GENBANK ACCESSION X74930]
[METASEQUOIA        GENBANK ACCESSION X80297]
[PSEUDOTSUGA        GENBANK ACCESSION U22458]
[PICEA              GENBANK ACCESSION X80298]
[EPHEDRA            GENBANK ACCESSION X80292]
[GNETUM             GENBANK ACCESSION X80295]
[PHYC_PIPER         GENBANK ACCESSION X80321]
[PHYB_PIPER         GENBANK ACCESSION X80322]
[PHYA_PIPER         GENBANK ACCESSION X80323]
[PHYB_ARABIDOPSIS   GENBANK ACCESSION X17342]
[PHYB_GLYCINE*      GENBANK ACCESSION L34843]
[PHYB_NICOTIANA     GENBANK ACCESSION L10114]
[PHYB_SOLANUM       GENBANK ACCESSION S51538]
[PHYD_ARABIDOPSIS   GENBANK ACCESSION X76609]
[PHYE_ARABIDOPSIS   GENBANK ACCESSION X76610]
[PHYB_ORYZA         GENBANK ACCESSION X57563]
[PHYA_AVENA         GENBANK ACCESSION X03242]
[PHYA_ORYZA         GENBANK ACCESSION X14172]
[PHYA_ZEA           GENBANK ACCESSION NONE]
[PHYA_ARABIDOPSIS   GENBANK ACCESSION X17341]
[PHYA_CUCURBITA     GENBANK ACCESSION M15265]
[PHYA_PISUM         GENBANK ACCESSION M37217]
[PHYA_GLYCINE*      GENBANK ACCESSION L34844]
[PHYA_NICOTIANA     GENBANK ACCESSION X66784]
[PHYA_SOLANUM       GENBANK ACCESSION S84872]
[PHYC_ARABIDOPSIS   GENBANK ACCESSION X17343]
```

*Note: the sequences from *Glycine* were not included in preliminary analyses.

```
MOUGEOTIA        -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
mougeotia        -??????????????????? ???????????????????? ????????????????????
CHARA            -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
chara            -??????????????????? ???????????????????? ????????????????????
MARCHANTIA       -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
marchantia       -??????????????????? ???????????????????? ????????????????????
FUNARIA          -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
funaria          -??????????????????? ???????????????????? ????????????????????
PHYSCOMITRELLA   -Q--S--F--G--C--M--I --A--V-----E--G--T-- ------------G--F--R-
physcomitrella   -CAAAGTTTCGGGTGTATGA TAGCAGTT---GAAGGAACA ------------GGATTTCG
CERATODON        -Q--N--F--G--C--M--V --A--V-----E--E--P-- ------------N--F--C-
ceratodon        -CAAAATTTTGGGTGTATGG TAGCAGTT---GAAGAGCCG ------------AATTTCTG
SELAGINELLA      -Q--P--F--G--C--M--L --A--V-----E--E--G-- ------------S--F--R-
selaginella      -CAGCCGTTTGGCCTAGTGC TGGCGGTG---GAGGAGGGC ------------AGCTTCCG
EQUISETUM        -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
equisetum        -??????????????????? ???????????????????? ????????????????????
MARSILEA         -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
marsilea         -??????????????????? ???????????????????? ????????????????????
ADIANTUM         -Q--Q--F--G--C--L--I --A--V-----E--E--E-- ------------T--F--R-
adiantum         -CAGCAATTTGGGTGCTTGA TTGCAGTG---GAGGAAGAG ------------ACCTTCAG
PSILOTUM         -Q--P--F--G--C--L--L --A--V-----D--E--N-- ------------S--F--R-
psilotum         -CAGCCCTTTGGTTGCCTCT TGGCTGTG---GACGAGAAT ------------TCTTTTCG
PSILOTUM2        -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
psilotum2        -??????????????????? ???????????????????? ????????????????????
METASEQUOIA      -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
metasequoia      -??????????????????? ???????????????????? ????????????????????
PSEUDOTSUGA      -Q--P--F--G--C--M--V --A--V-----N--D--K-- ------------T--F--E-
pseudotsuga      -CAGCCGTTTGGGTGCATGG TTGCAGTA---AATGATAAG ------------ACTTTTGA
PICEA            -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
picea            -??????????????????? ???????????????????? ????????????????????
EPHEDRA          -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
ephedra          -??????????????????? ???????????????????? ????????????????????
GNETUM           -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
```

```
gnetum             -??????????????????? ???????????????????? ????????????????????
PHYC_PIPER         -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
phyc_piper         -??????????????????? ???????????????????? ????????????????????
PHYB_PIPER         -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
phyb_piper         -??????????????????? ???????????????????? ????????????????????
PHYA_PIPER         -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
phya_piper         -??????????????????? ???????????????????? ????????????????????
PHYB_ARABIDOPSIS   -Q--P--F--G--C--M--I --A--V-----D--E--S-- ------------S--F--R-
phyb_arabidopsis   -CAGCCTTTCGGATGTATGA TCGCCGTC---GATGAATCC ------------AGTTTCCG
PHYB_GLYCINE       -Q--P--F--G--S--M--I --A--V-----D--E--P-- ------------S--F--R-
phyb_glycine       -CAGCCCTTCGGCTCCATGA TCGCCGTC---GACGAGCCC ------------TCCTTCCG
PHYB_NICOTIANA     -Q--P--F--G--C--M--I --A--V-----D--E--A-- ------------S--F--R-
phyb_nicotiana     -CAGCCTTTTGGTTGTATGA TAGCTGTA---GATGAGGCT ------------AGTTTTCG
PHYB_SOLANUM       -Q--P--F--G--C--M--I --A--V-----D--E--A-- ------------S--F--R-
phyb_solanum       -CAGCCTTTTGGTTGTATGA TAGCTGTA---GATGAGGCT ------------AGTTTTCG
PHYD_ARABIDOPSIS   -Q--P--F--G--C--L--I --A--V-----E--E--S-- ------------T--F--T-
phyd_arabidopsis   -CAGCCTTTTGGCTGCTTGA TCGCCGTC---GAAGAATCC ------------ACTTTCAC
PHYE_ARABIDOPSIS   -Q--P--F--G--C--L--I --A--V-----E--E--P-- ------------S--F--R-
phye_arabidopsis   -CAGCCCTTTGGTTGTTTGA TTGCTGTC---GAAGAACCT ------------AGTTTTAG
PHYB_ORYZA         -Q--P--F--G--C--T--L --A--V--A--D--D--S-- ------------S--F--R-
phyb_oryza         -CAGCCCTTCGGCTGCACGC TCGCCGTCGCCGACGACTCC ------------TCCTTCCG
PHYA_AVENA         -Q--T--F--G--C--L--L --A--L-----D--E--K-- ------------S--F--N-
phya_avena         -CAAACATTTGGTTGCCTGT TGGCCCTT---GATGAGAAG ------------AGCTTCAA
PHYA_ORYZA         -Q--P--F--G--C--L--L --A--L-----D--E--K-- ------------T--F--N-
phya_oryza         -CAACCATTTGGTTGCTTGT TGGCCCTT---GATGAGAAG ------------ACCTTCAA
PHYA_ZEA           -Q--P--F--G--C--L--L --A--L-----D--E--K-- ------------S--F--R-
phya_zea           -CAACCATTTGGTTGCCTGT TGGCCCTT---GACGAGAAG ------------AGCTTCAG
PHYA_ARABIDOPSIS   -Q--P--F--G--C--L--L --A--L-----D--E--K-- ------------T--F--K-
phya_arabidopsis   -CAGCCCTTCGGTTGTTTAC TTGCCTTG---GATGAGAAG ------------ACCTTCAA
PHYA_CUCURBITA     -Q--P--F--G--C--L--L --A--L-----D--D--K-- ------------T--F--K-
phya_cucurbita     -CAACCATTTGGTTGCTTGT TGGCCTTA---GATGACAAA ------------ACATTCAA
PHYA_PISUM         -Q--P--F--G--C--L--L --A--L-----D--E--K-- ------------T--C--K-
phya_pisum         -CAGCCTTTCGGGTGCTTGC TAGCTTTA---GATGAGAAA ------------ACGTGCAA
PHYA_GLYCINE       -Q--P--F--G--C--L--L --A--I-----D--E--K-- N--H--M--Q--T--C--K-
phya_glycine       -CAGCCTTTTGGTTGCTTGT TGGCCATT---GATGAGAAA AATCACATGCAGACATGTAA
PHYA_NICOTIANA     -Q--P--F--G--C--L--L --A--L-----D--E--K-- ------------T--F--K-
phya_nicotiana     -CAGCCATTTGGCTGTTTGT TAGCCCTT---GATGAGAAA ------------ACATTTAA
PHYA_SOLANUM       -Q--P--F--G--C--L--L --A--L-----D--E--K-- ------------T--L--K-
phya_solanum       -CAGCCATTTGGTTGTCTGT TAGCCCTG---GATGAGAAA ------------ACATTAAA
PHYC_ARABIDOPSIS   -Q--P--F--G--C--L--I --V--V-----D--E--K-- ------------N--L--K-
phyc_arabidopsis   -CAACCCTTTGGTTGTTTAA TCGTTGTT---GATGAGAAA ------------AACCTTAA

MOUGEOTIA          -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
mougeotia          ???????????????????? ???????????????????? ????????????????????
CHARA              -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
chara              ???????????????????? ???????????????????? ????????????????????
MARCHANTIA         -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
marchantia         ???????????????????? ???????????????????? ????????????????????
FUNARIA            -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
funaria            ???????????????????? ???????????????????? ????????????????????
PHYSCOMITRELLA     -V--I--A--Y--S--E--N --A--P--E--I--L--D-- L--V--P--Q-----A--V-
physcomitrella     TGTGATAGCGTACAGTGAAA ATGCACCCGAGATTCTAGAC CTGGTGCCACAG---GCCGT
CERATODON          -V--I--A--Y--S--E--N --A--S--E--F--L--D-- L--I--P--Q-----A--V-
ceratodon          TGTTATAGCGTACAGTGAGA ATGCGTCCGAGTTTCTAGAT CTGATACCCCAG---GCCGT
SELAGINELLA        -V--I--A--F--S--D--N --A--G--E--M--L--D-- L--M--P--Q-----S--V-
selaginella        TGTGATTGCCTTCAGTGACA ATGCGGGCGACATGCTGGAC CTCATGCCTCAG---TCCGT
EQUISETUM          -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
equisetum          ???????????????????? ???????????????????? ????????????????????
MARSILEA           -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
marsilea           ???????????????????? ???????????????????? ????????????????????
ADIANTUM           -V--L--H--M--C--E--- --A--P--E--M--L--D-- V--A--T--Q-----A--V-
adiantum           AGTCTTGCATATGTGCGAA- --GCACCTGAGATGCTAGAT GTAGCCACTCAA---GCCGT
PSILOTUM           -V--I--A--Y--S--E--N --A--P--E--V--L--D-- L--I--P--Q-----F--V-
psilotum           GGTGATTGCCTACAGCGAAA ATGCACCAGAAGTACTCGAT CTTATACCACAG---TTCGT
PSILOTUM2          -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
psilotum2          ???????????????????? ???????????????????? ????????????????????
METASEQUOIA        -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
metasequoia        ???????????????????? ???????????????????? ????????????????????
PSEUDOTSUGA        -V--I--A--Y--S--E--N --A--P--D--M--L--Q-- V--S--A--H-----A--V-
pseudotsuga        AGTGATTGCTTACAGTGAAA ATGCCCCAGATATGCTTCAA GTGTCAGCTCAT---GCAGT
PICEA              -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
picea              ???????????????????? ???????????????????? ????????????????????
EPHEDRA            -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
```

```
ephedra            ???????????????????? ???????????????????? ????????????????????
GNETUM             -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
gnetum             ???????????????????? ???????????????????? ????????????????????
PHYC_PIPER         -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
phyc_piper         ???????????????????? ???????????????????? ????????????????????
PHYB_PIPER         -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
phyb_piper         ???????????????????? ???????????????????? ????????????????????
PHYA_PIPER         -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
phya_piper         ???????????????????? ???????????????????? ????????????????????
PHYB_ARABIDOPSIS   -I--I--G--Y--S--E--N --A--R--E--M--L--G-- I--M--P--Q-----S--V-
phyb_arabidopsis   GATCATCGGTTACAGTGAAA ACGCCAGAGAAATGTTAGGG ATTATGCCTCAA---TCTGT
PHYB_GLYCINE       -I--L--G--Y--S--D--N --A--R--D--M--L--G-- I--T--P--Q-----S--V-
phyb_glycine       CATCCTCGGTTACTCCGACA ACGCCCGCGACATGCTCGGC ATTACTCCGCAG---TCCGT
PHYB_NICOTIANA     -V--I--A--Y--S--E--N --A--C--E--M--L--S-- L--T--P--Q-----S--V-
phyb_nicotiana     TGTTATTGCTTACAGCGAAA ATGCGTGCGAAATGCTTAGT CTAACTCCACAA---TCAGT
PHYB_SOLANUM       -V--I--A--Y--S--E--N --A--C--E--M--L--S-- L--T--P--Q-----S--V-
phyb_solanum       TGTAATAGCTTATAGTGAAA ATGCCTGTGAAATGCTTAGT TTAACTCCACAA---TCTGT
PHYD_ARABIDOPSIS   -I--I--G--Y--S--E--N --A--R--E--M--L--G-- L--M--S--Q-----S--V-
phyd_arabidopsis   AATCATCGGTTACAGTGAAA ATGCGCGGGAAATGCTAGGG CTCATGTCTCAA---TCTGT
PHYE_ARABIDOPSIS   -I--L--G--L--S--D--N --S--S--D--F--L--G-- L--L--S--L-----P--S-
phye_arabidopsis   GATACTTGGTCTTAGTGACA ACTCTTCTGACTTTCTTGGT TTGTTGTCTCTT---CCTTC
PHYB_ORYZA         -L--L--A--Y--S--E--N --T--A--D--L--L--D-- L--S--P--H--H--S--V-
phyb_oryza         CCTCCTCGCCTACTCCGAGA ACACCGCCGACCTGCTCGAC CTGTCGCCCCACCACTCCGT
PHYA_AVENA         -V--I--A--F--S--E--N --A--P--E--M--L--T-- T--V--S--H-----A--V-
phya_avena         TGTCATCGCGTTCAGTGAGA ACGCGCCAGAAATGCTTACA ACGGTCAGCCAT---GCGGT
PHYA_ORYZA         -V--I--A--L--S--E--N --A--P--E--M--L--T-- T--V--S--H-----A--V-
phya_oryza         TGTTATAGCGCTCAGCGAGA ATGCACCAGAGATGCTTACA ACTGTCAGCCAT---GCAGT
PHYA_ZEA           -V--I--A--F--S--E--N --A--P--E--M--L--T-- T--V--S--H-----A--V-
phya_zea           GGTCATTGCATTCAGTGAGA ATGCACCTGAAATGCTTACA ACGGTCAGCCAT---GCTGT
PHYA_ARABIDOPSIS   -V--I--A--Y--S--E--N --A--S--E--L--L--T-- M--A--S--H-----A--V-
phya_arabidopsis   AGTTATTGCATACAGCGAGA ATGCATCTGAGCTGTTGACA ATGGCCAGTCAT---GCAGT
PHYA_CUCURBITA     -V--I--A--Y--S--E--N --A--P--E--M--L--T-- M--V--S--H-----A--V-
phya_cucurbita     GGTTATTGCGTATAGTGAAA ATGCCCCTGAAATGTTGACC ATGGTGAGCCAT---GCTGT
PHYA_PISUM         -V--V--A--Y--S--E--N --A--P--E--M--L--T-- M--V--S--H-----A--V-
phya_pisum         GGTTGTTGCGTATAGTGAGA ATGCGCCTGAGATGCTGACT ATGGTGAGTCAT---GCTGT
PHYA_GLYCINE       -V--I--A--Y--S--E--N --E--P--E--M--L--T-- M--V--S--H-----A--V-
phya_glycine       GGTCATTGCATACAGTGAGA ACGAGCCCGAAATGCTGACC ATGGTTAGCCAT---GCTGT
PHYA_NICOTIANA     -V--I--A--F--S--E--N --A--P--E--M--L--T-- M--V--S--H-----A--V-
phya_nicotiana     GGTCATAGCATTCAGCGAGA ATGCCCCCGAAATGCTGACC ATGGTTAGCCAT---GCTGT
PHYA_SOLANUM       -V--I--A--F--S--E--N --A--P--E--M--L--T-- M--V--S--H-----A--V-
phya_solanum       GGTCATAGCATTCAGTGAAA ATGCCCCTGAAATGCTGACC ATGGTTAGCCAT---GCTGT
PHYC_ARABIDOPSIS   -V--I--A--F--S--E--N --T--Q--E--M--L--G-- L--I--P--H-----T--V-
phyc_arabidopsis   AGTCATTGCCTTTAGTGAAA ACACTCAAGAGATGTTGGGT TTGATTCCACAT---ACAGT

MOUGEOTIA          -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
mougeotia          ???????????????????? ???????????????????? ????????????????????
CHARA              -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
chara              ???????????????????? ???????????????????? ????????????????????
MARCHANTIA         -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
marchantia         ???????????????????? ???????????????????? ????????????????????
FUNARIA            -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
funaria            ???????????????????? ???????????????????? ????????????????????
PHYSCOMITRELLA     -P--S--V-----G------ --E--M--D--------T-- L--R--I--G--T--D--V-
physcomitrella     TCCAAGCGTG---GGG---- --GAGATGGAC------ACG CTACGAATCGGGACGGATGT
CERATODON          -P--S--M-----G------ --E--M--D--------V-- L--G--I--G--T--D--I-
ceratodon          CCCAAGTATG---GGG---- --GAGATGGAC------GTG CTAGGAATCGGGACGGATAT
SELAGINELLA        -P--S--L-----G--S--G --Q--Q--D--------V-- L--T--I--G--T--D--A-
selaginella        TCCAAGCCTG---GGCAGCG GGCAGCAGGAC------GTG CTGACCATCGGGACCGACGC
EQUISETUM          -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
equisetum          ???????????????????? ???????????????????? ????????????????????
MARSILEA           -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
marsilea           ???????????????????? ???????????????????? ????????????????????
ADIANTUM           -P--T--M-----G------ --Q--Y--S--------R-- L--C--I--G--A--D--V-
adiantum           GCCAACAATG---GGG---- --CAATATAGC------CGT CTCTGTATAGGTGCTGATGT
PSILOTUM           -P--T--V-----G------ --L--Q--E--------I-- L--R--I--G--T--D--A-
psilotum           TCCAACGGTG---GGA---- --CTGCAGGAG------ATA CTGAGAATTGGCACCGATGC
PSILOTUM2          -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
psilotum2          ???????????????????? ???????????????????? ????????????????????
METASEQUOIA        -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
metasequoia        ???????????????????? ???????????????????? ????????????????????
PSEUDOTSUGA        -P--S--V--E--------D --Q--Q-----------I-- L--C--I--G--S--D--A-
PSEUDOTSUGA        TCCCAGCGTGGAG------G ACCAGCAG---------ATT CTCTGTATTGGCAGTGATGC
PICEA              -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
```

```
picea              ???????????????????? ???????????????????? ????????????????????
EPHEDRA            -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
ephedra            ???????????????????? ???????????????????? ????????????????????
GNETUM             -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
gnetum             ???????????????????? ???????????????????? ????????????????????
PHYC_PIPER         -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
phyc_piper         ???????????????????? ???????????????????? ????????????????????
PHYB_PIPER         -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
phyb_piper         ???????????????????? ???????????????????? ????????????????????
PHYA_PIPER         -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
phya_piper         ???????????????????? ???????????????????? ????????????????????
PHYB_ARABIDOPSIS   -P--T--L------------ -----E--K--P--E--I-- L--A--M--G--T--D--V-
phyb_arabidopsis   TCCTACTCTT---------- -----GAGAAACCTGAGATT CTAGCTATGGGAACTGATGT
PHYB_GLYCINE       -P--S--L-----------D --D--K--N--D--A--A-- F--A--L--G--T--D--V-
phyb_glycine       CCCTTCGCTC---------G ACGACAAGAACGACGCCGCC TTCGCTCTCGGCACCGATGT
PHYB_NICOTIANA     -P--S--L------------ -----E--R--P--E--I-- L--T--V--G--T--D--V-
phyb_nicotiana     TCCAAGCCTT---------- -----GAGCGGCCTGAGATC CTCACTGTTGGAACTGATGT
PHYB_SOLANUM       -P--S--L------------ -----E--K--C--E--I-- L--T--I--G--T--D--V-
bsolanum           TCCAAGCCTT---------- -----GAGAAGTGTGAGATC CTCACTATTGGAACTGATGT
PHYD_ARABIDOPSIS   -P--S--I------------ --E--D--K--S--E--V-- L--T--I--G--T--D--L-
phyd_arabidopsis   ACCAAGCATC---------- --GAGGACAAATCAGAGGTT TTAACGATTGGTACGGATTT
PHYE_ARABIDOPSIS   -T--S--H-----S--G--- --E--F--D--K--V--K-- G--L--I--G--I--D--A-
phye_arabidopsis   CACCTCCCAT---TCTGGT- --GAGTTTGATAAAGTCAAG GGTTTGATTGGAATCGATGC
PHYB_ORYZA         -P--S--L--D--S--S--A --V--P--P--P--V----- ---S--L--G--A--D--A-
phyb_oryza         CCCCTCGCTCGACTCCTCCG CGGTGCCTCCCCCCGTC--- ---TCGCTCGGCGCAGACGC
PHYA_AVENA         -P--S--V------------ --D--D--P--P--R--L-- ---G--I--G--T--N--V-
phya_avena         ACCCAGTGTT---------- --GATGATCCCCCAAGGCTG ---GGGATTGGCACCAATGT
PHYA_ORYZA         -P--S--V------------ --D--D--P--P--K--L-- ---R--I--G--T--N--V-
phya_oryza         GCCAAGTGTT---------- --GATGATCCCCCAAAGCTA ---CGCATTGGCACCAATGT
PHYA_ZEA           -P--N--V------------ --D--D--P--P--K--L-- ---G--I--G--T--N--V-
phya_zea           GCCGAACGTT---------- --GATGATCCCCCGAAGCTA ---GGAATTGGCACCAATGT
PHYA_ARABIDOPSIS   -P--S--V-----G------ -----E--H--P--V--L-- ---G--I--G--T--D--I-
phya_arabidopsis   TCCTAGTGTT---GGC---- -----GAACACCCTGTTCTA ---GGCATTGGGACAGATAT
PHYA_CUCURBITA     -P--S--M-----G------ -----D--Y--P--V--L-- ---G--I--G--T--D--V-
phya_cucurbita     CCCAAGCATG---GGG---- -----GATTACCCTGTTCTT ---GGCATTGGCACAGATGT
PHYA_PISUM         -P--S--V-----G------ -----D--H--P--A--L-- ---G--I--G--T--D--I-
phya_pisum         TCCTAGCGTT---GGT---- -----GACCATCCTGCTCTT ---GGCATTGGAACTGACAT
PHYA_GLYCINE       -P--S--V-----G------ -----D--H--P--A--L-- ---G--I--G--T--D--I-
phya_glycine       CCCCAGTGTT---GGT---- -----GACCACCCTGCCCTT ---GGCATTGGCACTGACAT
PHYA_NICOTIANA     -P--S--V-----G------ -----E--L--P--A--L-- ---G--I--G--T--D--I-
phya_nicotiana     TCCAAGTGTG---GGT---- -----GAGCTTCCAGCTCTT ---GGCATTGGGACTGATAT
PHYA_SOLANUM       -P--S--V-----G------ -----E--H--P--V--L-- ---G--I--G--I--D--I-
phya_solanum       TCCAAGTGTT---GGT---- -----GAGCATCCAGTTCTT ---GGGATCGGGATTGATAT
PHYC_ARABIDOPSIS   -P--S--M------------ --E--Q--R-----E--A-- L--T--I--G--T--D--V-
phyc_arabidopsis   ACCAAGTATG---------- --GAGCAGCGT---GAAGCT TTGACTATAGGAACTGATGT

MOUGEOTIA          -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
mougeotia          ???????????????????? ???????????????????? ????????????????????
CHARA              -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
chara              ???????????????????? ???????????????????? ????????????????????
MARCHANTIA         -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
marchantia         ???????????????????? ???????????????????? ????????????????????
FUNARIA            -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
funaria            ???????????????????? ???????????????????? ????????????????????
PHYSCOMITRELLA     -R--T--L--F--T--A--S --S--V--A--S--L--E-- K--A--A--E--A--Q--E-
physcomitrella     GAGAACGCTGTTCACAGCTT CAAGTGTCGCCTCTCTTGAG AAGGCAGCTGAAGCACAAGA
CERATODON          -R--T--L--F--T--P--S --S--S--A--A--L--E-- K--A--A--A--T--Q--D-
ceratodon          AAGAACTTTATTCACACCGT CGAGTAGTGCCGCTCTTGAG AAGGCAGCTGCAACTCAGGA
SELAGINELLA        -R--T--L--F--T--A--A --A--S--A--L-----E-- K--A--A--G--A--V--D-
selaginella        CCGCACCCTCTTCACGGCAG CAGCCAGCGCCCTG---GAG AAGGCGGCAGGAGCCGTCGA
EQUISETUM          -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
equisetum          ???????????????????? ???????????????????? ????????????????????
MARSILEA           -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
marsilea           ???????????????????? ???????????????????? ????????????????????
ADIANTUM           -R--T--L--L--S--P--A --S--A--S--A--L--D-- R--V--I--G--V--V--D-
adiantum           AAGAACTCTGTTATCTCCGG CGAGTGCATCAGCATTGGAT CGAGTCATCGGTGTTGTCGA
PSILOTUM           -R--D--L--F--T--A--A --G--A--A--R--L--G-- R--V--V--G--A--V--D-
psilotum           TCGTGACCTCTTCACTGCAG CTGGTGCAGCTCGGCTTGGA AGAGTTGTTGGGGCTGTCGA
PSILOTUM2          -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
psilotum2          ???????????????????? ???????????????????? ????????????????????
METASEQUOIA        -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
metasequoia        ???????????????????? ???????????????????? ????????????????????
PSEUDOTSUGA        -R--T--L--F--T--P--S --S--T--M--A--L--Q-- K--A--A--K--F--G--D-
```

PSEUDOTSUGA	AAGGACCTTGTTCACCCCTT	CGAGCACCATGGCATTGCAG	AAGGCCGCTAAGTTTGGAGA
PICEA	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
picea	????????????????????	????????????????????	????????????????????
EPHEDRA	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
ephedra	????????????????????	????????????????????	????????????????????
GNETUM	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
gnetum	????????????????????	????????????????????	????????????????????
PHYC_PIPER	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
phyc_piper	????????????????????	????????????????????	????????????????????
PHYB_PIPER	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
phyb_piper	????????????????????	????????????????????	????????????????????
PHYA_PIPER	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
phya_piper	????????????????????	????????????????????	????????????????????
PHYB_ARABIDOPSIS	-R--S--L--F--T--S--S	--S--S--I--L--L--E--	R--A--F--V--A--R--E-
phyb_arabidopsis	GAGATCTTTGTTCACTTCTT	CGAGCTCGATTCTACTCGAG	CGTGCTTTCGTTGCTCGAGA
PHYB_GLYCINE	-R--A--L--F--T--H--S	--S--A--L--L--L--E--	K--A--F--S--A--R--E-
phyb_glycine	CCGAGCCCTCTTCACTCACT	CCAGCGCCTTACTCCTCGAA	AAGGCCTTCTCCGCACGCGA
PHYB_NICOTIANA	-R--T--L--F--T--P--S	--S--S--V--L--L--E--	R--A--F--G--A--R--E-
phyb_nicotiana	TAGGACCCTTTTTACTCCTT	CTAGCTCTGTTTTGCTAGAA	AGAGCATTTGGGGCGCGCGA
PHYB_SOLANUM	-R--T--L--F--T--P--S	--S--S--V--L--L--E--	R--A--F--G--A--R--E-
phyb_solanum	TAGGACCCTTTTTACCCCTT	CTAGCTCTGTTTTGCTGGAA	AGAGCATTTGGGGCACGTGA
PHYD_ARABIDOPSIS	-R--S--L--F--K--S--S	--S--Y--L--L--L--E--	R--A--F--V--A--R--E-
phyd_arabidopsis	GCGATCTCTCTTCAAGTCAT	CGAGCTACCTTCTCCTCGAG	CGCGCGTTCGTGGCTCGAGA
PHYE_ARABIDOPSIS	-R--T--L--F--T--P--S	--S--G--A--S--L--S--	K--A--A--S--F--T--E-
phye_arabidopsis	AAGGACGCTTTTTACTCCTT	CCTCTGGAGCTTCTTTGTCT	AAAGCTGCTTCCTTTACTGA
PHYB_ORYZA	-R--L--L--F--A--P--S	--S--A--V--L--L--E--	R--A--F--A--A--R--E-
phyb_oryza	GCGCCTCCTTTTCGCTCCCT	CGTCCGCCGTCCTCCTCGAG	CGCGCCTTCGCCGCGCGCGA
PHYA_AVENA	-R--S--L--F--S--D--Q	--G--A--T--A--L--H--	K--A--L--G--F--A--D-
phya_avena	ACGGTCTCTTTTCAGTGACC	AAGGTGCCACAGCACTGCAT	AAGGCACTAGGATTTGCTGA
PHYA_ORYZA	-R--S--L--F--T--D--P	--G--T--T--A--L--Q--	K--A--L--G--F--A--D-
phya_oryza	ACGGTCTCTTTTCACTGACC	CAGGTACCACAGCACTGCAG	AAGGCACTGGGATTTGCTGA
PHYA_ZEA	-R--S--L--F--T--D--P	--G--A--T--A--L--Q--	K--A--L--G--F--A--D-
phya_zea	GCGATCCCTTTTCACTGACC	CTGGTGCTACAGCACTGCAG	AAGGCACTTGGATTTGCTGA
PHYA_ARABIDOPSIS	-R--S--L--F--T--A--P	--S--A--S--A--L--Q--	K--A--L--G--F--G--D-
phya_arabidopsis	AAGGAGTCTTTTCACTGCTC	CTAGTGCGTCTGCATTGCAG	AAAGCCCTTGGATTTGGAGA
PHYA_CUCURBITA	-R--T--I--F--T--A--P	--S--A--S--A--L--L--	K--A--L--G--F--G--E-
phya_cucurbita	AAGGACTATTTTCACCGCAC	CTAGTGCTTCTGCACTGTTG	AAGGCCTTGGGCTTTGGAGA
PHYA_PISUM	-R--T--V--F--T--A--P	--S--A--S--A--L--Q--	K--A--L--G--F--A--E-
phya_pisum	AAGGACTGTTTTCACTGCGC	CGAGTGCTTCTGCCTTGCAG	AAGGCGCTAGGGTTTGCGGA
PHYA_GLYCINE	-K--T--L--F--T--A--P	--S--V--S--G--L--Q--	K--A--L--G--C--A--D-
phya_glycine	AAAAACTCTATTCACTGCAC	CAAGTGTTTCTGGATTGCAG	AAGGCTCTAGGATGTGCGGA
PHYA_NICOTIANA	-R--T--I--F--T--G--P	--S--A--A--A--L--Q--	K--A--L--G--F--G--E-
phya_nicotiana	CAGAACGATCTTCACTGGTC	CTAGCGCAGCTGCATTGCAG	AAGGCTTTGGGGTTCGGAGA
PHYA_SOLANUM	-R--T--I--F--T--G--P	--S--G--A--A--L--Q--	K--A--L--G--F--G--E-
phya_solanum	CAGAACGATCTTCACTGGTC	CTAGTGGCGCAGCATTGCAG	AAAGCCTTGGGGTTTGGGGA
PHYC_ARABIDOPSIS	-K--S--L--F--L--S--P	--G--C--S--A--L--E--	K--A--V--D--F--G--E-
phyc_arabidopsis	GAAATCATTGTTTCTGTCTC	CAGGTTGTTCTGCTTTGGAG	AAAGCTGTTGACTTTGGTGA
MOUGEOTIA	-?--?--?--?--?--?--?-	-?--?--?--?--?--?--?-	-?--?--?--?--?--?-
mougeotia	?????????????????????	?????????????????????	??????????????????
CHARA	-?--?--?--?--?--?--?-	-?--?--?--?--?--?--?-	-?--?--?--?--?--?-
chara	?????????????????????	?????????????????????	??????????????????
MARCHANTIA	-?--?--?--?--?--?--?-	-?--?--?--?--?--?--?-	-?--?--?--?--?--?-
marchantia	?????????????????????	?????????????????????	??????????????????
FUNARIA	-?--?--?--?--?--?--?-	-?--?--?--?--?--?--?-	-?--?--?--?--?--?-
funaria	?????????????????????	?????????????????????	??????????????????
PHYSCOMITRELLA	-M--S--L--L--N--P--I-	-T--V--N--C--R--R--S-	-G--K--Q--L--Y--A-
physcomitrella	AATGAGCCTACTCAATCCAAT	CACCGTTAACTGCAGAAGATC	TGGGAAGCAATTATATGC
CERATODON	-I--S--L--L--N--P--I-	-T--V--H--C--R--R--S-	-G--K--P--L--Y--A-
ceratodon	TATAAGCCTTCTTAACCCAAT	CACTGTTCATTGCAGACGCTC	AGGGAAACCGTTATATGC
SELAGINELLA	-L--S--M--L--N--P--I-	-W--V--Q--S--K--T--S-	-A--K--P--F--Y--A-
selaginella	CTTGAGCATGCTCAACCCGAT	ATGGGTACAGTCCAAGACGTC	GGCCAAGCCCTTCTACGC
EQUISETUM	-?--?--?--?--?--?--?-	-?--?--?--?--?--?--?-	-?--?--?--?--?--?-
equisetum	?????????????????????	?????????????????????	??????????????????
MARSILEA	-?--?--?--?--?--?--?-	-?--?--?--?--?--?--?-	-?--?--?--?--?--?-
marsilea	?????????????????????	?????????????????????	??????????????????
ADIANTUM	-V--S--M--F--N--P--I-	-T--V--Q--S--R--S--S-	-G--K--P--F--Y--A-
adiantum	TGTGAGCATGTTCAATCCCAT	AACCGTGCAAAGCCGAAGCTC	TGGAAAGCCCTTCTATGC
PSILOTUM	-V--S--L--F--N--P--I-	-I--V--Q--T--K--S--T-	-G--K--P--F--D--A-
psilotum	TGTTTCATTGTTTAATCCCAT	AATTGTTCAGACTAAAAGCAC	GGGCAAGCCTTTCGATGC
PSILOTUM2	-?--?--?--?--?--?--?-	-?--?--?--?--?--?--?-	-?--?--?--?--?--?-
psilotum2	?????????????????????	?????????????????????	??????????????????
METASEQUOIA	-?--?--?--?--?--?--?-	-?--?--?--?--?--?--?-	-?--?--?--?--?--?-

```
metasequoia        ????????????????????? ????????????????????? ??????????????????
PSEUDOTSUGA        -L--N--L--L--N--P--I- -L--V--H--S--K--S--S- -G--K--P--F--Y--A-
pseudotsuga        CCTGAATCTTCTGAATCCAT TTTGGTCCACTCCAAGAGTTC AGGTAAGCCCTTCTACGC
PICEA              -?--?--?--?--?--?--?- -?--?--?--?--?--?--?- -?--?--?--?--?--?-
picea              ????????????????????? ????????????????????? ??????????????????
EPHEDRA            -?--?--?--?--?--?--?- -?--?--?--?--?--?--?- -?--?--?--?--?--?-
ephedra            ????????????????????? ????????????????????? ??????????????????
GNETUM             -?--?--?--?--?--?--?- -?--?--?--?--?--?--?- -?--?--?--?--?--?-
gnetum             ????????????????????? ????????????????????? ??????????????????
PHYC_PIPER         -?--?--?--?--?--?--?- -?--?--?--?--?--?--?- -?--?--?--?--?--?-
phyc_piper         ????????????????????? ????????????????????? ??????????????????
PHYB_PIPER         -?--?--?--?--?--?--?- -?--?--?--?--?--?--?- -?--?--?--?--?--?-
phyb_piper         ????????????????????? ????????????????????? ??????????????????
PHYA_PIPER         -?--?--?--?--?--?--?- -?--?--?--?--?--?--?- -?--?--?--?--?--?-
phya_piper         ????????????????????? ????????????????????? ??????????????????
PHYB_ARABIDOPSIS   -I--T--L--L--N--P--V- -W--I--H--S--K--N--T- -G--K--P--F--Y--A-
phyb_arabidopsis   GATTACCTTGTTAAATCCGGT TTGGATCCATTCCAAGAATAC TGGTAAACCGTTTTACGC
PHYB_GLYCINE       -I--S--L--M--N--P--I- -W--I--H--S--R--T--S- -G--K--P--F--Y--G-
phyb_glycine       AATTAGCCTCATGAACCCTAT CTGGATCCACTCCAGAACCTC CGGGAAGCCTTTCTATGG
PHYB_NICOTIANA     -I--T--L--L--N--P--I- -W--I--H--S--K--N--S- -G--K--P--F--Y--A-
phyb_nicotiana     GATCACTTTGCTGAATCCTAT TTGGATTCATTCCAAGAATTC TGGCAAGCCATTTTACGC
PHYB_SOLANUM       -I--T--L--L--N--P--I- -W--I--H--S--K--N--S- -G--K--P--F--Y--A-
phyb_solanum       GATCACTTTACTCAACCCAAT TTGGATACATTCCAAGAATTC TGGAAAGCCCTTTTATGC
PHYD_ARABIDOPSIS   -I--T--L--L--N--P--I- -W--I--H--S--N--N--T- -G--K--P--F--Y--A-
phyd_arabidopsis   GATCACGCTTCTGAATCCTAT TTGGATTCACTCTAACAACAC TGGTAAACCTTTCTACGC
PHYE_ARABIDOPSIS   -I--S--L--L--N--P--V- -L--V--H--S--R--T--T- -Q--K--P--F--Y--A-
phye_arabidopsis   GATTTCCTTGTTGAATCCTGT TTTGGTCCATTCTAGGACGAC CCAGAAGCCTTTTTATGC
PHYB_ORYZA         -I--S--L--L--N--P--L- -W--I--H--S--R--V--S- -S--N--P--F--Y--A-
phyb_oryza         GATCTCGCTGCTCAACCCGCT CTGGATCCACTCCAGGGTCTC CTCTAACCCTTTCTACGC
PHYA_AVENA         -V--S--L--L--N--P--I- -L--V--Q--C--K--T--S- -G--K--P--F--Y--A-
phya_avena         TGTGTCTTTGCTGAATCCTAT CCTGGTTCAGTGCAAGACATC AGGCAAGCCTTTCTATGC
PHYA_ORYZA         -V--S--L--L--N--P--I- -L--V--Q--C--K--T--S- -G--K--P--F--Y--A-
phya_oryza         TGTTTCCTTGCTGAACCCTAT CCTAGTTCAATGCAAGACCTC AGGCAAGCCTTTCTATGC
PHYA_ZEA           -V--S--L--L--N--P--I- -L--V--Q--C--K--T--S- -G--K--P--F--Y--A-
phya_zea           TGTTTCTTTGCTGAATCCTAT CCTGGTTCAGTGCAAGACCTC AGGCAAGCCATTCTATGC
PHYA_ARABIDOPSIS   -V--S--L--L--N--P--I- -L--V--H--C--R--T--S- -A--K--P--F--Y--A-
phya_arabidopsis   TGTCTCTCTTTTGAATCCCAT TCTTGTGCACTGCAGGACTTC TGCAAAGCCCTTTTATGC
PHYA_CUCURBITA     -V--T--L--L--N--P--I- -L--V--H--C--K--T--S- -G--K--P--F--Y--A-
phya_cucurbita     GGTTACACTTCTTAATCCTAT CCTGGTGCATTGCAAGACTTC TGGAAAACCCTTCTATGC
PHYA_PISUM         -V--S--L--L--N--P--I- -L--V--H--C--K--T--S- -G--K--P--F--Y--A-
phya_pisum         GGTTTCGCTTCTTAACCCGAT TCTTGTTCATTGCAAGACTTC TGGGAAGCCGTTTTACGC
PHYA_GLYCINE       -V--S--L--L--N--P--I- -L--V--H--C--K--T--S- -G--K--P--F--Y--A-
phya_glycine       CGTTTCGCTTCTTAACCCCAT ACTTGTCCATTGCAAGACCTC TGGGAAGCCCTTTTATGC
PHYA_NICOTIANA     -V--S--L--L--N--P--V- -L--V--H--C--K--T--S- -G--K--P--Y--Y--A-
phya_nicotiana     GGTTTCTCTGTTAAATCCGGT CCTCGTTCACTGCAAAACTTC TGGGAAGCCATATTATGC
PHYA_SOLANUM       -V--S--L--L--N--P--V- -L--V--H--C--K--N--S- -G--K--P--F--Y--A-
phya_solanum       AGTTTCTCTGTTAAATCCTGT CCTTGTTCACTGCAAAAATTC TGGGAAGCCATTTTATGC
PHYC_ARABIDOPSIS   -I--S--I--L--N--P--I- -T--L--H--C--R--S--S- -S--K--P--F--Y--A-
phyc_arabidopsis   GATTAGTATTTTGAATCCTAT CACGCTTCATTGTAGGTCTTC AAGTAAGCCTTTTTATGC

MOUGEOTIA          -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
mougeotia          ???????????????????? ???????????????????? ????????????????????
CHARA              -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
chara              ???????????????????? ???????????????????? ????????????????????
MARCHANTIA         -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
marchantia         ???????????????????? ???????????????????? ????????????????????
FUNARIA            -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
funaria            ???????????????????? ???????????????????? ????????????????????
PHYSCOMITRELLA     -I--A--H--R--I--D--I --G--I--V--I--D--F-- E--A--V--K--T--D--D-
physcomitrella     CATTGCTCATCGCATTGACA TAGGTATAGTGATCGACTTT GAGGCAGTAAAAACAGACGA
CERATODON          -I--A--H--R--I--D--I --G--I--V--I--D--F-- E--A--V--K--M--I--D-
ceratodon          CATTGCCCATCGCATAGACA TTGGTATAGTCATTGACTTT GAGGCGGTGAAAATGATTGA
SELAGINELLA        -I--V--H--R--I--D--V --G--L--V--M--D--L-- E--P--V--K--A--S--D-
selaginella        CATCGTGCACCGCATCGACG TCGGCCTGGTCATGGACCTG GATCCCGTCAAGGCCAGCGA
EQUISETUM          -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
equisetum          ???????????????????? ???????????????????? ????????????????????
MARSILEA           -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
marsilea           ???????????????????? ???????????????????? ????????????????????
ADIANTUM           -I--L--H--R--N--D--V --G--L--V--I--D--L-- E--P--I--R--P--D--D-
adiantum           AATTTTGCACAGAAATGATG TGGGGTTGGTAATTGACCTT GAGCCTATCAGGCCTGATGA
PSILOTUM           -I--V--H--R--I--D--V --G--L--V--I--D--F-- E--P--L--R--S--T--D-
psilotum           CATTGTCCATCGCATAGACG TGGGATTGGTGATTGACTTC GAGCCTTTGAGGTCTACTGA
PSILOTUM2          -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
```

psilotum2	????????????????????	????????????????????	????????????????????
METASEQUOIA	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
metasequoia	????????????????????	????????????????????	????????????????????
PSEUDOTSUGA	-I--L--H--R--I--D--V	--A--L--V--M--D--F--	E--P--I--K--P--S--D-
PSEUDOTSUGA	TATACTTCACAGAATTGATG	TGGCTTTGGTGATGGATTTT	GAGCCTATTAAGCCTTCGGA
PICEA	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
picea	????????????????????	????????????????????	????????????????????
EPHEDRA	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
ephedra	????????????????????	????????????????????	????????????????????
GNETUM	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
gnetum	????????????????????	????????????????????	????????????????????
PHYC_PIPER	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
phyc_piper	????????????????????	????????????????????	????????????????????
PHYB_PIPER	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
phyb_piper	????????????????????	????????????????????	????????????????????
PHYA_PIPER	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
phya_piper	????????????????????	????????????????????	????????????????????
PHYB_ARABIDOPSIS	-I--L--H--R--I--D--V	--G--V--V--I--D--L--	E--P--A--R--T--E--D-
phyb_arabidopsis	CATTCTTCATAGGATTGATG	TTGGTGTTGTTATTGATTTA	GAGCCAGCTAGAACTGAAGA
PHYB_GLYCINE	-I--L--H--R--I--D--V	--G--I--V--I--D--L--	E--P--A--R--T--E--D-
phyb_glycine	AATCCTCCACCGAATTGACG	TCGGAATTGTCATCGATTTG	GAGCCTGCGCGTACGGAGGA
PHYB_NICOTIANA	-I--L--H--R--V--D--V	--G--I--V--I--D--L--	E--P--A--R--T--E--D-
phyb_nicotiana	AATTTTGCATAGGGTTGATG	TCGGGATTGTAATTGATTTG	GAGCCTGCTAGAACAGAGGA
PHYB_SOLANUM	-I--L--H--R--V--D--V	--G--I--V--I--D--L--	E--P--A--R--T--E--D-
phyb_solanum	AATTTTGCACAGGGTTGATG	TTGGTATTGTCATTGATTTG	GAGCCTGCTAGAACTGAGGA
PHYD_ARABIDOPSIS	-I--L--H--R--V--D--V	--G--I--L--I--D--L--	E--P--A--R--T--E--D-
phyd_arabidopsis	GATTCTCCACAGGGTTGATG	TTGGAATTTTGATCGATTTA	GAGCCGGCTCGAACCGAAGA
PHYE_ARABIDOPSIS	-I--L--H--R--I--D--A	--G--I--V--M--D--L--	E--P--A--K--S--G--D-
phye_arabidopsis	TATTCTTCACAGGATTGATG	CAGGGATTGTCATGGATTTG	GAGCCTGCTAAATCAGGTGA
PHYB_ORYZA	-I--L--H--R--I--D--V	--G--V--V--I--D--L--	E--P--A--R--T--E--D-
phyb_oryza	CATCCTCCACCGCATCGATG	TCGGCGTCGTCATCGACCTC	GAGCCCGCCCGCACCGAGGA
PHYA_AVENA	-I--V--H--R--A--T--G	--C--L--V--V--D--F--	E--P--V--K--P--T--E-
phya_avena	CATTGTTCATCGAGCAACTG	GTTGTTTGGTGGTAGACTTT	GAGCCTGTAAAGCCTACAGA
PHYA_ORYZA	-I--V--H--R--A--T--G	--C--L--V--V--D--F--	E--P--V--K--P--T--E-
phya_oryza	CATTGTTCATCGGGCAACTG	GTTGTTTGGTGGTAGACTTT	GAGCCTGTGAAACCTACAGA
PHYA_ZEA	-I--V--H--R--A--T--G	--C--L--V--V--D--F--	E--P--V--K--P--T--E-
phya_zea	CATTGTTCATAGGGCAACTG	GTTGTCTGGTGGTAGATTTT	GAGCCTGTGAAGCCTACAGA
PHYA_ARABIDOPSIS	-I--I--H--R--V--T--G	--S--I--I--I--D--F--	E--P--V--K--P--Y--E-
phya_arabidopsis	GATTATCCACAGGGTTACAG	GGAGCATCATCATCGACTTT	GAACCCGTGAAGCCTTATGA
PHYA_CUCURBITA	-I--V--H--R--V--T--G	--S--L--I--I--D--F--	E--P--V--K--P--Y--E-
phya_cucurbita	AATTGTTCATCGTGTTACTG	GAAGCTTAATCATTGACTTT	GAGCCTGTGAAGCCTTATGA
PHYA_PISUM	-I--I--H--R--V--T--G	--S--L--I--I--D--F--	E--P--V--K--P--Y--E-
phya_pisum	GATCATTCATCGTGTTACTG	GTAGTTTGATCATTGACTTT	GAGCCGGTGAAGCCTTATGA
PHYA_GLYCINE	-I--V--H--R--V--T--G	--S--L--I--V--D--F--	E--P--V--K--P--Y--E-
phya_glycine	AATTGTCCATCGCGTCACTG	GTAGTTTGATTGTTGACTTT	GAGCCAGTCAAGCCTTATGA
PHYA_NICOTIANA	-I--V--H--R--V--T--G	--S--L--I--I--D--F--	E--P--V--K--P--Y--E-
phya_nicotiana	AATTGTTCATAGGGTTACTG	GTAGCTTAATCATTGATTTC	GAGCCTGTGAAGCCCTATGA
PHYA_SOLANUM	-I--V--H--R--V--T--G	--S--L--I--I--D--F--	E--P--V--K--P--Y--E-
phya_solanum	AATTGTTCATAGGGTTACAG	GTAGCTTAATGATTGATTTT	GAGCCTGTGAAGCCCTATGA
PHYC_ARABIDOPSIS	-I--L--H--R--I--E--E	--G--L--V--I--D--L--	E--P--V--S--P--D--E-
phyc_arabidopsis	GATTCTGCATCGGATTGAGG	AAGGTCTTGTTATAGATTTG	GAGCCTGTGAGTCCTGATGA
MOUGEOTIA	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
mougeotia	????????????????????	????????????????????	????????????????????
CHARA	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
chara	????????????????????	????????????????????	????????????????????
MARCHANTIA	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
marchantia	????????????????????	????????????????????	????????????????????
FUNARIA	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
funaria	????????????????????	????????????????????	????????????????????
PHYSCOMITRELLA	-H--L--V-----S--A--A	--G--A--L--Q--S--H--	K--L--A--A--K--A--I-
physcomitrella	TCATTTAGTT---TCAGCTG	CTGGTGCACTGCAATCTCAT	AAGCTTGCAGCCAAGGCAAT
CERATODON	-V--P--V--S--A--A--A	--G--A--L--Q--S--H--	K--L--A--A--R--A--I-
ceratodon	TGTTCCAGTTTCAGCTGCTG	CCGGTGCACTGCAATCTCAC	AAACTTGCGGCCCGGGCTAT
SELAGINELLA	-T--R--V--G--S--A--A	--G--A--L--Q--S--H--	K--L--A--A--K--A--I-
selaginella	CACGAGGGTGGGATCCGCTG	CCGGGGCTCTGCAGTCCCAC	AAGCTGGCGGCCAAGGCCAT
EQUISETUM	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
equisetum	????????????????????	????????????????????	????????????????????
MARSILEA	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
marsilea	????????????????????	????????????????????	????????????????????
ADIANTUM	----A--S--I--T-----G	--G--A--L--Q--S--H--	K--L--A--A--K--A--I-
adiantum	T---GCATCTATCACT---G	GCGGTGCCTTGCAGTCACAT	AAGCTGGCAGCGAAGGCCAT
PSILOTUM	-L--A--A--F--S--A--A	--G--I--L--Q--S--H--	K--L--A--A--K--A--I-

```
psilotum            TTTAGCTGCCTTCTCGGCTG CGGGAATATTGCAATCTCAC AAACTTGCCGCCAAAGCAAT
PSILOTUM2           -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
psilotum2           ???????????????????? ???????????????????? ????????????????????
METASEQUOIA         -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
metasequoia         ???????????????????? ???????????????????? ????????????????????
PSEUDOTSUGA         ----V--S--T--S--A--M --G--A--L--Q--S--Y-- K--L--A--A--K--A--I-
PSEUDOTSUGA         T---GTTTCTACTTCTGCCA TGGGAGCATTGCAATCCTAT AAGCTAGCTGCAAAGGCCAT
PICEA               -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
picea               ???????????????????? ???????????????????? ????????????????????
EPHEDRA             -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
ephedra             ???????????????????? ???????????????????? ????????????????????
GNETUM              -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
gnetum              ???????????????????? ???????????????????? ????????????????????
PHYC_PIPER          -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
phyc_piper          ???????????????????? ???????????????????? ????????????????????
PHYB_PIPER          -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
phyb_piper          ???????????????????? ???????????????????? ????????????????????
PHYA_PIPER          -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
phya_piper          ???????????????????? ???????????????????? ????????????????????
PHYB_ARABIDOPSIS    ----P--A--L--S--I--A --G--A--V--Q--S--Q-- K--L--A--V--R--A--I-
phyb_arabidopsis    T---CCTGCGCTTTCTATTG CTGGTGCTGTTCAATCGCAG AAACTCGCGGTTCGTGCGAT
PHYB_GLYCINE        ----P--A--L--S--I--A --G--A--V--Q--S--Q-- E--A--L--V--R--A--I-
phyb_glycine        T---CCTGCCCTCTCTATCG CTGGAGCTGTCCAGTCGCAA GAAGCGCTCGTTCGCGCGAT
PHYB_NICOTIANA      ----P--A--L--S--I--A --G--A--V--Q--S--Q-- K--L--A--V--R--A--I-
phyb_nicotiana      C---CCTGCTTTATCCATTG CTGGCGCAGTGCAGTCACAA AAACTTGCTGTGAGGGCTAT
PHYB_SOLANUM        ----P--A--L--S--I--A --G--A--V--Q--S--Q-- K--L--R--S--E--G--L-
phyb_solanum        C---CCTGCTTTATCTATTG CTGGAGCAGTGCAGTCACAG AAACTTCGCAGTGAGGGCCT
PHYD_ARABIDOPSIS    ----P--A--L--S--I--A --G--A--V--Q--S--Q-- K--L--A--V--R--A--I-
phyd_arabidopsis    T---CCGGCACTTTCAATCG CCGGAGCAGTCCAATCGCAG AAACTTGCGGTACGTGCGAT
PHYE_ARABIDOPSIS    ----P--A--L--T--L--A --G--A--V--Q--S--Q-- K--L--A--V--R--A--I-
phye_arabidopsis    T---CCGGCTTTGACCCTTG CAGGCGCAGTTCAGTCTCAG AAGCTAGCCGTTAGGGCCAT
PHYB_ORYZA          ----P--A--L--S--I--A --G--A--V--Q--S--Q-- K--L--V--V--R--A--I-
phyb_oryza          T---CCTGCACTCTCCATCG CTGGCGCAGTCCAGTCTCAG AAGCTCGTGGTCCGTGCCAT
PHYA_AVENA          -F--P--A-----T--A--A --G--A--L--Q--S--Y-- K--L--A--A--K--A--I-
phya_avena          ATTTCCTGCC---ACTGCTG CTGGGGCTTTGCAGTCCTAC AAGCTTGCTGCCAAGGCAAT
PHYA_ORYZA          -F--P--A-----T--A--A --G--A--L--Q--S--Y-- K--L--A--A--K--A--I-
phya_oryza          ATTTCCAGCA---ACTGCCG CTGGGGCTTTGCAATCTTAC AAACTTGCTGCCAAGGCAAT
PHYA_ZEA            -F--P--A-----T--A--A --G--A--L--Q--S--Y-- K--L--A--A--K--A--I-
phya_zea            ATTTCCTGCC---ACTGCTG CTGGGGCTTTGCAGTCCTAC AAGCTTGCTGCCAAGGCAAT
PHYA_ARABIDOPSIS    ----V--P--M--T--A--A --G--A--L--Q--S--Y-- K--L--A--A--K--A--I-
phya_arabidopsis    A---GTCCCCATGACAGCTG CTGGTGCCTTACAATCATAC AAGCTCGCTGCCAAAGCAAT
PHYA_CUCURBITA      ----G--P--V--T--A--A --G--A--L--Q--S--Y-- K--L--A--A--K--A--I-
phya_cucurbita      A---GGTCCAGTGACTGCAG CTGGAGCTCTACAATCATAT AAACTTGCCGCCAAAGCGAT
PHYA_PISUM          ----V--P--M--T--A--A --G--A--L--Q--S--Y-- K--L--A--A--K--A--I-
phya_pisum          A---GTTCCCATGACTGCTG CGGGTGCCTTGCAATCTTAC AAACTTGCTGCTAAAGCAAT
PHYA_GLYCINE        ----V--P--M--T--A--A --G--A--L--Q--S--Y-- K--L--A--A--K--A--I-
phya_glycine        A---GTTCCCATGACTGCAG CAGGTGCCTTGCAATCTTAC AAGCTTGCTGCCAAAGCAAT
PHYA_NICOTIANA      ----V--P--M--T--A--A --G--A--L--Q--S--Y-- K--L--A--A--K--A--I-
phya_nicotiana      A---GTACCCATGACTGCTG CAGGGGCCCTGCAGTCGTAT AAACTTGCAGCCAAAGCCAT
PHYA_SOLANUM        ----V--P--M--T--A--A --G--A--L--Q--S--Y-- K--L--A--A--K--A--I-
phya_solanum        A---GTTCCCATGACTGCTG CAGGGGCCCTGCAGTCATAT AAACTTGCAGCCAAAGCCAT
PHYC_ARABIDOPSIS    ----V--P--V--T--A--A --G--A--L--R--S--Y-- K--L--A--A--K--S--I-
phyc_arabidopsis    G---GTGCCTGTGACTGCTG CCGGGGCTTTAAGATCGTAT AAGCTTGCGGCGAAATCGAT

MOUGEOTIA           -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
mougeotia           ???????????????????? ???????????????????? ????????????????????
CHARA               -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
chara               ???????????????????? ???????????????????? ????????????????????
MARCHANTIA          -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
marchantia          ???????????????????? ???????????????????? ????????????????????
FUNARIA             -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
funaria             ???????????????????? ???????????????????? ????????????????????
PHYSCOMITRELLA      ----T--R--L--Q--A--L --P--G--G--N--I--G-- L--L--C--D--T--V--V-
physcomitrella      C---ACACGGCTTCAAGCAT TACCTGGAGGTAACATAGGA TTGCTCTGTGACACTGTTGT
CERATODON           ----T--R--L--Q--A--L --P--G--G--D--I--E-- L--L--C--D--T--I--V-
ceratodon           T---ACACGACTTCAAGCAT TACCTGGAGGCGACATAGAG TTGCTTTGTGATACTATTGT
SELAGINELLA         ----S--R--L--Q--S--L --P--G--G--D--I--G-- L--L--C--D--T--V--V-
selaginella         A---TCGCGCCTCCAGTCGC TGCCGGGCGGCGACATCGGC CTGTTGTGCCACACGGTGGT
EQUISETUM           -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
equisetum           ???????????????????? ???????????????????? ????????????????????
MARSILEA            -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
marsilea            ???????????????????? ???????????????????? ????????????????????
ADIANTUM            ----A--R--L--Q--S--L --P--G--G--D--I--G-- L--L--C--D--S--V--V-
```

adiantum	C---GCTCGGTTGCAGTCTT	TGCCAGGAGGTGATATTGGC	CTCTTATGTGACTCAGTAGT
PSILOTUM	----S--R--L--Q--S--L	--P--V--G--D--I--G--	L--L--C--D--T--V--V-
psilotum	T---TCGAGGTTGCAATCTT	TACCGGTTGGGGATATCGGT	CTTCTCTGCGATACCGTTGT
PSILOTUM2	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
psilotum2	????????????????????	????????????????????	????????????????????
METASEQUOIA	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
metasequoia	????????????????????	????????????????????	????????????????????
PSEUDOTSUGA	----S--R--L--Q--S--I	--P--S--G--N--I--G--	L--L--C--D--M--I--A-
pseudotsuga	A---TCGCGGTTGCAATCGA	TCCCTAGTGGAAATATTGGG	CTACTGTGTGACATGATTGC
PICEA	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
picea	????????????????????	????????????????????	????????????????????
EPHEDRA	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
ephedra	????????????????????	????????????????????	????????????????????
GNETUM	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
gnetum	????????????????????	????????????????????	????????????????????
PHYC_PIPER	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
phycc_piper	????????????????????	????????????????????	????????????????????
PHYB_PIPER	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
phyb_piper	????????????????????	????????????????????	????????????????????
PHYA_PIPER	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
phya_piper	????????????????????	????????????????????	????????????????????
PHYB_ARABIDOPSIS	----S--Q--L--Q--A--L	--P--G--G--D--I--K--	L--L--C--D--T--V--V-
phyb_arabidopsis	T---TCTCAGTTACAGGCTC	TTCCTGGTGGAGATATTAAG	CTTTTGTGTGACACTGTCGT
PHYB_GLYCINE	----S--Q--L--Q--S--L	--P--S--A--D--V--K--	L--L--C--D--T--V--V-
phyb_glycine	T---TCGCAGCTTCAATCTC	TCCCAAGCGCTGATGTTAAG	CTTCTCTGTGACACTGTTGT
PHYB_NICOTIANA	----S--H--L--Q--S--L	--P--G--G--D--V--K--	L--L--C--D--T--V--V-
phyb_nicotiana	T---TCTCATTTGCAATCAC	TTCCTGGTGGGGATGTTAAG	CTTTTGTGTGATACTGTGGT
PHYB_SOLANUM	-F--L--I--C--N--H--F	--L--V--G--T--L--K--	L--L--C--D--T--V--V-
phyb_solanum	ATTTCTCATTTGCAATCACT	TCCTGGTGGGGACATTAAAG	CTTTTGTGTGATACTGTTGT
PHYD_ARABIDOPSIS	----S--H--L--Q--S--L	--P--S--G--D--I--K--	L--L--C--D--T--V--V-
phyd_arabidopsis	T---TCTCATTTACAATCGT	TGCCTAGCGGCGACATTAAG	CTTCTATGTGACACTGTTGT
PHYE_ARABIDOPSIS	----S--R--L--Q--S--L	--P--G--G--D--I--G--	A--L--C--D--T--V--V-
phye_arabidopsis	T---TCTAGGCTGCAGTCAC	TTCCCGGAGGAGATATTGGT	GCCTTGTGTGATACTGTTGT
PHYB_ORYZA	----S--R--L--Q--A--L	--P--G--G--D--V--K--	L--L--C--D--T--V--V-
phyb_oryza	C---TCCCGCCTCCAGGCCC	TTCCCGGCGGTGACGTCAAG	CTCCTTTGCGACACCGTTGT
PHYA_AVENA	----S--K--I--Q--S--L	--P--G--G--S--M--E--	V--L--C--N--T--V--V-
phya_avena	A---TCCAAGATCCAGTCAT	TGCCAGGTGGAAGCATGGAG	GTGCTATGCAATACTGTGGT
PHYA_ORYZA	----S--K--I--Q--S--L	--P--G--G--S--M--E--	V--L--C--N--T--V--V-
phya_oryza	C---TCTAAGATCCAGTCAC	TGCCAGGTGGAAGCATGGAG	GTGCTATGCAATACGGTGGT
PHYA_ZEA	----S--K--I--Q--S--L	--P--G--G--S--M--E--	A--L--C--N--T--V--V-
phya_zea	C---TCCAAGATCCAGTCAC	TACCAGGTGGAAGCATGGAG	GCCTTATGCAATACCGTGGT
PHYA_ARABIDOPSIS	----T--R--L--Q--S--L	--P--S--G--S--M--E--	R--L--C--D--T--M--V-
phya_arabidopsis	C---ACTAGGCTGCAATCTT	TACCCAGCGGGAGTATGGAA	AGGCTTTGTGATACAATGGT
PHYA_CUCURBITA	----T--R--L--Q--S--L	--P--S--G--S--M--A--	R--L--C--D--T--M--V-
phya_cucurbita	T---ACTAGATTGCAGTCTT	TGCCTAGTGGAAGCATGGCT	AGGCTTTGTGACACAATGGT
PHYA_PISUM	----T--R--L--Q--S--L	--A--S--G--S--M--E--	R--L--C--D--T--M--V-
phya_pisum	T---ACAAGATTGCAATCTT	TGGCTAGTGGCAGCATGGAA	AGGCTTTGTGATACCATGGT
PHYA_GLYCINE	----T--R--L--Q--S--L	--P--S--G--N--M--E--	R--L--C--D--T--M--V-
phya_glycine	T---ACCCGATTGCAATCAT	TGCCTAGTGGGAACATGGAA	AGACTATGTGATACTATGGT
PHYA_NICOTIANA	----T--R--L--Q--A--L	--P--S--G--S--M--E--	R--L--C--D--T--M--V-
phya_nicotiana	T---ACTCGCTTGCAGGCCT	TGCCCAGCGGCAGTATGGAA	AGACTTTGTGACACTATGGT
PHYA_SOLANUM	----T--R--L--Q--S--L	--P--S--G--S--M--E--	R--L--C--D--T--M--V-
phya_solanum	T---ACTCGCTTGCAGTCCT	TGCCCAGTGGCAGTATGGAA	AGACTTTGTGACACAATGGT
PHYC_ARABIDOPSIS	----S--R--L--Q--A--L	--P--S--G--N--M--L--	L--L--C--D--A--L--V-
phyc_arabidopsis	T---TCGAGGTTGCAGGCAT	TGCCTAGTGGGAATATGTTG	TTGTTGTGTGATGCTTTGGT

MOUGEOTIA	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--D--D-
mougeotia	????????????????????	????????????????????	???????????????GATGA
CHARA	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--D-
chara	????????????????????	????????????????????	??????????????????GA
MARCHANTIA	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--D-
marchantia	????????????????????	????????????????????	??????????????????GA
FUNARIA	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--D-
funaria	????????????????????	????????????????????	??????????????????GA
PHYSCOMITRELLA	-E--E--V--R--E--L--T	--G--Y--D--R--V--M--	A--Y--R--F--H--E--D-
physcomitrella	CGAAGAGGTTCGGGAGCTCA	CTGGTTATGACAGGGTAATG	GCTTATAGATTTCATGAGGA
CERATODON	-E--E--V--R--E--L--T	--G--Y--D--R--V--M--	A--F--K--F--H--E--D-
ceratodon	TGAGGAGGTGCGGGAACTTA	CTGGGTATGACAGGGTGATG	GCTTTTAAATTTCATGAAGA
SELAGINELLA	-E--E--V--R--D--V--T	--G--Y--D--L--V--M--	A--Y--K--F--H--E--D-
selaginella	GGAGGAGGTGCGGGACGTCA	CCGGTTACGACCTGGTCATG	GCATACAAGTTTCACGAAGA
EQUISETUM	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--D-
equisetum	????????????????????	????????????????????	??????????????????GA

```
MARSILEA           -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--D-
marsilea           ???????????????????? ???????????????????? ??????????????????GA
ADIANTUM           -E--E--V--H--E--L--T --G--F--D--R--V--M-- A--Y--K--F--H--E--D-
adiantum           GGAGGAGGTACATGAGCTTA CGGGTTTCGACAGAGTGATG GCATATAAGTTTCATGAAGA
PSILOTUM           -E--E--V--R--Q--L--T --G--Y--D--R--V--M-- A--Y--K--F--H--E--D-
psilotum           TGAGGAAGTTAGGCAGCTCA CCGGTTATGATAGGGTGATG GCATATAAATTTCATGAAGA
PSILOTUM2          -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
psilotum2          ???????????????????? ???????????????????? ????????????????????
METASEQUOIA        -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--D-
metasequoia        ???????????????????? ???????????????????? ??????????????????GA
PSEUDOTSUGA        -Q--E--V--H--E--L--M --G--Y--D--R--V--M-- A--Y--K--F--H--D--D-
pseudotsuga        ACAAGAGGTGCATGAGCTTA TGGGTTATGACCGCGTTATG GCTTATAAATTTCACGATGA
PICEA              -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--D-
picea              ???????????????????? ???????????????????? ??????????????????GA
EPHEDRA            -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--D-
ephedra            ???????????????????? ???????????????????? ??????????????????GA
GNETUM             -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--D-
gnetum             ???????????????????? ???????????????????? ??????????????????GA
PHYC_PIPER         -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--D-
phycc_piper        ???????????????????? ???????????????????? ??????????????????GA
PHYB_PIPER         -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--D-
phyb_piper         ???????????????????? ???????????????????? ??????????????????GA
PHYA_PIPER         -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--D-
phya_piper         ???????????????????? ???????????????????? ??????????????????GA
PHYB_ARABIDOPSIS   -E--S--V--R--D--L--T --G--Y--D--R--V--M-- V--Y--K--F--H--E--D-
phyb_arabidopsis   GGAAAGTGTGAGGGACTTGA CTGGTTATGATCGTGTTATG GTTTATAAGTTTCATGAAGA
PHYB_GLYCINE       -E--S--V--R--E--L--T --G--Y--D--R--V--M-- V--Y--K--F--H--E--D-
phyb_glycine       GGAAAGTGTTAGGGAATTGA CGGGTTATGATAGGGTTATG GTTTATAAGTTTCATGAGGA
PHYB_NICOTIANA     -E--S--V--R--E--L--T --G--Y--D--R--V--M-- V--Y--K--F--H--E--D-
phyb_nicotiana     TGAGAGTGTGAGGGAGTTAA CTGGGTATGATCGGGTTATG GTATATAAATTTCATGAGGA
PHYB_SOLANUM       -E--S--V--R--E--L--T --G--Y--D--R--V--M-- V--Y--K--F--H--E--D-
phyb_solanum       TGAGAGTGTCAGGGAGTTAA CCGGGTATGACCGGGTTATG GTATATAAATTTCATGAGGA
PHYD_ARABIDOPSIS   -E--S--V--R--D--L--T --G--Y--D--R--V--M-- V--Y--K--F--H--E--D-
phyd_arabidopsis   GGAAAGCGTTAGAGATCTTA CTGGCTACGACCGCGTTATG GTGTACAAGTTTCATGAAGA
PHYE_ARABIDOPSIS   -E--D--V--Q--R--L--T --G--Y--D--R--V--M-- V--Y--Q--F--H--E--D-
phye_arabidopsis   GGAAGATGTTCAGAGACTTA CCGGTTATGACCGTGTTATG GTCTATCAGTTTCATGAAGA
PHYB_ORYZA         -E--H--V--R--E--L--T --G--Y--D--R--V--M-- V--Y--R--F--H--E--D-
phyb_oryza         TGAGCATGTTAGAGAGCTCA CAGGTTATGACCGCGTTATG GTGTACAGGTTCCATGAGGA
PHYA_AVENA         -K--E--V--F--D--L--T --G--Y--D--R--V--M-- A--Y--K--F--H--E--D-
phya_avena         GAAGGAAGTCTTTGACCTTA CCGGGTATGACAGGGTTATG GCTTACAAGTTTCATGAAGA
PHYA_ORYZA         -K--E--L--F--D--L--T --G--Y--D--R--V--M-- A--Y--K--F--H--E--D-
phya_oryza         CAAGGAACTCTTTGACCTCA CAGGATATGATAGAGTTATG GCTTATAAGTTCCATGAAGA
PHYA_ZEA           -K--E--V--F--D--L--T --G--Y--D--R--V--M-- A--Y--K--F--H--E--D-
phya_zea           TAAGGAAGTCTTTGACCTGA CAGGTTATGACAGGGTTATG GCTTACAAGTTCCATGAAGA
PHYA_ARABIDOPSIS   -Q--E--V--F--E--L--T --G--Y--D--R--V--M-- A--Y--K--F--H--E--D-
phya_arabidopsis   TCAAGAGGTTTTTGAACTCA CGGGGTATGACAGGGTGATG GCTTATAAGTTTCATGAAGA
PHYA_CUCURBITA     -Q--E--V--F--E--L--T --G--Y--D--R--V--M-- A--Y--K--F--H--D--D-
phya_cucurbita     TCAAGAAGTTTTTGAACTAA CAGGTTATGATCGAGTTATG GCTTATAAATTCCATGATGA
PHYA_PISUM         -Q--E--V--F--E--L--T --G--Y--D--R--V--M-- A--Y--K--F--H--E--D-
phya_pisum         TCAAGAAGTTTTTGAACTAA CGGGTTATGACAGGGTGATG GCTTATAAATTTCACGAGGA
PHYA_GLYCINE       -Q--E--V--F--E--L--T --G--Y--D--R--V--M-- A--Y--K--F--H--E--D-
phya_glycine       TCAGGAAGTTTTTGAACTCA CAGGTTATGATAGGGTGATG GCTTATAAATTTCATGAGGA
PHYA_NICOTIANA     -Q--E--V--F--E--L--T --G--Y--D--R--V--M-- T--Y--K--F--H--D--D-
phya_nicotiana     TCAGGAGGTTTTCGAACTCA CAGGCTATGACAGGGTGATG ACGTATAAGTTTCACGATGA
PHYA_SOLANUM       -Q--E--V--F--E--L--T --G--Y--D--R--V--M-- G--Y--K--F--H--D--D-
phya_solanum       TCAGGAGGTTTTTGAACTCA CAGGTTATGACAGGGTGATG GGATATAAGTTTCACGATGA
PHYC_ARABIDOPSIS   -K--E--V--S--E--L--T --G--Y--D--R--V--M-- V--Y--K--F--H--E--D-
phyc_arabidopsis   TAAGGAAGTTAGTGAATTAA CTGGTTATGATAGGGTGATG GTGTATAAGTTCCATGAGGA

MOUGEOTIA          -E--H--G--E--V--V--A --E--I--R--R--S--D-- L--E--P--Y--L--G--L-
mougeotia          CGAACATGGTGAAGTTGTGG CAGAGATTCGAAGGTCTGAC CTTGAGCCATACCTGGGTCT
CHARA              -E--H--G--E--V--I--A --E--I--R--R--A--D-- L--E--P--Y--L--G--L-
chara              CGAACATGGGGAGGTGATTG CAGAGATTCGCAGGGCGGAT CTTGAACCCTACCTTGGACT
MARCHANTIA         -E--H--G--E--V--V--A --E--I--R--R--S--D-- L--E--P--Y--V--G--L-
marchantia         TGAACACGGAGAGGTTGTGG CGGAAATAAGGCGGTCTGAT CTAGAACCTTATGTTGGATT
FUNARIA            -E--H--G--E--V--V--A --E--I--R--R--T--D-- L--E--P--Y--L--G--L-
funaria            TGAGCATGGGGAGGTCGTTG CAGAAATTCGTCGCACAGAT CTTGAGCCGTACTTGGGTCT
PHYSCOMITRELLA     -E--H--G--E--V--V--A --E--I--R--R--A--D-- L--E--P--Y--L--G--L-
physcomitrella     TGAGCATGGAGAGGTTGTGG CTGAGATACGTCGCGCGGAT CTTGAGCCTTACTTAGGTCT
CERATODON          -E--H--G--E--V--V--A --E--I--R--R--M--D-- L--E--P--Y--M--G--L-
ceratodon          TGAGCATGGCGAAGTTGTGG CAGAAATACGTCGCATGGAT CTTGAGCCCTATATGGGTCT
SELAGINELLA        -E--H--G--E--V--V--A --E--I--R--R--S--D-- L--E--P--Y--L--G--L-
selaginella        CGAGCATGGCGAGCTGCTGG CGGAGATTCGGCGCTCCGAC CTCGAACCGTACCTCGGGCT
```

```
EQUISETUM          -E--H--G--E--V--V--A --E--I--R--R--A--D-- L--E--P--Y--L--G--L-
equisetum          CGAGCATGGTGAGGTTGTCG CTGAGATACGCCGTGCCGAC CTTGAGCCCTACCTCGGGCT
MARSILEA           -E--H--G--E--V--V--A --E--I--R--K--S--D-- L--E--P--Y--L--G--L-
marsilea           TGAACATGGTGAAGTTGTTG CAGAGATAAGGAAGTCTGAC TTAGAGCCATACCTTGGCTT
ADIANTUM           -E--H--G--E--V--V--A --E--I--R--R--T--D-- L--E--P--Y--I--G--L-
adiantum           TGAGCATGGAGAGGTGGTTG CTGAGATAAGGCGTACGGAT CTTGAGCCTTATATTGGGCT
PSILOTUM           -E--H--G--E--V--V--A --E--I--R--R--S--D-- L--E--P--F--V--G--I-
psilotum           TGAACACGGTGAGGTGGTTG CGGAGATACGTCGTTCAGAT CTGGAGCCTTTTGTCGGTAT
PSILOTUM2          -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
psilotum2          ???????????????????? ???????????????????? ????????????????????
METASEQUOIA        -E--H--G--E--V--V--A --E--I--R--R--S--D-- L--E--P--Y--L--G--L-
metasequoia        TGAGCACGGGGAGGTTGTTG CAGAAATCAGGCGCTCTGAC TTGGAGCCCTACCTTGGGCT
PSEUDOTSUGA        -D--H--G--E--V--I--S --E--I--R--R--S--D-- L--E--P--Y--L--G--L-
pseudotsuga        TGATCATGGTGAGGTGATAT CTGAGATTAGGAGATCTGAT TTAGAACCTTATTTGGGTCT
PICEA              -E--H--G--E--V--V--A --E--I--R--R--S--D-- L--E--P--Y--L--G--L-
picea              CGAACACGGGGAAGTTGTTG CGGAAATCAGGCGTTCAGAC TTGGAGCCCTATCTTGGGTT
EPHEDRA            -D--H--G--E--V--I--S --E--I--R--R--S--D-- L--E--P--Y--L--G--L-
ephedra            TGACCATGGAGAAGTTATCT CGGAGATCAGGAGATCTGAC CTCGAGCCTTATCTCGGCCT
GNETUM             -E--H--G--E--V--V--A --E--I--R--R--S--D-- L--E--P--Y--L--G--L-
gnetum             CGAGCATGGTGAAGTTGTTG CAGAAATCAGGCGCTCTGAT CTAGAGCCTTATCTGGGTTT
PHYC_PIPER         -E--H--G--E--V--I--A --E--C--K--R--S--D-- L--E--P--Y--L--G--L-
phyc_piper         TGAGCATGGGGAGGTGATAG CCGAGTGCAAGAGGTCGGAT TTGGAGCCTTACCTTGGCCT
PHYB_PIPER         -Q--H--G--E--V--M--A --E--S--K--R--A--D-- L--E--P--Y--I--G--L-
phyb_piper         CCAGCACGGGGAGGTGATGG CCGAGAGCAAGAGGGCGGAC CTGGAACCGTACATCGGCCT
PHYA_PIPER         -D--H--G--E--V--I--S --E--V--T--A--P--D-- L--D--P--Y--L--G--L-
phya_piper         CGACCACGGGGAGGTCATCT CCGAGGTGACGGCCCCTGAT TTGGACCCTTACCTGGGCCT
PHYB_ARABIDOPSIS   -E--H--G--E--V--V--A --E--S--K--R--D--D-- L--E--P--Y--I--G--L-
phyb_arabidopsis   TGAGCATGGAGAAGTTGTAG CTGAGAGTAAACGAGACGAT TTAGAGCCTTATATTGGACT
PHYB_GLYCINE       -E--H--G--E--V--V--S --E--S--K--R--P--D-- L--E--P--Y--I--G--L-
phyb_glycine       TGAGCATGGAGAGGTTGTTT CTGAGAGTAAGAGGCCTGAT TTGGAGCCTTACATTGGATT
PHYB_NICOTIANA     -E--H--G--E--V--V--A --E--S--K--I--P--D-- L--E--P--Y--I--G--L-
phyb_nicotiana     TGAGCACGGGGAGGTAGTGG CTGAGAGCAAAATACCAGAT TTAGAGCCCTATATTGGTTT
PHYB_SOLANUM       -E--H--G--E--V--V--A --E--S--K--R--S--D-- L--E--P--Y--I--G--L-
phyb_solanum       TGAACATGGAGAGGTAGTGG CTGAGAGTAAAAGATCAGAT TTAGAGCCCTATATCGGTTT
PHYD_ARABIDOPSIS   -E--H--G--E--V--V--A --E--S--K--R--N--D-- L--E--P--Y--I--G--L-
phyd_arabidopsis   TGAACATGGTGAAGTCGTAG CCGAGAGTAAACGGAACGAT TTAGAGCCTTACATTGGTCT
PHYE_ARABIDOPSIS   -D--H--G--E--V--V--S --E--I--R--R--S--D-- L--E--P--Y--L--G--L-
phye_arabidopsis   TGATCATGGTGAAGTTGTTT CTGAGATTAGAAGGTCTGAT TTGGAGCCTTATTTGGGTTT
PHYB_ORYZA         -E--H--G--E--V--V--A --E--S--R--R--S--N-- L--E--P--Y--I--G--L-
phyb_oryza         TGAGCATGGAGAAGTCGTTG CCGAGAGCCGGCGCAGTAAC CTTGAGCCCTACATCGGGTT
PHYA_AVENA         -D--H--G--E--V--F--S --E--I--T--K--P--G-- L--E--P--Y--L--G--L-
phya_avena         TGACCATGGTGAGGTATTCT CCGAAATCACAAAGCCTGGT CTTGAGCCTTATCTAGGCCT
PHYA_ORYZA         -D--H--G--E--V--F--A --E--I--T--K--P--G-- L--E--P--Y--L--G--L-
phya_oryza         TGACCATGGTGAAGTCTTTG CTGAGATCACAAAGCCTGGT CTTGAACCTTATCTTGGCCT
PHYA_ZEA           -E--H--G--E--V--F--A --E--I--T--K--P--G-- I--E--P--Y--I--G--L-
phya_zea           TGAGCATGGGGAGGTCTTCG CTGAGATCACCAAACCTGGT ATTGAGCCCTATATAGGCCT
PHYA_ARABIDOPSIS   -D--H--G--E--V--V--S --E--V--T--K--P--G-- L--E--P--Y--L--G--L-
phya_arabidopsis   TGATCACGGTGAGGTTGTCT CCGAGGTTACAAAACCTGGG CTGGAGCCTTATCTTGGGCT
PHYA_CUCURBITA     -D--H--G--E--V--I--S --E--V--A--K--P--G-- L--Q--P--Y--L--G--L-
phya_cucurbita     TGATCATGGGGAAGTGATCT CTGAAGTCGCAAAGCCCGGC CTTCAGCCATATCTTGGTTT
PHYA_PISUM         -D--H--G--E--V--I--A --E--I--A--K--P--G-- L--E--P--Y--L--G--L-
phya_pisum         TGATCACGGGGAGGTGATTG CTGAGATAGCAAAGCCAGGC CTAGAGCCATATCTAGGTCT
PHYA_GLYCINE       -D--H--G--E--V--I--R --E--I--T--K--P--C-- L--E--P--Y--L--G--L-
phya_glycine       TGATCATGGAGAGGTGATTC GTGAGATAACAAAGCCCTGT CTTGAGCCATATCTGGGTTT
PHYA_NICOTIANA     -D--H--G--E--V--V--A --E--I--T--K--P--G-- L--D--P--Y--L--G--L-
phya_nicotiana     TGATCATGGAGAGGTGGTGG CCGAGATCACGAAGCCTGGC CTTGATCCTTACCTTGGTTT
PHYA_SOLANUM       -D--H--G--E--V--V--S --E--I--T--K--P--G-- L--E--P--Y--L--G--L-
phya_solanum       TGATCATGGAGAGGTGGTGT CTGAGATCACAAAGCCTGGC CTCGAGCCTTACCTTGGTTT
PHYC_ARABIDOPSIS   -G--H--G--E--V--I--A --E--C--C--R--E--D-- M--E--P--Y--L--G--L-
phyc_arabidopsis   TGGGCATGGGGAAGTGATTG CTGAATGCTGCCGGGAAGAT ATGGAACCTTATCTTGGGTT

MOUGEOTIA          -H--Y--P--A--T--D--I --P--Q--A--S--R--F-- L--F--I--K--N--R--I-
mougeotia          GCACTATCCTGCAACAGACA TTCCACAGGCATCCAGATTC CTGTTCATAAAGAATCGGAT
CHARA              -H--Y--P--A--T--D--V --P--Q--A--A--R--F-- L--F--M--K--S--R--I-
chara              TCATTACCCTGCGACGGATG TTCCTCAAGCCGCTCGCTTT CTGTTTATGAAGAGTCGCAT
MARCHANTIA         -H--Y--P--A--T--D--I --P--Q--A--S--R--F-- L--F--T--K--N--R--V-
marchantia         GCACTACCCTGCCACTGACA TTCCTCAGGCGTCCCGATTT CTGTTCACGAAAAACAGAGT
FUNARIA            -H--Y--P--A--T--D--I --P--Q--A--S--R--F-- L--F--M--K--N--R--V-
funaria            TCATTACCCGGCCACTGACA TTCCCCAGGCATCTCGGTTT CTATTTATGAAGAATAGGGT
PHYSCOMITRELLA     -H--Y--P--G--T--D--I --P--Q--A--S--R--F-- L--F--M--K--N--K--V-
physcomitrella     TCATTATCCGGGCACTGATA TTCCCCAGGCGTCCCGGTTT CTGTTTATGAAGAACAAGGT
CERATODON          -H--Y--P--A--T--D--I --P--Q--A--S--R--F-- L--L--M--K--N--R--V-
ceratodon          CCATTATCCGGCCACTGACA TTCCCCAGGCGTCCCGTTTT CTGTTAATGAAGAACAGGGT
```

```
SELAGINELLA        -H--Y--P--A--T--D--I --P--Q--A--S--R--F-- L--F--M--K--N--R--V-
selaginella        GCACTACCCAGCCACGGACA TACCCCAGGCTTCTCGATTC CTCTTCATGAAGAACCGGGT
EQUISETUM          -H--Y--P--A--T--D--I --P--Q--A--S--R--F-- L--L--M--K--N--R--V-
equisetum          TCACTACCCGGCCACAGACA TCCCCCAAGCCTCCCGTTTT CTCTTGATGAAGAATCGTGT
MARSILEA           -H--Y--P--A--T--D--I --P--Q--A--S--R--F-- L--F--M--K--N--R--V-
marsilea           ACACTACCCAGCAACTGATA TACCACAAGCATCACGGTTC CTCTTCATGAAAAACAGAGT
ADIANTUM           -H--Y--P--A--T--D--I --P--Q--A--A--R--F-- L--F--M--K--N--R--V-
adiantum           GCATTACCCAGCTACTGATA TCCCACAAGCTGCTCGTTTT CTCTTCATGAAAAACAGAGT
PSILOTUM           -H--Y--P--A--T--D--I --P--Q--A--C--R--F-- L--F--L--K--N--R--V-
psilotum           ACATTATCCAGCCACGGACA TTCCCCAGGCTTGTCGTTTT CTGTTCTTGAAGAACAGAGT
PSILOTUM2          -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
psilotum2          ???????????????????? ???????????????????? ????????????????????
METASEQUOIA        -H--Y--P--A--T--D--I --P--Q--A--S--R--F-- L--F--M--Q--N--R--V-
metasequoia        GCATTATCCTGCCACTGATA TTCCGCAGGCATCCCGATTT CTTTTCATGCAGAACAGGGT
PSEUDOTSUGA        -H--Y--P--A--T--D--I --P--Q--A--S--R--F-- L--F--M--R--N--R--V-
pseudotsuga        ACACTACCCCGCAACTGATA TTCCTCAAGCTTCAAGATTT TTGTTTATGAGGAACAGAGT
PICEA              -H--Y--P--A--T--G--I --P--Q--A--S--R--F-- L--F--M--R--T--R--A-
picea              ACATTACCCGGCCACCGGTA TCCCTCAGGCTTCTCGCTTT CTTTTTATGCGGACAAGGGC
EPHEDRA            -H--Y--P--A--T--D--V --P--Q--A--S--R--F-- L--F--M--K--N--R--V-
ephedra            TCATTATCCTGCTACTGATG TCCCTCAGGCGTCGCGGTTT TTGTTTATGAAAAACAGGGT
GNETUM             -H--Y--P--A--T--D--I --P--Q--A--S--R--F-- L--F--M--Q--N--R--V-
gnetum             GCATTATCCTGCCACCGACA TTCCTCAGGCATCTAGATTT CTCTTCATGCAAAACAGAGT
PHYC_PIPER         -H--Y--P--A--T--D--I --P--Q--A--S--R--F-- L--F--M--K--N--K--V-
phycc_piper        GCATTACCCTGCTACTGATA TACCTCAGGCTTCGAGGTTT TTGTTTATGAAGAATAAGGT
PHYB_PIPER         -H--Y--P--A--T--D--I --P--Q--A--S--R--F-- L--F--K--Q--N--R--V-
phyb_piper         CCACTACCCGGCCACCGACA TCCCCCAAGCGTCGCGTTTC CTTTTCAAGCAGAACCGGGT
PHYA_PIPER         -H--Y--P--A--T--D--I --P--Q--A--A--R--F-- L--F--M--K--N--K--V-
phya_piper         CCACTACCCAGCCACCGACA TTCCCCAGGCCGCCCGTTTC TTGTTCATGAAGAACAAGGT
PHYB_ARABIDOPSIS -H--Y--P--A--T--D--I --P--Q--A--S--R--F-- L--F--K--Q--N--R--V-
phyb_arabidopsis GCATTATCCTGCTACTGATA TTCCTCAAGCGTCAAGGTTC TTGTTTAAGCAGAACCGTGT
PHYB_GLYCINE       -H--Y--P--A--T--D--I --P--Q--A--S--R--F-- L--F--K--Q--N--R--V-
phyb_glycine       GCATTATCCTGCTACTGATA TTCCTCAGGCTTCTAGGTTT TTGTTTAAGCAAAATAGAGT
PHYB_NICOTIANA     -H--Y--P--A--T--D--I --P--Q--A--S--R--F-- L--F--K--Q--N--R--V-
phyb_nicotiana     GCATTATCCTGCTACCGACA TTCCTCAAGCTTCGCGGTTT TTGTTTAAGCAGAACAGGGT
PHYB_SOLANUM       -H--Y--P--A--T--D--I --P--Q--A--S--R--F-- L--F--K--Q--N--R--V-
phyb_solanum       GCATTATCCTGCTACTGATA TTCCTCAAGCTTCACGGTTT TTGTTTAAGCAGAACAGGGT
PHYD_ARABIDOPSIS -H--Y--P--A--T--D--I --P--Q--A--S--R--F-- L--F--K--Q--N--R--V-
phyd_arabidopsis GCATTATCCCGCTACTGATA TTCCTCAGGCATCTCGGTTC TTGTTCAAGCAAAACCGTGT
PHYE_ARABIDOPSIS -H--Y--P--A--T--D--I --P--Q--A--A--R--F-- L--F--K--Q--N--R--V-
phye_arabidopsis ACATTATCCTGCAACAGATA TTCCTCAGGCTGCTCGGTTC TTGTTCAAACAGAACCGTGT
PHYB_ORYZA         -H--Y--P--A--T--D--I --P--Q--A--S--R--F-- L--F--R--Q--N--R--V-
phyb_oryza         GCATTATCCTGCTACAGATA TCCCACAGGCATCACGCTTC CTGTTCCGGCAGAACCGTGT
PHYA_AVENA         -H--Y--P--A--T--D--I --P--Q--A--A--R--L-- L--F--M--K--N--K--V-
phya_avena         GCACTATCCAGCCACTGACA TCCCTCAAGCAGCCAGGCTT CTTTTCATGAAGAACAAAGT
PHYA_ORYZA         -H--Y--P--A--T--D--I --P--Q--A--A--R--F-- L--F--M--K--N--K--V-
phya_oryza         GCATTATCCAGCTACTGATA TCCCTCAGGCAGCCAGGTTT CTTTTCATGAAGAACAAAGT
PHYA_ZEA           -H--Y--P--A--T--D--I --P--Q--A--A--R--F-- L--F--M--K--N--K--V-
phya_zea           GCACTATCCAGCCACTGATA TCCCTCAAGCTGCCAGGTTT CTCTTCATGAAGAACAAAGT
PHYA_ARABIDOPSIS -H--Y--P--A--T--D--I --P--Q--A--A--R--F-- L--F--M--K--N--K--V-
phya_arabidopsis GCATTATCCTGCCACCGACA TCCCTCAAGCAGCCCGTTTT CTGTTTATGAAGAACAAGGT
PHYA_CUCURBITA     -H--Y--P--A--T--D--I --P--Q--A--A--R--F-- L--F--M--K--N--K--V-
phya_cucurbita     GCATTATCCAGCAACTGACA TTCCTCAAGCTGCACGTTTT TTGTTCATGAAAAATAAGGT
PHYA_PISUM         -H--Y--P--A--T--D--I --P--Q--A--A--R--F-- L--F--M--K--N--K--V-
phya_pisum         GCACTATCCGGCGACAGATA TTCCCCAGGCTGCGCGGTTT CTATTTATGAAGAACAAGGT
PHYA_GLYCINE       -H--Y--P--A--T--D--I --P--Q--A--S--R--F-- L--F--R--K--N--K--V-
phya_glycine       GCACTATCCAGCCACCGACA TTCCCCAGGCTTCACGCTTT TTATTTAGGAAGAACAAGGT
PHYA_NICOTIANA     -H--Y--P--A--T--D--I --P--Q--A--A--R--F-- L--F--M--K--N--K--V-
phya_nicotiana     ACACTATCCTGCTACGGATA TCCCACAAGCTGCACGCTTT TTGTTCATGAAGAATAAGGT
PHYA_SOLANUM       -H--Y--P--A--T--D--I --P--Q--A--A--R--F-- L--F--M--K--N--K--V-
phya_solanum       ACATTATCCTGCTACGGATA TTCCACAGGCTGCACGGTTT TTGTTTATGAAGAATAAGGT
PHYC_ARABIDOPSIS -H--Y--S--A--T--D--I --P--Q--A--S--R--F-- L--F--M--R--N--K--V-
phyc_arabidopsis GCATTACTCCGCTACTGATA TACCGCAAGCTTCGAGATTT CTGTTTATGAGAAACAAGGT

MOUGEOTIA          -R--M--I--C--D--C--T --S--P--Q--V--K--V-- V--Q--D--S--R--I--P-
mougeotia          TCGAATGATATGTGATTGCA CCAGCCCTCAGGTGAAGGTG GTGCAGGATTCAAGGATCCC
CHARA              -R--M--I--A--D--C--S --A--P--Q--V--K--V-- L--Q--D--K--S--H--V-
chara              CAGGATGATCGCTGATTGCT CTGCTCCTCAGGTCAAAGTA CTCCAGGACAAGAGCCATGT
MARCHANTIA         -R--M--I--C--D--C--C --A--Q--P--V--Q--V-- I--Q--D--K--E--L--R-
marchantia         GCGTATGATATGTGATTGCT GCGCCCAGCCAGTTCAGGTC ATCCAAGACAAGGAGTTGAG
FUNARIA            -R--M--I--G--D--C--F --A--P--P--V--K--I-- V--Q--D--P--N--L--R-
funaria            CCGGATGATTGGTGATTGTT TTGCACCTCCAGTGAAAATT GTACAAGATCCAAACTTGAG
PHYSCOMITRELLA     -R--I--I--A--D--C--S --A--P--P--V--K--V-- I--Q--D--P--T--L--R-
physcomitrella     GCGGATAATTGCTGATTGTT CCGCACCTCCAGTGAAAGTT ATACAAGATCCAACCTTGAG
```

```
CERATODON          -R--L--I--A--D--C--Y --A--S--P--V--K--L-- I--Q--D--P--D--I--R-
ceratodon          GCGGTTGATAGCTGATTGCT ATGCGTCTCCAGTGAAACTC ATACAAGATCCAGACATTAG
SELAGINELLA        -R--M--I--C--D--C--S --A--P--P--V--K--I-- T--Q--D--K--E--L--R-
selaginella        ACGCATGCTCTGTGACTGCT CCGCCCCGCCCGTGAAGATC ACCCAGGACAAGGAGCTGAG
EQUISETUM          -R--M--I--H--D--C--R --A--P--P--I--R--V-- I--Q--D--E--H--L--P-
equisetum          GCGCATGATCCATGATTGCC GTGCACCCCCCATCCGTGTC ATCCAAGATGAACACCTCCC
MARSILEA           -R--M--I--C--D--C--R --S--N--P--V--R--V-- I--Q--D--K--N--L--R-
marsilea           GAGGATGATATGTGATTGTA GATCCAACCCTGTAAGAGTG ATTCAGGACAAGAATCTCAG
ADIANTUM           -R--M--I--C--D--C--R --L--P--P--V--K--L-- I--Q--D--K--T--L--S-
adiantum           GAGGATGATATGTGATTGCA GGTTGCCGCCAGTTAAGCTT ATTCAGGACAAGACGCTTAG
PSILOTUM           -T--M--I--C--D--C--Y --A--P--P--I--R--I-- I--Q--D--R--Q--L--K-
psilotum           CACGATGATTTGTGACTGTT ATGCTCCCCCTATTAGGATA ATCCAAGACAGGCAACTAAA
PSILOTUM2          -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
psilotum2          ???????????????????? ???????????????????? ????????????????????
METASEQUOIA        -R--M--I--C--D--C--R --A--S--P--V--R--V-- I--Q--A--E--E--L--K-
metasequoia        GCGTATGATCTGTGATTGCC GTGCTAGTCCTGTCAGAGTT ATACAGGCGGAGGAATTAAA
PSEUDOTSUGA        -R--M--I--C--D--C--C --A--P--P--V--I--V-- I--Q--D--K--R--L--P-
pseudotsuga        AAGAATGATTTGTGATTGCT GTGCACCACCTGTAATCGTG ATTCAGGACAAGAGATTGCC
PICEA              -R--M--I--C--D--C--M --A--T--P--V--Q--V-- I--Q--S--E--E--L--M-
picea              GCGAATGATCTGCGATTGCA TGGCAACTCCCGTGCAGGTT ATCCAGTCCGAGGAATTGAT
EPHEDRA            -R--M--I--C--D--C--Y --S--P--P--V--R--V-- I--Q--D--K--R--L--K-
ephedra            GAGAATGATTTGTGATTGTT ACTCTCCACCGGTGAGAGTC ATTCAGGACAAGAGACTGAA
GNETUM             -R--M--I--C--D--C--R --A--N--P--V--N--V-- I--Q--S--A--E--L--K-
gnetum             GAGGATGATTTGTGATTGTC GAGCCAATCCTGTGAATGTG ATACAGTCTGCGGAGCTGAA
PHYC_PIPER         -R--M--I--C--D--C--L --A--P--P--V--K--V-- V--Q--D--K--R--L--G-
phycc_piper        GAGGATGATTTGTGACTGTC TCGCTCCCCCGGTGAAGGTG GTGCAGGACAAACGGCTGGG
PHYB_PIPER         -R--M--I--S--D--C--Y --A--Q--P--V--K--V-- I--Q--D--K--W--L--E-
phyb_piper         GCGGATGATATCTGATTGCT ATGCGCAGCCGGTGAAGGTG ATCCAGGACAAGTGGCTGGA
PHYA_PIPER         -R--L--I--C--D--C--H --A--K--P--V--K--V-- Y--Q--D--E--K--L--P-
phya_piper         CCGCCTGATTTGCGATTGCC ATGCTAAACCCGTCAAGGTG TATCAGGACGAGAAGCTTCC
PHYB_ARABIDOPSIS -R--M--I--V--D--C--N --A--T--P--V--L--V-- V--Q--D--D--R--L--T-
phyb_arabidopsis CCGAATGATAGTAGATTGCA ATGCCACACCTGTTCTTGTG GTCCAGGACGATAGGCTAAC
PHYB_GLYCINE       -R--M--I--V--D--C--H --A--S--A--V--R--V-- V--Q--D--E--A--L--V-
phyb_glycine       TAGGATGATTGTGGATTGTC ATGCTTCTGCTGTGAGGGTG GTGCAGGATGAGGCTCTTGT
PHYB_NICOTIANA     -R--M--I--V--D--C--H --A--T--P--V--R--V-- V--Q--D--E--S--L--M-
phyb_nicotiana     AAGAATGATTGTGGACTGCC ATGCCACCCCTGTGCGGGTT GTTCAGGATGAATCACTGAT
PHYB_SOLANUM       -R--M--I--V--D--C--H --A--T--P--V--R--V-- T--Q--D--E--S--L--M-
phyb_solanum       GAGAATGATTGTGGACTGCC ATGCTACCCCTGTGCGGGTT ACTCAGGATGAATCACTGAT
PHYD_ARABIDOPSIS -R--M--I--V--D--C--Y --A--S--P--V--R--V-- V--Q--D--D--R--L--T-
phyd_arabidopsis TAGGATGATAGTAGATTGCT ATGCGTCACCGGTTCGTGTG GTTCAAGACGATAGGCTCAC
PHYE_ARABIDOPSIS -R--M--I--C--D--C--N --A--T--P--V--K--V-- V--Q--S--E--E--L--K-
phye_arabidopsis CCGAATGATTTGTGACTGCA ATGCAACTCCGGTTAAGGTT GTTCAGAGTGAGGAACTCAA
PHYB_ORYZA         -R--M--I--A--D--C--H --A--A--P--V--R--V-- I--Q--D--P--A--L--T-
phyb_oryza         GCGGATGATTGCTGATTGCC ATGCTGCGCCGGTGAGGGTC ATCCAGGATCCTGCACTAAC
PHYA_AVENA         -R--M--I--C--D--C--R --A--R--S--I--K--V-- I--E--A--E--A--L--P-
phya_avena         ACGGATGATTTGTGATTGCC GTGCGAGATCCATAAAGGTC ATTGAGGCTGAGGCACTCCC
PHYA_ORYZA         -R--M--I--C--D--C--R --A--R--S--I--K--I-- I--E--D--E--S--L--H-
phya_oryza         CCGGATGATTTGTGATTGCC GTGCAAGATCTATCAAGATT ATCGAAGATGAGTCGCTCCA
PHYA_ZEA           -R--M--I--C--D--C--R --A--R--S--V--K--I-- I--E--D--E--A--L--S-
phya_zea           CAGAATGATCTGTGATTGCC GTGCAAGATCCGTGAAGATT ATTGAAGATGAGGCACTCTC
PHYA_ARABIDOPSIS -R--M--I--V--D--C--N --A--K--H--A--R--V-- L--Q--D--E--K--L--S-
phya_arabidopsis CCGGATGATAGTTGATTGCA ATGCAAAACATGCTAGGGTG CTTCAAGATGAAAAGCTTTC
PHYA_CUCURBITA     -R--M--I--V--D--C--R --A--K--H--L--K--V-- L--Q--D--E--K--L--Q-
phya_cucurbita     CCGGATGATTGTTGATTGTC GTGCAAAACATTTAAAAGTA CTCCAAGATGAGAAATTACA
PHYA_PISUM         -R--M--I--V--D--C--N --A--K--H--V--K--V-- L--Q--D--E--K--L--P-
phya_pisum         CCGTATGATAGTTGATTGTA ATGCAAAACATGTGAAGGTT CTTCAAGACGAAAAACTCCC
PHYA_GLYCINE       -R--M--I--V--D--C--H --A--K--H--V--R--V-- L--Q--D--E--K--L--Q-
phya_glycine       TCGTATGATAGTTGACTGTC ATGCAAAACACGTGAGGGTT CTTCAAGATGAAAAACTCCA
PHYA_NICOTIANA     -R--M--I--C--D--C--R --A--K--H--V--K--V-- V--Q--D--E--K--L--P-
phya_nicotiana     CCGAATGATTTGTGATTGCC GAGCAAAACATGTGAAGGTA GTCCAAGATGAGAAGCTTCC
PHYA_SOLANUM       -R--M--I--C--D--C--R --A--K--H--V--K--V-- V--Q--D--E--K--L--P-
phya_solanum       CCGAATGATTTGTGATTGCC GAGCAAAACATGTGAAGGTA GTCCAAGATGAGAAGCTTCC
PHYC_ARABIDOPSIS -R--M--I--C--D--C--S --A--V--P--V--K--V-- V--Q--D--K--S--L--S-
phyc_arabidopsis TAGGATGATTTGTGATTGTT CAGCGGTTCCGGTTAAAGTC GTTCAAGATAAGAGTCTCTC


MOUGEOTIA          -Q--E--M--S--L--A--G --S--T--M--R--G--V-- H--G--C--H--T--Q--Y-
mougeotia          ACAGGAAATGAGTTTGGCAG GGTCAACAATGAGAGGAGTA CATGGCTGCCATACGCAGTA
CHARA              -Q--L--I--S--L--S--G --S--T--I--R--G--V-- H--G--C--H--A--Q--Y-
chara              GCAACTCATCAGCTTGTCAG GATCTACGATTAGAGGCGTG CACGGATGTCATGCGCAGTA
MARCHANTIA         -Q--P--L--S--L--A--G --S--T--L--R--A--L-- H--G--C--H--V--Q--Y-
marchantia         GCAACCTCTCAGTCTGGCCG GTTCCACTCTGAGAGCCCTT CATGGCTGCCATGTTCAGTA
FUNARIA            -Q--P--V--S--L--A--G --S--T--L--R--S--P-- H--G--C--H--A--Q--Y-
```

funaria	GCAGCCAGTCAGTTTAGCAG	GGTCGACTTTACGATCCCCT	CATGGATGCCATGCTCAGTA
PHYSCOMITRELLA	-Q--P--V--S--L--A--G	--S--T--L--R--S--P--	H--G--C--H--A--Q--Y-
physcomitrella	GCAGCCGGTCAGCTTGGCAG	GTTCGACTTTACGTTCTCCT	CACGGATGCCATGCCCAGTA
CERATODON	-Q--P--V--S--L--A--G	--S--T--L--R--A--P--	H--G--C--H--A--Q--Y-
ceratodon	GCAGCCAGTCAGCTTGGCAG	GTTCGACTTTACGTGCCCCG	CATGGATGTCATGCCCAGTA
SELAGINELLA	-Q--P--I--S--L--A--G	--S--T--L--R--A--P--	H--G--C--H--A--Q--Y-
selaginella	GCAACCGATCAGCCTGGCTG	GCTCCACGCTGCGAGCACCG	CACGGATGCCATGCGCAGTA
EQUISETUM	-Q--P--L--S--L--A--G	--S--A--L--R--A--P--	H--G--C--H--A--Q--Y-
equisetum	GCAACCTCTAAGCTTAGCCG	GGTCTGCTTTGCGAGCTCCC	CACGGTTGCCATGCGCAGTA
MARSILEA	-Q--P--L--S--L--A--A	--S--T--L--R--A--P--	H--G--C--H--S--Q--Y-
marsilea	ACAGCCATTGAGTTTAGCTG	CTTCCACTCTACGTGCCCCA	CATGGTTGTCATAGCCAGTA
ADIANTUM	-Q--P--M--S--L--T--G	--S--T--L--R--A--P--	H--G--C--H--T--Q--Y-
adiantum	CCAGCCCATGAGCTTGACAG	GGTCAACGCTCCGCGCGCCG	CATGGATGCCACACCCAATA
PSILOTUM	-Q--P--L--S--L--A--G	--S--T--L--R--A--P--	H--G--C--H--A--H--Y-
psilotum	ACAGCCTCTCAGCTTGGCTG	GCTCTACACTGAGGGCCCCT	CATGGTTGTCATGCGCACTA
PSILOTUM2	-?--?--?--?--?--?--?	--?--?--?--?--?--?--	?--?--?--?--?--?--?-
psilotum2	????????????????????	????????????????????	????????????????????
METASEQUOIA	-Q--P--L--C--L--V--G	--S--T--L--R--A--P--	H--G--C--H--A--Q--Y-
metasequoia	GCAGCCCCTCTGTTTGGTGG	GTTCTACGCTGAGGGCACCC	CATGGTTGTCATGCTCAGTA
PSEUDOTSUGA	-R--D--L--S--F--C--G	--S--T--L--R--A--P--	H--G--C--H--A--Q--Y-
pseudotsuga	TAGAGACTTGAGTTTTTGTG	GGTCTACATTGAGAGCACCT	CATGGTTGCCATGCTCAATA
PICEA	-Q--P--L--C--L--V--G	--S--T--L--R--A--P--	H--G--C--H--A--Q--Y-
picea	GCAGCCTCTATGTTTAGTGG	GTTCAACGCTTCGGGCGCCC	CATGGGTGCCACGCCCAATA
EPHEDRA	-K--E--L--S--F--C--G	--S--T--L--R--A--P--	H--G--C--H--A--Q--Y-
ephedra	AAAAGAGCTGAGTTTCTGCG	GTTCTACTCTGAGGGCACCT	CATGGATGTCATGCTCAATA
GNETUM	-Q--P--L--C--L--V--G	--S--T--L--R--A--P--	H--G--C--H--A--Q--Y-
gnetum	GCAGCCTTTGTGTTTAGTGG	GTTCTACTCTGCGAGCTCCG	CATGGCTGCCATGCTCAATA
PHYC_PIPER	-Q--P--L--S--L--A--G	--S--T--L--R--A--P--	H--G--C--H--S--Q--Y-
phycc_piper	ACAGCCGCTTAGTCTGGCTG	GATCGACGTTGAGAGCACCC	CATGGCTGCCATTCACAGTA
PHYB_PIPER	-Q--P--F--S--F--L--P	--S--T--L--R--A--P--	L--G--C--H--A--Q--Y-
phyb_piper	GCAGCCATTTTCTTTCCTCC	CTTCGACTCTCCGAGCCCCC	CTCGGCTGCCATGCTCAGTA
PHYA_PIPER	-F--D--L--T--L--C--G	--S--T--L--R--A--P--	H--S--C--H--L--H--Y-
phya_piper	CTTTGATCTCACCCTGTGCG	GCTCCACCCTCCGGGCTCCC	CACAGTTGCCATCTCCACTA
PHYB_ARABIDOPSIS	-Q--S--M--C--L--V--G	--S--T--L--R--A--P--	H--G--C--H--S--Q--Y-
phyb_arabidopsis	TCAGTCTATGTGCTTGGTTG	GTTCTACTCTTAGGGCTCCT	CATGGTTGTCACTCTCAGTA
PHYB_GLYCINE	-Q--P--L--C--L--V--G	--S--T--L--G--A--P--	H--G--C--H--A--Q--Y-
phyb_glycine	GCAGCCTTTGTGTTTGGTTG	GGTCCACCTTAGGGGCTCCT	CACGGTTGTCATGCTCAGTA
PHYB_NICOTIANA	-Q--P--L--C--L--V--G	--S--T--L--R--A--P--	H--G--C--H--A--Q--Y-
phyb_nicotiana	GCAGCCTTTATGTTTAGTTG	GTTCCACACTTAGAGCCCCT	CATGGTTGCCACGCGCAGTA
PHYB_SOLANUM	-Q--P--L--C--L--V--G	--S--T--L--R--A--P--	H--G--C--H--A--Q--Y-
phyb_solanum	GCAGCCTTTATGTCTAGTTG	GTTCCACACTTAGAGCACCT	CATGGTTGCCACGCACAGTA
PHYD_ARABIDOPSIS	-Q--F--I--C--L--V--G	--S--T--L--R--A--P--	H--G--C--H--A--Q--Y-
phyd_arabidopsis	GCAGTTTATATGCTTGGTGG	GTTCGACTTTGCGAGCTCCT	CATGGCTGTCATGCTCAATA
PHYE_ARABIDOPSIS	-R--P--L--C--L--V--N	--S--T--L--R--A--P--	H--G--C--H--T--Q--Y-
phye_arabidopsis	GAGACCACTTTGTTTAGTTA	ATTCTACTCTAAGAGCTCCT	CATGGCTGCCATACGCAGTA
PHYB_ORYZA	-Q--P--L--C--L--V--G	--S--T--L--R--S--P--	H--G--C--H--G--Q--Y-
phyb_oryza	ACAGCCGCTGTGCTTGGTTG	GGTCCACGCTGCGTTCGCCG	CATGGTTGCCATGGCCAGTA
PHYA_AVENA	-F--D--I--S--L--C--G	--S--A--L--R--A--P--	H--S--C--H--L--Q--Y-
phya_avena	CTTTGATATTAGCCTATGTG	GTTCAGCACTCAGGGCACCA	CACAGTTGTCACCTTCAGTA
PHYA_ORYZA	-L--D--I--S--L--C--G	--S--T--L--R--A--P--	H--S--C--H--L--Q--Y-
phya_oryza	CTTGGATATTAGCTTATGTG	GTTCAACACTGAGGGCACCA	CACAGTTGTCATCTTCAGTA
PHYA_ZEA	-I--D--I--S--L--C--G	--S--T--L--R--A--P--	H--S--C--H--L--K--Y-
phya_zea	CATTGATATTAGCTTGTGTG	GTTCAACTCTTAGAGCACCA	CATAGCTGTCACCTTAAGTA
PHYA_ARABIDOPSIS	-F--D--L--T--L--C--G	--S--T--L--R--A--P--	H--S--C--H--L--Q--Y-
phya_arabidopsis	CTTTGACCTTACCTTGTGTG	GCTCCACCCTTAGAGCACCG	CACAGCTGCCATTTGCAGTA
PHYA_CUCURBITA	-F--D--L--T--L--C--G	--S--T--L--R--A--P--	H--S--C--H--L--Q--Y-
phya_cucurbita	GTTTGATCTAACTTTATGTG	GTTCAACTTTAAGAGCTCCA	CACAGTTGCCATTTACAGTA
PHYA_PISUM	-F--D--L--T--L--C--G	--S--T--L--R--A--P--	H--S--C--H--L--Q--Y-
phya_pisum	ATTTGATTTGACTCTGTGCG	GTTCGACCTTGAGAGCTCCA	CATAGTTGCCATTTGCAGTA
PHYA_GLYCINE	-F--D--L--I--L--C--G	--S--T--L--R--A--P--	H--S--C--H--A--Q--Y-
phya_glycine	ATTTGATTTGATTTTGTGTG	GTTCCACCTTAAGAGCTCCT	CATAGTTGCCACGCGCAGTA
PHYA_NICOTIANA	-F--D--L--T--L--C--G	--S--T--L--R--A--P--	H--Y--C--H--L--Q--Y-
phya_nicotiana	ATTTGATTTAACATTGTGCG	GCTCTACTCTTAGGGCTCCT	CACTACTGCCATCTACAGTA
PHYA_SOLANUM	-F--D--L--T--L--C--G	--S--T--L--R--A--P--	H--Y--C--H--L--Q--Y-
phya_solanum	ATTTGACTTAACATTGTGCG	GCTCTACTCTTAGGGCCCCT	CACTACTGCCATTTACAGTA
PHYC_ARABIDOPSIS	-Q--P--I--S--L--S--G	--S--T--L--R--A--P--	H--G--C--H--A--Q--Y-
phyc_arabidopsis	ACAGCCAATAAGTCTTTCTG	GATCTACTTTGAGAGCTCCT	CATGGTTGTCACGCACAGTA
MOUGEOTIA	-M--M--N--M--G--S--T	--A--S--L--V--M--C--	V--T--I--N--D--T--N-
mougeotia	CATGATGAACATGGGCTCCA	CGGCTTCTCTGGTGATGTGT	GTCACTATCAACGACACCAA
CHARA	-M--K--N--M--G--S--A	--A--S--L--V--M--A--	V--I--I--N--D--A--E-
chara	CATGAAGAACATGGGTTCGG	CGGCGTCTCTGGTGATGGCC	GTGATCATTAACGATGCGGA
MARCHANTIA	-M--G--N--M--G--Q--I	--A--S--L--V--M--A--	V--I--I--K--P--Q--R-

marchantia	CATGGGAAATATGGGGCAGA	TAGCGTCTCTGGTCATGGCA	GTGATAATCAAGCCTCAACG
FUNARIA	-M--G--N--M--G--S--I	--A--S--L--V--M--A--	V--I--I--N--D--N--E-
funaria	CATGGGCAACATGGGGTCCA	TTGCATCGTTAGTCATGGCT	GTGATTATCAATGATAACGA
PHYSCOMITRELLA	-M--G--N--M--G--S--I	--A--S--L--V--M--A--	V--I--I--N--D--N--E-
physcomitrella	CATGGGCAATATGGGGTCCA	TTGCGTCCCTAGTCATGGCT	GTGATCATCAACGATAACGA
CERATODON	-M--G--N--M--G--S--I	--A--S--L--V--M--A--	V--I--I--N--D--N--E-
ceratodon	CATGGGTAACATGGGCTCCA	TTGCGTCGCTTGTCATGGCC	GTAATCATCAATGATAACGA
SELAGINELLA	-M--G--N--M--G--S--V	--A--S--L--V--M--A--	M--I--I--N--D--N--D-
selaginella	CATGGGCAACATGGGCTCTG	TTGCGTCCCTGGTCATGGCC	ATGATCATCAACGACAACGA
EQUISETUM	-M--E--N--M--G--S--I	--A--S--L--V--M--A--	V--T--L--N--D--N--E-
equisetum	CATGGAGAACATGGGGTCGA	TTGCATCACTAGTCATGGCT	GTAACTCTTAATGACAATGA
MARSILEA	-M--A--N--M--G--S--I	--A--S--L--V--M--A--	V--I--V--N--D--N--E-
marsilea	TATGGCAAACATGGGCTCTA	TTGCTTCCTTGGTGATGGCT	GTTATTGTCAATGACAATGA
ADIANTUM	-M--A--N--M--N--S--I	--S--S--L--V--M--A--	V--I--V--N--D--S--D-
adiantum	CATGGCCAACATGAATTCAA	TCTCCTCCTTGGTGATGGCT	GTAATTGTGAATGATAGTGA
PSILOTUM	-M--G--N--M--G--S--I	--A--S--L--V--M--A--	V--I--V--K--R--H--G-
psilotum	TATGGGTAATATGGGTTCTA	TTGCCTCTTTGGTGATGGCT	GTTATAGTAAAGCGTCATGG
PSILOTUM2	-?--?--?--?--G--S--I	--G--S--L--V--M--A--	V--I--V--N--D--N--D-
psilotum2	????????????GGATCCA	TTGGCTCTTTGGTAATGGCT	GTTATAGTGAACGACAACGA
METASEQUOIA	-M--A--N--M--G--S--I	--A--S--L--V--M--A--	V--I--V--N--G--N--G-
metasequoia	CATGGCTAATATGGGCTCCA	TTGCCTCTCTTGTTATGGCT	GTAATTGTCAATGGAAATGG
PSEUDOTSUGA	-M--A--N--M--G--S--I	--A--S--L--V--M--A--	V--V--V--T--E--K--E-
pseudotsuga	CATGGCCAACATGGGCTCTA	TAGCATCTTTAGTAATGGCT	GTAGTAGTCACTGAGAAGGA
PICEA	-M--A--N--M--G--S--I	--A--S--L--V--M--A--	V--I--I--N--G--N--D-
picea	CATGGCCAACATGGGTTCCA	TTGCTTCTCTTGTTATGGCC	GTGATTATTAACGGGAATGA
EPHEDRA	-M--A--N--M--G--S--V	--A--S--L--V--M--A--	V--I--V--N--E--K--E-
ephedra	CATGGCTAATATGGGTTCCG	TTGCCTCACTTGTTATGGCC	GTCATCGTGAATGAGAAGGA
GNETUM	-M--A--N--M--S--S--V	--A--S--L--V--M--A--	V--I--V--N--G--N--E-
gnetum	CATGGCCAACATGAGCTCCG	TTGCATCATTGGTCATGGCT	GTGATAGTGAATGGAAATGA
PHYC_PIPER	-M--S--N--M--G--S--I	--A--S--L--V--M--S--	V--M--V--C--D--G--D-
phycc_piper	CATGTCCAATATGGGTTCCA	TTGCATCTCTTGTGATGTCT	GTGATGGTGTGTGACGGCGA
PHYB_PIPER	-M--A--N--M--G--S--V	--A--S--L--A--M--A--	V--I--I--N--G--G--D-
phyb_piper	CATGGCCAACATGGGCTCCG	TCGCCTCCCTGGCCATGGCC	GTCATCATCAACGGGGGCGA
PHYA_PIPER	-M--A--N--M--G--S--I	--A--S--L--V--M--A--	V--V--I--N--E--A--A-
phya_piper	CATGGCCAACATGGGCTCTA	TCGCGTCCCTTGTCATGGCG	GTGGTCATCAATGAGGCGGC
PHYB_ARABIDOPSIS	-M--A--N--M--G--S--I	--A--S--L--A--M--A--	V--I--I--N--G--N--E-
phyb_arabidopsis	TATGGCTAACATGGGATCTA	TTGCGTCTTTAGCAATGGCG	GTTATAATCAATGGAAATGA
PHYB_GLYCINE	-M--A--N--M--G--S--I	--A--S--L--V--M--A--	V--I--I--N--G--N--D-
phyb_glycine	TATGGCTAACATGGGCTCGA	TTGCGTCTTTGGTGATGGCA	GTTATTATCAATGGGAATGA
PHYB_NICOTIANA	-M--A--N--M--G--S--I	--A--S--L--T--L--A--	V--I--I--N--G--N--D-
phyb_nicotiana	CATGGCAAATATGGGGTCTA	TTGCGTCATTAACACTAGCA	GTTATTATCAATGGAAACGA
PHYB_SOLANUM	-M--A--N--M--G--S--I	--A--S--L--T--L--A--	V--I--I--N--G--N--D-
phyb_solanum	CATGGCAAATATGGGGTCTA	TTGCCTCATTAACACTGGCA	GTTATTATCAACGGAAATGA
PHYD_ARABIDOPSIS	-M--T--N--M--G--S--I	--A--S--L--A--M--A--	V--I--I--N--G--N--E-
phyd_arabidopsis	CATGACTAACATGGGCTCTA	TTGCGTCGTTAGCTATGGCA	GTTATAATAAATGGAAACGA
PHYE_ARABIDOPSIS	-M--A--N--M--G--S--V	--A--S--L--A--L--A--	I--V--V--K----------
phye_arabidopsis	TATGGCGAATATGGGCTCTG	TAGCTTCTCTTGCACTCGCA	ATTGTAGTAAAA--------
PHYB_ORYZA	-M--A--N--M--G--S--I	--A--S--L--V--M--A--	V--I--I--S--S--G--G-
phyb_oryza	TATGGCGAACATGGGTTCCA	TTGCATCTCTTGTTATGGCA	GTGATCATTAGTAGTGGTGG
PHYA_AVENA	-M--E--N--M--N--S--I	--A--S--L--V--M--A--	V--V--V--N--E--N--E-
phya_avena	TATGGAGAACATGAACTCGA	TTGCATCCCTTGTCATGGCT	GTTGTGGTTAATGAGAATGA
PHYA_ORYZA	-M--E--N--M--N--S--I	--A--S--L--V--M--A--	V--V--V--N--E--N--E-
phya_oryza	TATGGAGAACATGAACTCGA	TTGCATCCCTTGTCATGGCT	GTTGTGGTTAATGAGAATGA
PHYA_ZEA	-M--E--N--M--N--S--I	--A--S--L--V--M--A--	V--V--V--N--E--N--E-
phya_zea	TATGGAGAACATGAACTCGA	TTGCATCCCTTGTCATGGCT	GTTGTGGTCAATGAAAATGA
PHYA_ARABIDOPSIS	-M--A--N--M--D--S--I	--A--S--L--V--M--A--	V--V--V--N--E--E--D-
phya_arabidopsis	CATGGCCAACATGGATTCAA	TTGCATCTCTGGTTATGGCG	GTTGTAGTTAACGAGGAAGA
PHYA_CUCURBITA	-M--E--N--M--N--S--I	--A--S--L--V--M--A--	V--V--V--N--E--G--D-
phya_cucurbita	TATGGAAAACATGAACTCTA	TAGCCTCTTTGGTTATGGCA	GTTGTGGTTAATGAAGGGGA
PHYA_PISUM	-M--A--N--M--D--S--I	--A--S--L--V--M--A--	V--V--V--N--D--S--D-
phya_pisum	CATGGCTAACATGGATTCAA	TTGCTTCGTTGGTTATGGCA	GTTGTCGTCAATGACAGCGA
PHYA_GLYCINE	-M--A--N--M--D--S--I	--A--S--L--V--L--A--	V--V--V--N--D--N--E-
phya_glycine	CATGGCTAACATGGATTCAA	TTGCTTCCCTGGTTTTGGCA	GTTGTAGTCAATGACAACGA
PHYA_NICOTIANA	-M--E--N--M--S--S--I	--A--S--L--V--M--A--	V--V--V--N--D--G--D-
phya_nicotiana	TATGGAGAACATGAGTTCAA	TTGCATCCCTTGTAATGGCA	GTTGTGGTCAATGACGGGGA
PHYA_SOLANUM	-M--E--N--M--N--S--I	--A--S--L--V--M--A--	V--V--V--N--D--G--D-
phya_solanum	TATGGAGAACATGAATTCAA	TTGCATCACTTGTAATGGCA	GTTGTGGTCAATGACGGGGA
PHYC_ARABIDOPSIS	-M--S--N--M--G--S--V	--A--S--L--V--M--S--	V--T--I--N--G--S--D-
phyc_arabidopsis	TATGAGTAATATGGGATCAG	TGGCGTCTCTTGTCATGTCT	GTAACTATCAATGGTAGTGA
MOUGEOTIA	-E--I--A--G--G--P---	--------------------	--------------------
mougeotia	TGAGATTGCTGGAGGACCA-	--------------------	--------------------
CHARA	-D--E--G--G--G--S--G	--G--G--G--A--S--A--	S--I----------------

```
chara             GGATGAAGGAGGTGGTAGTG GAGGAGGCGGTGCTTCTGCT TCCATA--------------
MARCHANTIA        -E--E--Y--Q--T--T--G --V--N-------------- --------------------
marchantia        TGAGGAGTACCAGACGACCG GGGTTAAC------------ --------------------
FUNARIA           -E--D--S--R--G--S--- -------------------- --------------------
funaria           GGAAGATTCTCGTGGATCA- -------------------- --------------------
PHYSCOMITRELLA    -E--D--------------S --------H--G-----S-- V-------------------
physcomitrella    GGAAGAT------------T CG------CATGGC---TCA GTC-----------------
CERATODON         -E--Y--------------S --------R--G-----A-- I-------------------
ceratodon         GGAATAT------------T CA------CGTGGG---GCA ATT-----------------
SELAGINELLA       -E--P--S--G--------G --------G--G--G--G-- G-------------------
selaginella       TGAACCGTCTGGC------G GC------GGCGGCGGCCGC GGG-----------------
EQUISETUM         -F--D--N--A--N--E--- -------------------- --------------------
equisetum         ATTTGACAATGCTAATGAA- -------------------- --------------------
MARSILEA          -E--E--T--S--N--R--N --Q----------------- --------------------
marsilea          AGAAGAAACAAGTAACCGGA ATCAA--------------- --------------------
ADIANTUM          -D--D--------------S --------A--G--H--S-- S-------------------
adiantum          TGACGAT------------A GT------GCCGGCCATTCT TCG-----------------
PSILOTUM          -E--E--D--R--------S --------L--G--F--Q-- S-------------------
psilotum          CGAGGAGGATAGG------A GC------CTGGGCTTTCAG TCA-----------------
PSILOTUM2         -A--E--P--S--------G --------R--G--N--Q-- P-------------------
psilotum2         TGCAGAGCCGAGC------G GC------CGCGGAAATCAG CCG-----------------
METASEQUOIA       -N--D--E--E--------- --------G--G--N--G-- E-------------------
metasequoia       AAATGATGAAGAA------- --------GGCGGCAACGGC GAA-----------------
PSEUDOTSUGA       -E--D--------A--D--S --E--------G-----Q-- I-------------------
pseudotsuga       AGAGGAC------GCAGATA GCGAG------GGG---CAG ATA-----------------
PICEA             -D--E--G--A--G--D--- -------------------- --------------------
picea             TGATGAAGGTGCGGGGGAT- -------------------- --------------------
EPHEDRA           -D--E--E--A--E--N--E --G--Q-------------- S-------------------
ephedra           AGATGAGGAAGCAGAAAATG AAGGACAG------------ AGC-----------------
GNETUM            -E--D--A--G--H--P--M --T--I--A--A--I--I-- A--G--T--G--N--R--N-
gnetum            AGAGGATGCAGGGCATCCAA TGACAATAGCAGCAATAATA GCAGGAACAGGTAACAGAAA
PHYC_PIPER        -E--D--S--N--G--E--K --T--A-------------- --------------------
phycc_piper       CGAAGATTCGAATGGGGAAA AAACGGCT------------ --------------------
PHYB_PIPER        -E--D--A--A--G--G--R --S----------------- --------------------
phyb_piper        CGAGGACGCCGCCGGGGGGC GGAGC--------------- --------------------
PHYA_PIPER        -T--T--S--T--T--T--T --D--L--H--P--A--V-- G--D--N--S--F--S--S-
phya_piper        TACGACGAGTACTACTACTA CTGATCTTCATCCTGCTGTT GGTGACAATTCTTTTTCTTC
PHYB_ARABIDOPSIS  -D--D--G--S--------N --------V--A-------- S-------------------
phyb_arabidopsis  AGATGATGGGAGC------A AT------GTAGCT------ AGT-----------------
PHYB_GLYCINE      -E--E--G--V--G------ -------------------- --------------------
phyb_glycine      CGAGGAAGGCGTTGGT---- -------------------- --------------------
PHYB_NICOTIANA    -E--E--------------A --------V----------- G-------------------
phyb_nicotiana    TGAGGAA------------G CT------GTT--------- GGG-----------------
PHYB_SOLANUM      -E--E--------------A --------V--G-------- G-------------------
phyb_solanum      TGAGGAA------------G CT------GTGGGT------ GGC-----------------
PHYD_ARABIDOPSIS  -E--D--G--N--------G --------V-----N--T-- G-------------------
phyd_arabidopsis  AGAAGATGGTAAT------G GG------GTT---AATACT GGA-----------------
PHYE_ARABIDOPSIS  -------------------- -------------------- --------------------
phye_arabidopsis  -------------------- -------------------- --------------------
PHYB_ORYZA        -D--D--D--H--------N --I--A--R--G-------- S-------------------
phyb_oryza        GGATGATGATCAT------A ACATTGCACGGGGC------ AGC-----------------
PHYA_AVENA        -E--D--D--E--A--E--S --E--Q--P--A-------- Q-------------------
phya_avena        AGAGGATGATGAAGCTGAGT CTGAACAACCAGCA------ CAG-----------------
PHYA_ORYZA        -D--D--D--E--V--G--A --D--Q--P--A-------- Q-------------------
phya_oryza        GGATGATGATGAAGTTGGGG CTGATCAACCTGCA------ CAA-----------------
PHYA_ZEA          -E--D--D--E--P--E--P --E--Q--P--P-------- Q-------------------
phya_zea          AGAGGATGATGAACCCGAGC CTGAACAACCACCA------ CAA-----------------
PHYA_ARABIDOPSIS  -G--E--G--D--------A --P--D--A--T--T----- Q-------------------
phya_arabidopsis  TGGAGAAGGGGAT------G CTCCTGATGCTACTACA--- CAG-----------------
PHYA_CUCURBITA    -E--E--N--E--------G --------P--A-----L-- Q-------------------
phya_cucurbita    TGAAGAAAATGAA------G GC------CCTGCT---TTG CAG-----------------
PHYA_PISUM        -E--D--G--D--------S --A--D--A--------V-- L-------------------
phya_pisum        TGAAGATGGAGAT------A GCGCTGACGCA------GTT CTC-----------------
PHYA_GLYCINE      -E--D--G--D--------T -----D--A--------V-- Q-------------------
phya_glycine      AGAAGATGGGGAC------A CT---GATGCT------GTT CAG-----------------
PHYA_NICOTIANA    -E--E--G--E--------S --S--D--S--T-------- Q-------------------
phya_nicotiana    CGAAGAGGGAGAA------A GCTCTGATTCGACA------ CAA-----------------
PHYA_SOLANUM      -E--E--G--E--------S --S--D--S--S-------- Q-------------------
phya_solanum      TGAAGAAGGAGAA------A GCTCTGATTCTTCA------ CAA-----------------
PHYC_ARABIDOPSIS  -S--D--E--M--------N --R--D--L----------- --------------------
phyc_arabidopsis  TAGTGATGAGATG------A ACAGAGATTTA--------- --------------------

MOUGEOTIA         -------------------- -------------------- ---------G--M--K--G-
```

```
mougeotia           -------------------- -------------------- ---------GGGATGAAAGG
CHARA               -------------------- -------------------- ---------T--P--R--G-
chara               -------------------- -------------------- ---------ACTCCTCGTGG
MARCHANTIA          -------------------- -------------------- ---------H--H--K--G-
marchantia          -------------------- -------------------- ---------CACCACAAGGG
FUNARIA             -------------------- -------------------- ---------V--Q--R--G-
funaria             -------------------- -------------------- ---------GTTCAAAGAGG
PHYSCOMITRELLA      -------------------- -------------------- ---------Q--R-----G-
physcomitrella      -------------------- -------------------- ---------CAAAGA---GG
CERATODON           -------------------- -------------------- ---------Q--R-----G-
ceratodon           -------------------- -------------------- ---------CAAAGA---GG
SELAGINELLA         -------------------- -------------------- ---------Q--H--K--G-
selaginella         -------------------- -------------------- ---------CAGCACAAGGG
EQUISETUM           -------------------- -------------------- ---------S--L--K--R-
equisetum           -------------------- -------------------- ---------TCATTGAAAAG
MARSILEA            -------------------- -------------------- ---------Q--P--K--M-
marsilea            -------------------- -------------------- ---------CAGCCAAAAAT
ADIANTUM            -------------------- -------------------- ---------Q--------G-
adiantum            -------------------- -------------------- ---------CAA------GG
PSILOTUM            -------------------- -------------------- ---------Q--N-----G-
psilotum            -------------------- -------------------- ---------CAAAAC---GG
PSILOTUM2           -------------------- -------------------- ---------K--N-------
psilotum2           -------------------- -------------------- ---------AAGAAC-----
METASEQUOIA         -------------------- -------------------- ---------V--Q--W----
metasequoia         -------------------- -------------------- ---------GTTCAGTGG--
PSEUDOTSUGA         -------------------- -------------------- ---------Q--Q--K--G-
pseudotsuga         -------------------- -------------------- ---------CAGCAGAAAGG
PICEA               -------------------- -------------------- ------------R--N--S-
picea               -------------------- -------------------- ------------AGAAATTC
EPHEDRA             -------------------- -------------------- ---------Q--Q--K--G-
ephedra             -------------------- -------------------- ---------CAACAAAAGGG
GNETUM              -N--S--------------- -------------------- ---------S--S--S--M-
gnetum              CAACAGC------------- -------------------- ---------AGCAGCAGCAT
PHYC_PIPER          -------------------- -------------------- --------------------
phycc_piper         -------------------- -------------------- --------------------
PHYB_PIPER          -------------------- -------------------- ---------G--S--N--R-
phyb_piper          -------------------- -------------------- ---------GGTTCGAATCG
PHYA_PIPER          -S--S--A--T--T--A--A --A--A--S--D--T--T-- T--T--A--P--S--R--R-
phya_piper          TTCATCTGCTACCACTGCTG CTGCTGCTAGTGATACCACC ACCACCGCCCCTAGTCGGAG
PHYB_ARABIDOPSIS    -------------------- -------------------- ---------G--R--S--S-
phyb_arabidopsis    -------------------- -------------------- ---------GGAAGAAGCTC
PHYB_GLYCINE        -------------------- -------------------- ---------G--R--S--S-
phyb_glycine        -------------------- -------------------- ---------GGTCGCAGTTC
PHYB_NICOTIANA      -------------------- -------------------- ---------G--R--S--S-
phyb_nicotiana      -------------------- -------------------- ---------GGCCGAAGTTC
PHYB_SOLANUM        -------------------- -------------------- ---------G--R--N--S-
phyb_solanum        -------------------- -------------------- ---------GGTCGAAATTC
PHYD_ARABIDOPSIS    -------------------- -------------------- ---------G--R--N--S-
phyd_arabidopsis    -------------------- -------------------- ---------GGAAGAAACTC
PHYE_ARABIDOPSIS    -------------------- -------------------- ---------G--K--D--S-
phye_arabidopsis    -------------------- -------------------- ---------GGCAAAGATTC
PHYB_ORYZA          -------------------- -------------------- ---------I--P--S--A-
phyb_oryza          -------------------- -------------------- ---------ATCCCGTCGGC
PHYA_AVENA          -------------------- -------------------- ---------Q--Q--K--K-
phya_avena          -------------------- -------------------- ---------CAGCAGAAAAA
PHYA_ORYZA          -------------------- -------------------- ---------Q--Q--K--R-
phya_oryza          -------------------- -------------------- ---------CAGCAGAAGAG
PHYA_ZEA            -------------------- -------------------- ---------Q--Q--K--K-
phya_zea            -------------------- -------------------- ---------CAGCAGAAGAA
PHYA_ARABIDOPSIS    -------------------- -------------------- ---------P--Q--K--R-
phya_arabidopsis    -------------------- -------------------- ---------CCTCAAAAGAG
PHYA_CUCURBITA      -------------------- -------------------- ---------Q--Q--K--R-
phya_cucurbita      -------------------- -------------------- ---------CAGCAAAAGAG
PHYA_PISUM          -------------------- -------------------- ---------P--Q--K--K-
phya_pisum          -------------------- -------------------- ---------CCACAAAAGAA
PHYA_GLYCINE        -------------------- -------------------- ---------P--Q--K--T-
phya_glycine        -------------------- -------------------- ---------CCACAAAAGAC
PHYA_NICOTIANA      -------------------- -------------------- ---------S--Q--K--R-
phya_nicotiana      -------------------- -------------------- ---------TCTCAAAAGAG
PHYA_SOLANUM        -------------------- -------------------- ---------S--Q--K--R-
phya_solanum        -------------------- -------------------- ---------TCTCAAAAAAG
PHYC_ARABIDOPSIS    -------------------- -------------------- ------------Q--T--G-
phyc_arabidopsis    -------------------- -------------------- ------------CAGACTGG
```

```
MOUGEOTIA          -R--K--L--W--G--L--I --V--C--H--H--S--T-- P--R--H--I--P--F--P-
mougeotia          GCGAAAGTTGTGGGGATTAA TCGTCTGCCATCACTCGACT CCTCGCCACATTCCGTTCCC
CHARA              -R--K--L--W--G--L--V --V--C--H--H--A--T-- P--R--V--V--P--F--P-
chara              TAGAAAGCTATGGGGTCTTG TTGTGTGCCATCATGCAACC CCTAGGGTGGTTCCATTCCC
MARCHANTIA         -R--K--L--W--G--L--V --V--C--H--H--T--T-- P--R--S--V--P--F--P-
marchantia         CAGGAAGCTATGGGGCCTTG TAGTCTGTCACCACACAACT CCTCGATCAGTTCCTTTTCC
FUNARIA            -R--K--L--W--G--F--V --V--C--H--H--T--S-- P--R--T--V--P--F--P-
funaria            TAGAAAGCTATGGGGGTTTG TAGTGTGCCATCACACATCT CCACGGACTGTGCCATTTCC
PHYSCOMITRELLA     -R--K--L--W--G--L--V --V--C--H--H--T--S-- P--R--T--V--P--F--P-
physcomitrella     TAGAAAGCTATGGGGACTTG TAGTGTGTCATCATACATCT CCACGAACTGTACCGTTTCC
CERATODON          -R--K--L--W--G--L--V --V--C--Q--H--T--S-- P--R--T--V--P--F--P-
ceratodon          TAGAAAGCTGTGGGGACTCG TAGTGTGTCAGCATACATCT CCACGAACTGTACCGTTTCC
SELAGINELLA        -R--R--L--W--G--L--V --V--C--H--H--T--S-- P--R--S--V--P--F----
selaginella        AAGAAGGCTGTGGGCGCTGG TGGTGTGTCACCACACGTCC CCGCGATCGGTGCCCTTC--
EQUISETUM          -R--K--L--W--G--L--I --V--C--H--H--T--S-- P--R--A--V--P--F----
equisetum          AAGGAAATTGTGGGGACTCA TCGTGTGCCATCATACGAGC CCAAGAGCAGTTCCATTT--
MARSILEA           -R--K--L--W--G--L--V --V--C--H--H--T--T-- P--R--A--V--P--F--Q-
marsilea           GAGGAAGCTTTGGGGACTGG TTGTTTGTCATCATACAACC CCAAGGGCAGTGCCATTTCA
ADIANTUM           -I--K--L--W--G--L--V --V--C--H--H--T--S-- P--R--Y--V--P--F--P-
adiantum           CATTAAACTTTGGGGCTTGG TCGTATGCCACCATACGTCA CCACGCTATGTACCTTTCCC
PSILOTUM           ----R--L--W--G--M--V --V--C--H--H--T--T-- P--R--A--V--P--F--A-
psilotum           A---AGGCTATGGGGGATGG TGGTTTGTCACCACACGACT CCCAGAGCTGTCCCTTTCGC
PSILOTUM2          -R--R--L--W--G--M--V --V--C--H--H--T--T-- P--R--A--V--P--F--S-
psilotum2          -AGGAGGCTCTGGGGAATGG TTGTTTGCCACCACACCACT CCCAGGGCGGTCCCGTTCTC
METASEQUOIA        ----K--L--W--G--L--V --V--C--H--H--T--S-- P--R--A--V--P--F--P-
metasequoia        ----AAGCTATGGGGGTTGG TTGTTTGTCATCACACCTCA CCCCGAGCCGTTCCGTTTCC
PSEUDOTSUGA        -R--R--L--W--G--L--V --V--C--H--H--T--S-- P--R--Y--V--P--F--P-
pseudotsuga        AAGAAGGCTATGGGGACTGG TGGTTTGTCATCACACTAGT CCACGCTATGTTCCATTTCC
PICEA              -M--K--L--W--G--L--V --V--C--H--H--T--S-- P--R--P--V--P--F--P-
picea              CATGAAGCTCTGGGGTTTGG TTGTGTGCCACCATACCTCC CCACGGCCGGTTCCGTTTCC
EPHEDRA            -R--K--L--W--G--L--V --V--C--H--H--T--G-- P--R--F--V--P--F--P-
ephedra            GAGGAAGCTTTGGGGACTCG TCGTTTGTCATCATACAGGT CCTAGATTTGTGCCATTTCC
GNETUM             -M--K--L--W--G--L--I --V--C--H--H--T--S-- P--R--A--V--P--F--P-
gnetum             GATGAAGCTTTGGGGTCTGA TTGTGTGCCACCATACTTCT CCCAGAGCTGTTCCATTCCC
PHYC_PIPER         -R--K--L--W--G--L--V --V--C--H--H--T--S-- P--R--F--I--P--F--P-
phycc_piper        -AGGAAGCTTTGGGGATTGG TGGTATGCCACCATACCAGT CCGAGGTTCATACCTTTTCC
PHYB_PIPER         -S--K--L--W--G--L--V --V--C--H--H--T--S-- P--R--C--I--P--F--P-
phyb_piper         TTCGAAGCTGTGGGGGCTGG TGGTGTGCCACCACACGTCC CCTCGCTGTATCCCCTTTCC
PHYA_PIPER         -K--M--L--W--G--L--V --V--C--H--N--T--K-- P--R--F--V--P--F--P-
phya_piper         GAAGATGCTGTGGGGCCTTG TAGTTTGCCACAACACCAAG CCGAGGTTCGTTCCCTTCCC
PHYB_ARABIDOPSIS   -M--R--L--W--G--L--V --V--C--H--H--T--S-- S--R--C--I--P--F--P-
phyb_arabidopsis   GATGAGGCTTTGGGGTTTGG TTGTTTGCCATCACACTTCT TCTCGCTGCATACCGTTTCC
PHYB_GLYCINE       -M--R--L--W--G--L--V --V--C--H--H--T--S-- A--R--C--I--P--F--P-
phyb_glycine       GATGAGGCTGTGGGGGCTTG TTGTCTGCCACCATACCTCT GCCAGGTGTATTCCTTTTCC
PHYB_NICOTIANA     -M--R--L--W--G--L--V --V--G--H--H--T--S-- A--R--C--I--P--F--P-
phyb_nicotiana     AATGAGGCTGTGGGGCTTGG TTGTTGGACACCATACTTCT GCTAGGTGCATTCCATTCCC
PHYB_SOLANUM       -M--R--L--W--G--L--V --V--G--H--H--T--S-- V--R--S--I--P--F--P-
phyb_solanum       AATGAGGCTATGGGGCTTGG TTGTTGGACACCACACTTCT GTTCGGTCCATTCCTTTCCC
PHYD_ARABIDOPSIS   -M--R--L--W--G--L--V --V--C--H--H--T--S-- A--R--C--I--P--F--P-
phyd_arabidopsis   GATGAGGCTTTGGGGTTTAG TTGTTTGCCATCACACATCA GCTCGTTGCATACCTTTTCC
PHYE_ARABIDOPSIS   -S--K--L--W--G--L--V --V--G--H--H--C--S-- P--R--Y--V--P--F--P-
phye_arabidopsis   GAGCAAGCTTTGGGGATTAG TTGTTGGTCATCATTGTTCT CCTAGATACGTTCCATTCCC
PHYB_ORYZA         -M--K--L--W--G--L--V --V--C--H--H--T--S-- P--R--C--I--P--F--P-
phyb_oryza         GATGAAGTTGTGGGGGTTGG TAGTATGCCACCACACATCT CCACGGTGCATCCCTTTCCC
PHYA_AVENA         -K--K--L--W--G--L--L --V--C--H--H--E--S-- P--R--Y--V--P--F--P-
phya_avena         GAAGAAACTATGGGGCCTCC TTGTTTGCCACCATGAGAGC CCCAGATATGTCCCTTTTCC
PHYA_ORYZA         -K--K--L--W--G--L--L --V--C--H--H--E--S-- P--R--Y--V--P--F--P-
phya_oryza         GAAGAAACTATGGGGACTCC TTGTTTGCCACCATGAGAGC CCCAGATATGTTCCTTTCCC
PHYA_ZEA           -K--R--L--W--G--L--I --V--C--H--H--E--S-- P--R--Y--V--P--F--P-
phya_zea           GAAGAGGCTGTGGGGTCTCA TTGTTTGCCACCATGAGAGC CCCAGATATGTCCCTTTCCC
PHYA_ARABIDOPSIS   -K--R--L--W--G--L--V --V--C--H--N--T--T-- P--R--F--V--P--F--P-
phya_arabidopsis   AAAGAGACTATGGGGTTTAG TGGTTTGTCACAATACGACT CCGAGGTTTGTTCCATTTCC
PHYA_CUCURBITA     -K--R--L--W--G--L--V --V--C--H--N--S--S-- P--R--F--V--P--F--P-
phya_cucurbita     AAAGAGATTATGGGGCTTGG TAGTATGTCATAATTCAAGT CCCAGATTTGTTCCGTTTCC
PHYA_PISUM         -K--R--L--W--G--L--V --V--C--H--N--T--T-- P--R--F--V--P--F--P-
phya_pisum         AAAGAGACTTTGGGGTTTGG TAGTTTGTCATAACACTACT CCAAGGTTTGTTCCTTTTCC
PHYA_GLYCINE       -E--R--L--W--G--L--V --V--C--H--N--T--T-- P--R--F--V--P--F--P-
phya_glycine       GGAGAGACTTTGGGGTTTGG TAGTTTGCCATAACACTACT CCCAGGTTTGTTCCCTTTCC
PHYA_NICOTIANA     -K--R--L--W--G--L--V --V--C--H--N--T--T-- P--R--F--V--P--F--P-
phya_nicotiana     AAAAAGGCTTTGGGGCCTGG TGGTTTGCCACAACACCACC CCGAGGTTTGTTCCCTTCCC
PHYA_SOLANUM       -K--R--L--W--G--L--V --V--S--H--N--T--T-- P--R--F--A--P--F--P-
phya_solanum       AAAAAGGCTATGGGGCTTGG TTGTTTCCCACAACACGACC CCAAGGTTCGCGCCCTTCCC
```

```
PHYC_ARABIDOPSIS -R--H--L--W--G--L--V --V--C--H--H--A--S-- P--R--F--V--P--F--P-
phyc_arabidopsis CAGACACTTATGGGGCTTGG TGGTTTGTCATCACGCAAGT CCTAGATTTGTTCCGTTTCC

MOUGEOTIA        -I--H--S--A--C--E--F --L--M--Q--V--F--G-- L--Q--L--N--M--E--A-
mougeotia        CATACATTCTGCGTGTGAAT TCCTGATGCAAGTTTTTGGC CTCCAGTTGAATATGGAGGC
CHARA            -L--R--A--A--C--E--F --L--M--Q--V--F--G-- L--Q--L--N--M--E--V-
chara            TCTCCGTGCAGCATGCGAGT TTTTGATGCAAGTTTTTGGT CTTCAGTTGAACATGGAAGT
MARCHANTIA       -L--R--S--A--C--E--F --R--M--Q--V--F--G-- L--Q--L--N--M--E--V-
marchantia       ATTGCGTTCAGCCTGTGAAT TCCGCATGCAGGTATTCGGC CTGCAGCTGAACATGGAGGT
FUNARIA          -L--R--S--A--C--E--F --L--M--Q--V--F--G-- L--Q--L--N--M--E--V-
funaria          ACTTCGGTCTGCTTGTGAGT TTTTGATGCAGGTGTTTGGT TTACAGCTGAATATGGAGGT
PHYSCOMITRELLA   -L--R--S--A--C--G--F --L--M--Q--V--F--G-- L--Q--L--N--M--E--V-
physcomitrella   CCTTCGGTCCGCGTGTGGGT TTTTGATGCAGGTGTTCGGA TTACAGCTTAACATGGAGGT
CERATODON        -L--R--S--V--C--E--F --L--M--Q--V--F--G-- M--Q--L--N--L--H--V-
ceratodon        ACTTCGGTCTGTGTGCGAGT TTTTGATGCAGGTATTTGGT ATGCAGCTCAACCTCCATGT
SELAGINELLA      -L--R--S--A--C--E--F --L--M--Q--V--F--G-- L--Q--L--N--M--E--A-
selaginella      -CTGCGCTCGGCGTGCGACT TCCTCATGCAGGTGTTTGGC CTGCAGCTCAACATGGAGGC
EQUISETUM        -L--R--S--A--C--E--F --L--M--Q--V--F--G-- L--Q--L--N--M--E--V-
equisetum        -CTACGCTCTGCATGTGAGT TCTTAATGCAAGTCTTCGGA TTGCAGCTCAACATGGAAGT
MARSILEA         -L--R--S--A--C--E--F --L--M--Q--V--F--G-- L--Q--L--N--M--E--L-
marsilea         ACTGAGGTCTGCTTGTGAGT TTCTAATGCAGGTGTTTGGC CTTCAGCTGAACATGGAACT
ADIANTUM         -V--R--S--A--C--E--F --L--M--Q--V--F--S-- L--Q--L--N--M--E--V-
adiantum         CGTAAGGTCGGCTTGTGAGT TCTTGATGCAGGTCTTTAGT CTTCAACTAAATATGGAGGT
PSILOTUM         -L--R--C--A--C--E--F --F--A--Q--V--F--A-- L--Q--L--N--M--E--L-
psilotum         TCTCAGGTGTGCTTGCGAGT TTTTCGCACAGGTCTTTGCC CTTCAACTAAACATGGAGCT
PSILOTUM2        -L--R--S--A--C--E--F --L--M--Q--V--F--G-- L--Q--L--N--M--E--I-
psilotum2        TCTGAGGTCAGCTTGTGAGT TTTTGATGCAGGTCTTTGGC CTTCAACTAAACATGGAAAT
METASEQUOIA      -L--R--S--A--C--E--F --L--M--Q--T--F--G-- L--Q--I--N--M--E--L-
metasequoia      TCTTCGTTCTGCTTGTGAAT TTCTGATGCAGACATTTGGT CTTCAGATTAACATGGAATT
PSEUDOTSUGA      -L--R--Y--A--C--E--F --L--M--Q--V--F--G-- I--H--L--N--K--E--V-
pseudotsuga      TTTGAGGTATGCTTGTGAGT TCTTGATGCAAGTTTTTGGG ATTCATCTGAACAAGGAAGT
PICEA            -L--R--Y--A--C--E--F --M--M--Q--A--F--G-- L--Q--L--N--M--E--L-
picea            TCTCCGCTATGCTTGCGAAT TTATGATGCAAGCGTTCGGT CTTCAGCTGAACATGGAATT
EPHEDRA          -L--R--Y--A--C--E--F --M--M--Q--V--F--S-- I--Q--L--N--K--E--V-
ephedra          ATTGAGATATGCTTGTGAAT TCATGATGCAGGTTTTCAGC ATTCAGCTCAACAAGGAAGT
GNETUM           -L--R--Y--A--C--E--F --L--M--Q--A--F--G-- L--Q--L--N--M--E--L-
gnetum           TCTGCGTTATGCCTGCGAGT TTCTTATGCAGGCATTTGGG CTTCAGCTAAACATGGAGTT
PHYC_PIPER       -L--R--Y--A--C--E--F --L--I--Q--V--F--A-- I--Q--L--S--K--E--V-
phycc_piper      ATTGCGGTATGCTTGCGAGT TTTTGATCCAGGTGTTTGCA ATCCAGCTCAGTAAAGAAGT
PHYB_PIPER       -S--A--Y--A--C--E--F --L--N--A--A--F--G-- L--Q--L--N--M--E--L-
phyb_piper       CTCTGCCTACGCTTGCGAGT TTCTGAATGCAGCCTTCGGC CTTCAGTTGAACATGGAGCT
PHYA_PIPER       -L--R--Y--A--C--E--F --L--M--Q--V--F--A-- I--H--V--N--K--E--L-
phya_piper       CCTCAGATACGCCTGCGAGT TCCTGATGCAGGTGTTTGCC ATCCACGTGAACAAGGAACT
PHYB_ARABIDOPSIS -L--R--Y--A--C--E--F --L--M--Q--A--F--G-- L--Q--L--N--M--E--L-
phyb_arabidopsis GCTAAGGTATGCTTGTGAGT TTTTGATGCAGGCTTTCGGT TTACAGTTAAACATGGAATT
PHYB_GLYCINE     -L--R--Y--A--C--E--F --L--M--Q--A--F--G-- L--Q--L--N--M--E--L-
phyb_glycine     CTTGAGGTATGCTTGTGAGT TTCTGATGCAGGCGTTTGGG CTGCAGTTGAACATGGAGCT
PHYB_NICOTIANA   -L--R--Y--A--C--E--F --L--M--Q--A--F--G-- L--Q--L--N--M--E--L-
phyb_nicotiana   TCTTCGGTATGCCTGTGAAT TCCTTATGCAGGCCTTTGGA CTCCAATTGAACATGGAGTT
PHYB_SOLANUM     -L--R--Y--A--C--E--F --L--M--Q--A--F--G-- L--Q--L--N--M--E--L-
phyb_solanum     TCTTAGGTATGCATGTGAAT TCCTTATGCAGGCCTTTGGA CTCCAATTGAACATGGAGTT
PHYD_ARABIDOPSIS -L--R--Y--A--C--E--F --F--M--Q--A--F--G-- L--Q--L--N--M--E--L-
phyd_arabidopsis TTTGAGGTACGCTTGTGAGT TCTTTATGCAGGCCTTTGGC TTACAGCTAAACATGGAGTT
PHYE_ARABIDOPSIS -L--R--Y--A--C--E--F --L--M--Q--A--F--G-- L--Q--L--Q--M--E--L-
phye_arabidopsis GTTGCGGTATGCTTGTGAGT TTCTGATGCAAGCATTTGGG CTTCAGCTTCAAATGGAACT
PHYB_ORYZA       -L--R--Y--A--C--E--F --L--M--Q--A--F--G-- L--Q--L--N--M--E--L-
phyb_oryza       ACTACGGTATGCATGCGAGT TCCTCATGCAAGCCTTTGGG TTGCAGCTCAACATGGAGTT
PHYA_AVENA       -L--R--Y--A--C--E--F --L--A--Q--V--F--A-- V--H--V--N--R--E--F-
phya_avena       GCTGCGTTATGCTTGTGAGT TCTTAGCACAGGTGTTTGCT GTCCATGTCAACAGGGAGTT
PHYA_ORYZA       -L--R--Y--A--C--E--F --L--A--Q--V--F--A-- V--H--V--N--K--E--F-
phya_oryza       ATTGCGGTATGCCTGTGAGT TCTTAGCACAAGTGTTTGCT GTCCATGTTAACAAGGAGTT
PHYA_ZEA         -L--R--Y--A--C--E--F --L--A--Q--V--F--A-- V--H--V--N--K--E--F-
phya_zea         ACTGCGGTATGCCTGTGAAT TCTTGGCCCAAGTGTTTGCT GTCCATGTAAACAAGGAGTT
PHYA_ARABIDOPSIS -L--R--Y--A--C--E--F --L--A--Q--V--F--A-- I--H--V--N--K--E--V-
phya_arabidopsis TCTCAGGTATGCCTGTGAGT TTCTAGCTCAAGTGTTTGCC ATACACGTCAATAAGGAGGT
PHYA_CUCURBITA   -L--R--Y--A--C--E--F --L--A--Q--V--F--A-- I--H--V--N--K--E--L-
phya_cucurbita   TCTTAGGTATGCTTGTGAGT TCCTAGCTCAAGTATTTGCT ATTCATGTGAACAAGGAATT
PHYA_PISUM       -L--R--Y--A--C--E--F --L--A--Q--V--F--A-- I--H--V--N--K--E--I-
phya_pisum       TCTAAGGTATGCTTGTGAGT TTCTGGCTCAAGTGTTTGCC ATCCATGTGAACAAAGAAAT
PHYA_GLYCINE     -L--R--Y--A--R--E--F --L--P--Q--V--F--A-- D--H--V--H--K--E--I-
phya_glycine     TCTAAGGTATGCAAGAGAAT TTCTGCCTCAAGTATTTGCC GACCATGTGCACAAAGAAAT
PHYA_NICOTIANA   -L--R--Y--A--C--E--F --L--A--Q--V--F--A-- I--H--V--N--K--E--L-
phya_nicotiana   TCTGAGGTATGCATGTGAAT TTCTTGCGCAAGTCTTTGCC ATACACGTCAACAAGGAACT
```

```
PHYA_SOLANUM       -L--R--Y--A--C--E--F --L--A--Q--V--F--A-- I--L--V--N--K--E--L-
phya_solanum       ACTGAGGTATGCATGTGAGT TTCTTGCACAAGTCTTTGCC ATACTCGTTAACAAGGAACT
PHYC_ARABIDOPSIS   -L--R--Y--A--C--E--F --L--T--Q--V--F--G-- V--Q--I--N--K--E--A-
phyc_arabidopsis   ATTACGATATGCTTGTGAAT TCTTGACTCAAGTATTTGGC GTGCAGATCAACAAAGAAGC

MOUGEOTIA          -E--L--A--A--Q--H--R --E--K--H--I--L--R-- T--Q--T--L--L--C--D-
mougeotia          GGAGTTGGCTGCCCAGCACA GAGAGAAGCACATACTGCGG ACACAGACACTTCTGTGCGA
CHARA              -E--L--A--A--Q--I--R --E--K--H--I--L--R-- T--Q--T--L--L--C--D-
chara              CGAGCTTGCTGCTCAGATCA GGGAGAAGCATATCTTGAGA ACTCAGACTCTGCTGTGCGA
MARCHANTIA         -E--L--T--A--L--L--R --E--K--R--Y--L--R-- T--Q--T--L--L--C--D-
marchantia         CGAACTCACAGCTCTTTTGA GAGAAAAACGCTACCTCCGC ACGCAGACGCTCTTGTGTGA
FUNARIA            -E--L--A--A--Q--L--R --E--K--H--I--L--R-- T--Q--S--L--L--C--D-
funaria            AGAGTTAGCTGCTCAGCTAA GGGAAAAACATATTCTCAGA ACTCAAAGTCTTTTGTGCGA
PHYSCOMITRELLA     -E--S--A--A--Q--L--R --E--K--H--I--L--R-- T--Q--T--L--L--C--D-
physcomitrella     CGAGTCAGCCGCTCAGCTAA GGGAAAAACATATTCTCAGA ACTCAAACTCTACTTTGCGA
CERATODON          -E--L--A--A--Q--L--R --E--K--H--I--L--R-- T--Q--T--L--L--C--D-
ceratodon          TGAGCTGGCCGCTCAACTAA GGGAAAAACATATTCTCAGA ACTCAAACTCTTCTTTGTGA
SELAGINELLA        -A--V--A--A--H--V--R --E--K--H--I--L--R-- T--Q--T--L--L--C--D-
selaginella        GGCAGTGGCGGCGCACGTGC GGGAGAAGCACATACTGCGC ACCCAGACGCTCTTGTGCCA
EQUISETUM          -E--L--A--A--Q--M--R --E--K--H--I--L--R-- T--Q--T--L--L--C--D-
equisetum          TGAGCTAGCGGCGCAAATGA GGGAGAAACATATACTTCGC ACCCAAACACTCTTGTGCGA
MARSILEA           -E--L--A--A--Q--M--R --E--K--H--I--L--R-- T--Q--T--L--L--C--D-
marsilea           GGAACTGGCTGCACAGATGC GGGAGAAGCATATACTTCGC ACTCAAACTCTTCTATGTGA
ADIANTUM           -G--M--A--A--Q--V--R --E--K--H--I--L--R-- T--Q--T--L--L--C--D-
adiantum           GGGGATGGCTGCTCAGGTGA GAGAAAAGCACATCCTTCGC ACACAAACACTTCTTTGTGA
PSILOTUM           -E--L--A--A--Q--M--R --E--K--D--I--L--R-- T--Q--S--L--L--C--D-
psilotum           AGAACTAGCAGCTCAAATGA GAGAAAAGGATATTCTTCGC ACTCAAAGTCTTCTCTGTGA
PSILOTUM2          -E--L--A--A--Q--M--R --E--K--H--I--L--R-- T--Q--T--L--L--C--D-
psilotum2          CGAGTTGGCTGCTCAAATGC GAGAAAAACACATTCTTCGC ACTCAGACACTTCTATGCGA
METASEQUOIA        -Q--L--A--A--Q--L--T --E--K--N--I--L--R-- T--Q--T--L--L--C--D-
metasequoia        GCAATTGGCTGCTCAATTGA CAGAGAAGAACATTCTTAGG ACTCAAACACTTCTCTGCGA
PSEUDOTSUGA        -E--L--A--I--Q--M--K --E--K--R--I--L--R-- T--Q--T--L--L--C--D-
pseudotsuga        TGAATTGGCTATTCAAATGA AAGAGAAGCGAATTTTGCGC ACCCAAACATTGCTCTGCGA
PICEA              -Q--L--A--A--Q--L--T --E--K--H--I--L--R-- T--Q--T--L--L--C--D-
picea              GCAATTGGCAGCTCAATTAA CGGAGAAACACATTCTTAGG ACTCAAACGCTTCTCTGTGA
EPHEDRA            -E--L--A--C--Q--M--R --E--K--Q--I--L--R-- T--Q--T--L--L--C--D-
ephedra            TGAATTGGCCTGCCAGATGA GAGAGAAACAAATCCTCAGA ACCCAAACACTTCTTTGCGA
GNETUM             -Q--L--A--S--Q--M--T --E--K--H--I--L--K-- T--Q--T--L--L--C--D-
gnetum             GCAATTAGCGAGTCAGATGA CGGAGAAACATATCCTGAAA ACGCAAACGTTGCTTTGTGA
PHYC_PIPER         -E--L--A--A--Q--V--K --E--K--H--V--L--R-- M--Q--T--V--L--C--D-
phycc_piper        AGAATTAGCCGCGCAGGTGA AGGAGAAGCATGTACTGCGT ATGCAGACTGTGTTGTGTGA
PHYB_PIPER         -Q--L--A--S--Q--L--M --E--K--H--V--L--K-- T--Q--T--L--L--C--D-
phyb_piper         GCAGTTGGCGTCGCAGTTAA TGGAGAAGCATGTGCTCAAA ACCCAGACCTTGCTTTGCGA
PHYA_PIPER         -E--L--E--R--Q--I--R --E--K--N--I--L--R-- T--Q--T--L--L--C--D-
phya_piper         GGAACTGGAGAGACAAATTC GGGAGAAGAACATCTTGCGA ACCCAAACACTGCTCTGTGA
PHYB_ARABIDOPSIS   -Q--L--A--L--Q--M--S --E--K--R--V--L--R-- T--Q--T--L--L--C--D-
phyb_arabidopsis   GCAGTTAGCTTTGCAAATGT CAGAGAAACGCGTTTTGAGA ACGCAGACACTGTTATGTGA
PHYB_GLYCINE       -Q--L--A--A--Q--S--L --E--K--R--V--L--R-- T--Q--T--L--L--C--D-
phyb_glycine       TCAGTTGGCCGCGCAGTCGT TGGAGAAGCGGGTTTTGAGG ACACAGACTCTGTTGTGTGA
PHYB_NICOTIANA     -Q--L--A--S--Q--L--S --E--K--H--V--L--R-- T--Q--T--L--L--C--D-
phyb_nicotiana     GCAACTGGCATCACAGTTGT CTGAGAAACATGTTTTGAGG ACACAAACACTGTTATGTGA
PHYB_SOLANUM       -Q--L--A--S--Q--L--S --E--K--H--V--L--R-- T--Q--T--L--L--C--D-
phyb_solanum       GCAATTGGCGTCACAGTTGT CTGAGAAACATGTTTTAAGG ACACAAACACTGTTATGTGA
PHYD_ARABIDOPSIS   -Q--L--A--L--Q--V--S --E--K--R--V--L--R-- M--Q--T--L--L--C--D-
phyd_arabidopsis   GCAGTTAGCCTTGCAGGTGT CTGAAAAACGCGTTCTGAGA ATGCAGACACTATTATGTGA
PHYE_ARABIDOPSIS   -Q--L--A--S--Q--L--A --E--K--K--A--M--R-- T--Q--T--L--L--C--D-
phye_arabidopsis   TCAGTTAGCATCACAGTTAG CCGAGAAGAAGGCTATGCGG ACGCAGACCTTGTTGTGCGA
PHYB_ORYZA         -Q--L--A--H--Q--L--S --E--K--H--I--L--R-- T--G--T--L--L--C--D-
phyb_oryza         GCAGCTTGCACACCAACTGT CAGAGAAACACATTCTGCGG ACCGGAACACTGCTGTGTGA
PHYA_AVENA         -E--L--E--K--Q--L--R --E--K--N--I--L--K-- M--Q--T--M--L--S--D-
phya_avena         TGAATTAGAGAAACAGTTGC GTGAGAAGAACATACTGAAG ATGCAAACAATGCTCTCTGA
PHYA_ORYZA         -E--L--E--R--Q--V--R --E--K--S--I--L--R-- M--Q--T--M--L--S--D-
phya_oryza         TGAATTAGAGAGGCAAGTAC GCGAGAAAAGCATATTGAGG ATGCAAACAATGCTCTCTGA
PHYA_ZEA           -E--L--E--K--Q--I--R --E--K--N--I--L--R-- M--Q--T--M--L--S--D-
phya_zea           TGAATTGGAGAAGCAGATAC GAGAGAAAAACATTCTGCGA ATGCAAACAATGCTCTCTGA
PHYA_ARABIDOPSIS   -E--L--D--N--Q--M--V --E--K--N--I--L--R-- T--Q--T--L--L--C--D-
phya_arabidopsis   GGAACTCGATAACCAGATGG TGGAGAAGAACATTTTGCGC ACGCAGACACTCTTGTGCGA
PHYA_CUCURBITA     -E--L--E--N--Q--I--I --E--K--N--I--L--R-- T--Q--T--L--L--C--D-
phya_cucurbita     AGAGTTGGAAAATCAAATTA TAGAAAAGAATATTCTGCGT ACGCAGACACTCTTGTGCGA
PHYA_PISUM         -E--L--E--Y--Q--I--L --E--K--N--I--L--R-- T--Q--T--L--L--C--D-
phya_pisum         AGAGTTAGAATATCAGATTC TTGAGAAGAATATCCTGCGC ACGCAGACGCTGTTGTGTGA
PHYA_GLYCINE       -E--L--E--Y--Q--I--I --E--K--N--I--L--H-- H--P--G--H--L--L--C-
phya_glycine       AGAGTTAGAATATCAGATTA TTGAGAAGAATATCCTGCAC CACCCAGGGCACCTCTTGTG
```

```
PHYA_NICOTIANA    -E--L--E--S--Q--I--L --E--K--N--I--L--R-- T--Q--T--L--L--C--D-
phya_nicotiana    GGAATTGGAAAGTCAGATTC TTGAGAAAAATATCCTGCGT ACTCAGACTCTCTTGTGTGA
PHYA_SOLANUM      -E--L--E--N--Q--F--L --E--K--N--I--L--R-- T--Q--T--L--L--C--D-
phya_solanum      GGAATTGGAAAATCAATTCC TTGAGAAAAATATTCTGCGT ACTCAGACTCTCTTGTGTGA
PHYC_ARABIDOPSIS -E--S--A--V--L--L--K --E--K--R--I--L--Q-- T--Q--S--V--L--C--D-
phyc_arabidopsis GGAATCAGCTGTTCTGTTGA AAGAGAAGCGTATTTTGCAA ACTCAGAGTGTGCTATGTGA

MOUGEOTIA         -M--L--L--R--D-----A --P--M--G--I--V--S-- Q--S--P--N--?--?--?-
mougeotia         CATGCTGCTGAGGGAT---G CTCCCATGGGAATCGTCTCT CAGAGCCCCAAC????????
CHARA             -M--L--L--R--D-----A --P--T--G--I--V--T-- Q--S--P--N--I--?--?-
chara             CATGCTTCTCCGCGAC---G CACCCACTGGGATTGTCACA CAGTCACCCAACATC?????
MARCHANTIA        -M--L--L--R--D-----A --T--I--G--T--V--S-- Q--S--P--N--I--?--?-
marchantia        CATGCTGCTTCGCGAT---G CCACGATTGGAACTGTTTCT CAGTCACCTAACATA?????
FUNARIA           -M--L--L--Q--D-----A --P--I--G--I--V--S-- Q--V--P--N--I--?--?-
funaria           CATGCTTCTTCAAGAT---G CTCCCATTGGCATTGTGTCT CAGGTTCCAAATATT?????
PHYSCOMITRELLA    -M--L--L--R--D-----A --P--I--G--I--V--S-- Q--I--P--N--I--M--D-
physcomitrella    CATGCTTCTTCGAGAT---G CTCCTATTGGTATTGTTTCT CAGATTCCAAATATTATGGA
CERATODON         -M--L--L--R--D-----A --P--I--G--I--V--S-- Q--T--P--N--I--M--D-
ceratodon         CATGCTTCTTCGAGAT---G CTCCTATTGGAATTGTATCT CAAACTCCAAATATTATGGA
SELAGINELLA       -M--L--L--R--D-----A --P--I--G--I--V--S-- Q--S--P--N--I--M--D-
selaginella       CATGCTGCTGCGCGAC---G CTCCCATCGGCATTGTTTCT CAGTCGCCAAACATCATGGA
EQUISETUM         -M--L--L--R--D-----A --P--I--G--I--V--T-- Q--S--P--N--I--?--?-
equisetum         CATGCTACTCCGTGAT---G CGCCGATCGGGATCGTAACA CAATCGCCAAATATA?????
MARSILEA          -M--L--L--R--D-----A --P--I--G--I--V--S-- E--S--P--N--I--?--?-
marsilea          TATGCTCCTCAGAGAT---G CTCCAATTGGCATTGTTTCA GAGTCTCCCAATATC?????
ADIANTUM          -M--L--L--R--D-----A --P--I--G--I--V--S-- Q--S--P--N--I--M--D-
adiantum          TATGCTTCTCAGAGAT---G CTCCCATTGGTATCGTCTCG CAATCACCCAACATCATGGA
PSILOTUM          -M--L--L--R--D-----A --P--I--G--I--V--T-- R--S--P--N--V--M--D-
psilotum          TATGCTTCTTCGTGAC---G CTCCTATTGGCATTGTCACT CGTTCCCCGAACGTAATGGA
PSILOTUM2         -M--L--L--R--D-----A --P--I--G--I--V--T-- Q--S--P--N--I--M--D-
psilotum2         TATGCTTCTTAGAGAT---G CTCCTATCGGCATCGTCACG CAGTCCCCTAATATAATGGA
METASEQUOIA       -M--L--L--R--D-----A --P--I--G--I--V--T-- Q--S--P--R--S--?--?-
metasequoia       TATGCTTCTCCGAGAC---G CACCTATTGGAATTGTAACT CAGTCCCCCAGATCT?????
PSEUDOTSUGA       -M--L--L--R--D-----A --P--L--G--I--V--S-- Q--K--P--N--I--M--D-
pseudotsuga       CATGCTGCTTCGTGAT---G CACCCTTAGGAATTGTCAGT CAGAAGCCCAACATAATGGA
PICEA             -M--L--L--R--D-----A --P--I--G--I--V--T-- Q--S--P--S--I--?--?-
picea             TATGCTTCTCCGAGAT---G CCCCAATTGGAATTGTAACC CAGTCTCCCAGTATC?????
EPHEDRA           -M--L--L--R--D-----T --P--L--G--I--I--S-- Q--K--P--S--I--?--?-
ephedra           TATGCTTCTCAGGGAT---A CCCCTTTAGGAATCATCTCA CAAAAACCAAGTATA?????
GNETUM            -M--L--L--R--D-----A --P--I--G--I--V--T-- Q--S--P--S--I--?--?-
gnetum            TATGCTTCTGAGAGAT---G CCCCCATAGGGATTGTGACC CAATCTCCAAGCATC?????
PHYC_PIPER        -M--L--L--R--D-----A --P--M--G--I--M--T-- Q--S--P--N--V--?--?-
phycc_piper       CATGCTTCTCCGAGAT---G CTCCAATGGGAATCATGACA CAATCCCCAAATGT??????
PHYB_PIPER        -M--L--L--R--D-----S --P--T--A--M--V--T-- Q--S--P--S--I--?--?-
phyb_piper        CATGCTTCTTCGTGAC---T CCCCCACCGCGATGGTCACC CAGAGCCCGAGTATA?????
PHYA_PIPER        -M--L--L--R--D-----A --P--M--G--I--V--N-- Q--S--P--N--I--?--?-
phya_piper        CATGCTCCTCCGGGAC---G CCCCCATGGGCATTGTGAAC CAGAGTCCCAACATA?????
PHYB_ARABIDOPSIS -M--L--L--R--D-----S --P--A--G--I--V--T-- Q--S--P--S--I--M--D-
phyb_arabidopsis TATGCTTCTGCGTGAC---T CGCCTGCTGGAATTGTTACA CAGAGTCCCAGTATCATGGA
PHYB_GLYCINE      -M--L--L--R--D-----S --P--T--G--I--V--T-- Q--S--P--S--I--M--D-
phyb_glycine      TATGCTTCTTAGGGAC---T CGCCTACTGGCATTGTTACT CAGAGTCCTAGTATAATGGA
PHYB_NICOTIANA    -M--L--L--R--D-----S --P--T--G--I--V--I-- Q--S--P--S--I--M--D-
phyb_nicotiana    CATGCTCCTTCGAGAC---T CACCTACGGGGATTGTTATC CAGAGCCCCAGTATTATGGA
PHYB_SOLANUM      -M--L--L--R--D-----S --P--P--G--I--V--T-- Q--S--P--S--I--M--D-
phyb_solanum      CATGCTCCTTCGAGAC---T CTCCACCGGGGATTGTTACC CAAAGCCCCAGTATTATGGA
PHYD_ARABIDOPSIS -M--L--L--R--D-----S --P--A--G--I--V--T-- Q--R--P--S--I--M--D-
phyd_arabidopsis TATGCTTCTACGTGAC---T CACCAGCGGGGATTGTCACG CAGAGGCCTAGTATCATGGA
PHYE_ARABIDOPSIS -M--L--L--R--D-----T --V--S--A--I--V--T-- Q--S--P--G--I--M--D-
phye_arabidopsis TATGCTTCTCCGTGAT---A CTGTTTCCGCTATTGTTACA CAATCTCCGGGTATTATGGA
PHYB_ORYZA        -M--L--L--R--D-----S --P--T--G--I--V--T-- Q--S--P--S--I--M--D-
phyb_oryza        TATGCTACTCCGGGAT---T CACCAACTGGCATTGTCACA CAAAGCCCCAGCATCATGGA
PHYA_AVENA        -M--L--F--R--E--A--S --P--L--T--I--V--S-- G--T--P--N--I--M--D-
phya_avena        TATGTTGTTCCGAGAAGCCT CTCCCCTGACTATCGTATCA GGGACCCCCAATATCATGGA
PHYA_ORYZA        -M--L--L--R--E--S--S --P--L--S--I--V--S-- G--T--P--N--I--M--D-
phya_oryza        CATGCTTCTCAGGGAATCCT CTCCTCTGAGTATAGTATCA GGGACTCCCAACATCATGGA
PHYA_ZEA          -M--L--F--K--E--S--S --P--L--S--I--V--S-- G--S--P--N--I--M--D-
phya_zea          CATGCTGTTCAAGGAATCAT CTCCCTTGAGTATCGTGTCT GGGAGTCCAAATATCATGGA
PHYA_ARABIDOPSIS -M--L--M--R--D-----A --P--L--G--I--V--S-- Q--S--P--N--I--M--D-
phya_arabidopsis TATGCTGATGCGTGAT---G CTCCACTGGGTATTGTGTCG CAAAGCCCCAACATAATGGA
PHYA_CUCURBITA    -M--L--M--R--D-----A --P--L--G--I--V--S-- R--S--P--N--I--M--D-
phya_cucurbita    CATGCTAATGCGTGAT---G CTCCTTTAGGTATTGTGTCG AGGAGTCCTAACATAATGGA
PHYA_PISUM        -M--L--M--R--D-----A --P--L--G--I--V--S-- Q--S--P--N--I--M--D-
phya_pisum        TATGTTGATGCGAGAT---G CACCCTTAGGTATTGTATCA CAAAGCCCTAATATAATGGA
```

```
PHYA_GLYCINE      -M--L--M--R--D-----A --P--L--G--I--A--S-- E--S--P--N--I--M--D-
phya_glycine      TATGCTGATGCGAGAT---G CACCCCTAGGAATTGCATCA GAGAGTCCTAATATAATGGA
PHYA_NICOTIANA    -M--L--M--R--V-----A --P--L--G--I--V--S-- Q--S--P--N--I--M--D-
phya_nicotiana    TATGCTGATGCGAGTA---G CTCCCTTAGGTATAGTGTCA CAGAGTCCAAATATTATGGA
PHYA_SOLANUM      -M--L--M--R--D-----A --P--L--G--I--V--S-- Q--S--P--N--I--M--D-
phya_solanum      TATGCTGATGCGAGAT---G CTCCCCTAGGTATCGTGTCA CAGAGCCCCAACATTATGGA
PHYC_ARABIDOPSIS  -M--L--F--R--N-----A --P--I--G--I--V--T-- Q--S--P--N--I--M--D-
phyc_arabidopsis  CATGCTTTTCCGCAAT---G CACCAATAGGTATAGTCACT CAATCACCAAATATAATGGA

MOUGEOTIA         -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
mougeotia         ???????????????????? ???????????????????? ????????????????????
CHARA             -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
chara             ???????????????????? ???????????????????? ????????????????????
MARCHANTIA        -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
marchantia        ???????????????????? ???????????????????? ????????????????????
FUNARIA           -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
funaria           ???????????????????? ???????????????????? ????????????????????
PHYSCOMITRELLA    -L--V--K--C--D--G--A --A--L--Y--Y--G--K-- P--F--W--L--L--G--T-
physcomitrella    CTTGGTGAAATGCGATGGAG CTGCTCTTTATTATGGGAAG CCATTCTGGCTTCTGGGTAC
CERATODON         -L--V--K--C--D--G--A --A--L--Y--Y--G--K-- R--V--W--L--L--G--T-
ceratodon         TCTTGTGAAATGTGATGGAG CAGCTCTTTACTATGGGAAG CGAGTGTGGCTTCTTGGCAC
SELAGINELLA       -L--V--K--C--D--G--A --A--L--Y--Y--G--K-- R--F--W--L--L--G--I-
selaginella       CCTGGTCAAGTGCCACGGTG CCGCACTCTACTACGGCAAA CGCTTCTGGTTACTGGGCAT
EQUISETUM         -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
equisetum         ???????????????????? ???????????????????? ????????????????????
MARSILEA          -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
marsilea          ???????????????????? ???????????????????? ????????????????????
ADIANTUM          -L--V--T--C--D--G--A --A--L--Y--Y--G--K-- K--C--W--L--L--G--T-
adiantum          TTTGGTGACATGTGATGGTG CAGCATTGTATTATGGGAAG AAATGCTGGCTTTTAGGTAC
PSILOTUM          -L--V--K--C--D--G--A --A--L--Y--Y--G--E-- R--L--W--L--L--G--T-
psilotum          TCTTGTGAAGTGTGATGGGG CTGCGTTATATTACGGTGAA AGGCTTTGGTTGCTTGGAAC
PSILOTUM2         -L--V--K--C--D--G--A --A--L--Y--Y--G--K-- K--F--W--L--L--G--T-
psilotum2         TCTTGTGAAGTGTGATGGAG CTGCTTTGTATTACGGGAAA AAGTTTTGGTTGCTTGGAAC
METASEQUOIA       -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
metasequoia       ???????????????????? ???????????????????? ????????????????????
PSEUDOTSUGA       -L--V--K--S--D--G--A --A--?--?--?--?--?-- ?--?--?--?--?--?--?-
pseudotsuga       TCTAGTCAAATCTGATGGTG CTGCT??????????????? ????????????????????
PICEA             -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
picea             ???????????????????? ???????????????????? ????????????????????
EPHEDRA           -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
ephedra           ???????????????????? ???????????????????? ????????????????????
GNETUM            -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
gnetum            ???????????????????? ???????????????????? ????????????????????
PHYC_PIPER        -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
phycc_piper       ???????????????????? ???????????????????? ????????????????????
PHYB_PIPER        -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
phyb_piper        ???????????????????? ???????????????????? ????????????????????
PHYA_PIPER        -?--?--?--?--?--?--? --?--?--?--?--?--?-- ?--?--?--?--?--?--?-
phya_piper        ???????????????????? ???????????????????? ????????????????????
PHYB_ARABIDOPSIS  -L--V--K--C--D--G--A --A--F--L--Y--H--G-- K--Y--Y--P--L--G--V-
phyb_arabidopsis  CTTAGTGAAATGTGACGGTG CAGCATTTCTTTACCACGGG AAGTATTACCCGTTGGGTGT
PHYB_GLYCINE      -L--V--K--C--D--G--A --A--L--Y--F--Q--G-- N--Y--Y--P--L--G--V-
phyb_glycine      CTTGGTGAAGTGTGATGGGG CTGCCCTTTATTTCCAAGGG AACTATTATCCGTTGGGTGT
PHYB_NICOTIANA    -L--V--K--C--D--G--A --A--L--Y--C--Q--G-- K--Y--Y--P--L--G--V-
phyb_nicotiana    CCTTGTGAAGTGCGATGGCG CTGCTCTGTACTGCCAGGGG AAGTACTATCCATTAGGCGT
PHYB_SOLANUM      -L--V--K--C--D--G--A --A--L--Y--Y--Q--G-- K--Y--Y--P--L--G--V-
phyb_solanum      CCTTGTGAAGTGCGATGGTG CTGCTCTATACTACCAGGGG AAGTACTATCCATTAGGCGT
PHYD_ARABIDOPSIS  -L--V--K--C--N--G--A --A--F--L--Y--Q--G-- K--Y--Y--P--L--G--V-
phyd_arabidopsis  TTTAGTAAAATGTAATGGTG CGGCATTTCTTTACCAAGGG AAGTATTATCCGTTGGGTGT
PHYE_ARABIDOPSIS  -L--V--K--C--D--G--A --A--L--Y--Y--K--G-- K--C--W--L--V--G--V-
phye_arabidopsis  CCTTGTGAAATGTGATGGAG CTGCGTTATATTACAAGGGG AAATGTTGGTTGGTTGGTGT
PHYB_ORYZA        -L--V--K--C--D--G--A --A--L--Y--Y--H--G-- K--Y--Y--P--L--G--V-
phyb_oryza        CCTTGTGAAGTGTGATGGTG CTGCTCTGTATTACCATGGG AAGTACTACCCTCTTGGTGT
PHYA_AVENA        -L--V--K--C--D--G--A --A--L--L--Y--G--G-- K--V--W--R--L--R--N-
phya_avena        CCTAGTCAAATGTGATGGTG CTGCTCTTCTGTATGGGGGA AAAGTATGGCGTCTGCGTAA
PHYA_ORYZA        -L--V--K--C--D--G--A --A--L--L--Y--G--G-- K--V--W--R--L--Q--N-
phya_oryza        CCTTGTGAAATGTGATGGTG CTGCTCTTTTGTATGGGGGA AAAGTGTGGCGGCTACAGAA
PHYA_ZEA          -L--V--K--C--D--G--A --A--L--L--Y--G--D-- K--V--W--R--L--Q--T-
phya_zea          CCTAGTTAAGTGTGATGGCG CTGCTCTTTTGTATGGGGAC AAAGTATGGCGGCTTCAAAC
PHYA_ARABIDOPSIS  -L--V--K--C--D--G--A --A--L--L--Y--K--D-- K--I--W--K--L--G--T-
phya_arabidopsis  CCTTGTGAAATGTGATGGAG CAGCTCTCTTGTATAAAGAC AAGATATGGAAACTGGGAAC
PHYA_CUCURBITA    -L--V--K--S--D--G--A --A--L--L--Y--K--K-- K--I--W--R--L--G--L-
phya_cucurbita    TCTTGTCAAATCTGATGGGG CTGCCTTGTTATATAAGAAA AAAATTTGGCGATTAGGATT
```

```
PHYA_PISUM         -L--V--K--C--D--G--A --A--L--F--Y--R--N-- K--L--W--L--L--G--A-
phya_pisum         TCTAGTGAAATGTGATGGGG CTGCACTCTTTTATAGAAAC AAGTTATGGTTATTAGGAGC
PHYA_GLYCINE       -L--V--K--C--D--G--A --A--L--I--Y--R--N-- K--V--W--R--L--G--V-
phya_glycine       TCTAGTTAAATGTGATGGAG CTGCCCTCATATACAGGAAC AAAGTATGGAGATTAGGAGT
PHYA_NICOTIANA     -L--V--K--C--D--G--A --A--L--L--Y--K--N-- K--I--H--R--L--G--M-
phya_nicotiana     TCTTGTCAAATGTGATGGTG CTGCTTTGCTCTATAAGAAT AAGATACATCGACTTGGAAT
PHYA_SOLANUM       -L--I--K--C--D--G--A --A--L--L--Y--K--N-- K--I--H--R--L--G--M-
phya_solanum       TCTCATCAAATGTGATGGTG CTGCTTTGCTCTATAAGAAT AAGATACATCGATTAGGAAT
PHYC_ARABIDOPSIS   -L--V--K--C--D--G--A --A--L--Y--Y--R--D-- N--L--W--S--L--G--V-
phyc_arabidopsis   TCTTGTTAAATGTGATGGAG CAGCATTATATTACAGAGAC AACCTCTGGTCTCTAGGAGT

MOUGEOTIA          -?--?--?--?--?--?--
mougeotia          ???????????????????
CHARA              -?--?--?--?--?--?--
chara              ???????????????????
MARCHANTIA         -?--?--?--?--?--?--
marchantia         ???????????????????
FUNARIA            -?--?--?--?--?--?--
funaria            ???????????????????
PHYSCOMITRELLA     -T--P--T--E--S--Q--
physcomitrella     TACTCCTACGGAGAGTCAG
CERATODON          -T--P--T--E--N--Q--
ceratodon          GACACCGACTGAGAATCAG
SELAGINELLA        -T--P--S--E--A--Q--
selaginella        CACGCCCAGCGAGGCACAG
EQUISETUM          -?--?--?--?--?--?--
equisetum          ???????????????????
MARSILEA           -?--?--?--?--?--?--
marsilea           ???????????????????
ADIANTUM           -T--P--T--E--A--Q--
adiantum           TACACCAACAGAGGCGCAG
PSILOTUM           -T--P--T--K--A--Q--
psilotum           AACTCCCACAAAGGCTCAA
PSILOTUM2          -T--P--T--E--A--Q--
psilotum2          AACTCCCACGGAGGCTCAG
METASEQUOIA        -?--?--?--?--?--?--
metasequoia        ???????????????????
PSEUDOTSUGA        -?--?--?--?--?--?--
pseudotsuga        ???????????????????
PICEA              -?--?--?--?--?--?--
picea              ???????????????????
EPHEDRA            -?--?--?--?--?--?--
ephedra            ???????????????????
GNETUM             -?--?--?--?--?--?--
gnetum             ???????????????????
PHYC_PIPER         -?--?--?--?--?--?--
phycc_piper        ???????????????????
PHYB_PIPER         -?--?--?--?--?--?--
phyb_piper         ???????????????????
PHYA_PIPER         -?--?--?--?--?--?--
phya_piper         ???????????????????
PHYB_ARABIDOPSIS   -A--P--S--E--V--Q--
phyb_arabidopsis   TGCTCCTAGTGAAGTTCAG
PHYB_GLYCINE       -T--P--T--E--A--Q--
phyb_glycine       GACTCCAACTGAAGCTCAG
PHYB_NICOTIANA     -T--P--T--E--A--Q--
phyb_nicotiana     TACACCAACTGAAGCTCAG
PHYB_SOLANUM       -T--P--T--E--A--Q--
phyb_solanum       AACACCAACTGAAGCTCAG
PHYD_ARABIDOPSIS   -T--P--T--D--S--Q--
phyd_arabidopsis   GACTCCAACTGATTCTCAG
PHYE_ARABIDOPSIS   -T--P--N--E--S--Q--
phye_arabidopsis   TACTCCTAATGAGTCACAA
PHYB_ORYZA         -T--P--T--E--V--Q--
phyb_oryza         CACTCCCACAGAAGTTCAG
PHYA_AVENA         -A--P--T--E--S--Q--
phya_avena         TGCTCCAACGGAGTCTCAG
PHYA_ORYZA         -A--P--T--E--S--Q--
phya_oryza         TGCTCCAACTGAGTCTCAG
PHYA_ZEA           -A--P--T--E--S--Q--
phya_zea           GGCTCCAACCGAGTCTCAG
PHYA_ARABIDOPSIS   -T--P--S--E--F--H--
phya_arabidopsis   AACTCCAAGTGAGTTCCAC
```

```
PHYA_CUCURBITA    -T--P--N--D--F--Q--
phya_cucurbita    GACACCTAATGACTTCCAG
PHYA_PISUM        -T--P--T--E--S--Q--
phya_pisum        GACACCGACTGAATCTCAA
PHYA_GLYCINE      -T--P--S--E--P--Q--
phya_glycine      GACACCAAGTGAACCCCAG
PHYA_NICOTIANA    -T--P--S--D--F--Q--
phya_nicotiana    GACCCCAAGCGACTTTCAG
PHYA_SOLANUM      -N--P--S--D--F--Q--
phya_solanum      GAACCCAAGTGACTTTCAG
PHYC_ARABIDOPSIS  -T--P--T--E--T--Q--
phyc_arabidopsis  TACTCCCACAGAGACACAC
```

Appendix D

Alignment Of Nucleotide Sequences Of Monocot And Dicot Phytochrome PCR Clones

Source input fileC:\MEGA\DATA\MASTER.MEG
Title : phytochrome pcr clones
[ceratop = *Ceratophyllum*, cerato = *Ceratophyllum*, calamov = *Calamovilfa*, billber = *Billbergia*, aquile = *Aquilegia*, dianthu = *Dianthus*, dianth = *Dianthus*, curcurb = *Cucurbita*, arabido = *Arabidopsis*, hebest = *Hebestigma*, myrospe = *Myrospermum*, lycoper = *Lycopersicon*, antirrh = *Antirrhinum*, spinaci = *Spinacia*, pseudotsu = *Pseudotsuga*, a-e = *PHYA-PHYE*, aa & bb = duplicate *PHYA* & *PHYB*, uk = unknown]

```
equisetum TCGCGTTTCC TCTTCATGAA AAACCGAGTC CGAATGATAT GTGATTGCCG CACCAACCCT
gingko     ..T..G..T. .T..T...C. G...A....G A.......C. ........T. TG.A.CT...
a ceratop  ..TA.A..T. .G........ ...TAAG..T ..T....... ....C....A .G.A..A..C
b ceratop  C.T..C..T. .G....AAC. ...T...... ..CG.....G T......TAA TG.A..A..C
aa cerato  ?????????T .TG.T..... G..TAAG..T A.......T. ....C....A .G.A..A..A
a avena    G.CA.GC.T. .T........ G...AA...A ..G.....T. .......... TG.G.GAT.C
a hordeum  G.TACC..T. .G........ G...AA...G ..G.....T. .C...GT... .G.A.GAA.C
b hordeum  ..A..C.... .G...CG.C. ......T..G ..G.....TG C........A .G.AGCTG.G
c hordeum  ..CA.G..T. .G........ G..TAA...G ..G....... .......TGC TG.TTCA...
c calamov  ..CA.A..T. .G..T..... ....AA...G ..G....... ........TC .G....A..G
a oryza    G.CA.G..T. .T........ G...AA.... ..G.....T. .......... TG.A.GAT..
b oryza    ..A..C.... .G...CG.C. G.....T..G ..G.....TG C........A TG.TGCG..G
a zea      G.CA.G..T. .......... G...AA.... A.......C. .......... TG.A.GAT.C
a panicum  G.AA.G..T. .T........ G...AA.... ..C..C..CA ....C....A TG.G.GG..G
a lemna    ..C..C.... .....T.... G...AA...G A.G.....C. .C..C...AA GG.A..G..G
c lemna    ...A.G..T. .A...C.... G...AAG..T A.G.....C. .C..C..TTC .G.TC.G..G
c elodea   ...A.G..TT .G........ G...AAG..T A.G....... .CA......A TG.ATCA..A
a muscari  G....G.... .......... G.G.AAG... ..G.....C. .C..C..... .G....G..G
b muscari  ......G... .G.....A.C. G..TA.G..G A.G.....CG CC..C....A .G.GGTG..G
c muscari  ...A.G..T. .T........ G...AAG... A.G....... ....C...TC TG.TCCT..G
b billber  ...A.A..TT .G.....A.C. G...A....T A.G.....TG T...C...AA TG....G..A
a aquile   G.CA.A...T .A........ ...TAAGA.. ........C. ........T.. TG.A..G.A.
d aquile   ..AA.G..TT .G....AAC. G..TA....A A.........G T...C..T.A TG.T.CA..G
e aquile   G....G..T. .A..T.A.C. ...T..G..T .CG.....T. ........T.A AG.G.GTT.A
a urtica   G....G..T. .......... G...AAG... ........CG T...C...AA .G....G.AC
e urtica   G....A...T .G..T.G.C. ...T..G... ..G..C..C. .CA.C...AA .G..G.G..G
d quercus  ...A.G..TT .G..T.A.C. G..TA.G..T A.G.....TG TG.....TAA TG....TG..G
e quercus  G.C......T .A.....A.C. G..T.....A A.G.....T. .........A TG.A..AT..
a dianthu  ..AA.A...T .G..T..... G..TAA.... .......... .C..C..T.. GG.T..ATT.
b dianthu  ...A.G..TT .G..T.A.C. G..TA.G..T A..G....TG T...C..T.A TG.T..A..A
c dianthu  ..A..G..T. .T..TT.A.. ...TAAG..G A.......T. .C..C..T.T AT.TTCT..C
aa dianth  G....G.... .G........ G..TAA...T .......... ....C..T.. TG.A..ATT.
a cucurb   G.A.....TT .G........ ...TAAG... ..G.....TG T......T.. TG.A..A.A.
a arabido  G.C.....T. .G..T..... G...AAG... ..G.......G T.......AA TG.A..A.A.
b arabido  ..AA.G...T .G..T.A.C. G.....T... .........G TA......AA TG...CA...
c arabido  ...A.A..T. .G..T....G ....AAG..T A.G.....T. .......TTC AG.GGTT..G
d arabido  ..T..G...T .G....A.C. ......T..T A.G......G TA......TA TG.GTCA..G
e arabido  G.T..G...T .G....AAC. G.....T... ........T. ....C...AA TG.A.CT..G
a daucus   G.A..A..TT .A...CCTG. .G.TAAG..T ..T..T..C. .C.....T.. AG.G..GAA.
b daucus   ..TA.G..T. .G..TCA.C. ...T..G..T A.G.....TG TG.....T.A TG...CAT.A
a hebest   G.A..C..T. .G..T..... G...AAG... ..T......TG T........A TG.A..A.A..
e hebest   G.T..C...T .G.....A.C. ......T..T A.G......TA C.....C..A TG.A..G..A
a myrospe  G.A.....TT .A..T..... G...AAG... ..C......G T.......T.A TG.A..A.A.
b myrospe  ..AA.G..TT .G....AAC. G.....G..T A.G.....TG TG......AA TG.T.CT..A
e myrospe  G.T..C...T .G.....A.C. ......G... A.G.....T. ........TA TG.A..G...
a pisum    G....G..T. .A..T..... G...AAG... ..T......G T......TAA TG.A..A.A.
a lycoper  G.A..G..TT .G..T..... G..TAAG... ........T. .......... AG.A..G.A.
b lycoper  ..A..G..TT .G...T.A.C. G...A.G..G A.......TG TG..C....A TG.T.C....
d lycoper  ..AA.A...T .G..T.A.C. G...A.G..T A.......TG TG......AC TG...T....
e lycoper  G.C..C...T .G..A..AC. G...A....T A.G....... .......TAA TG.TC.G...
a solanum  G.A..G..TT .G..T..... G..TAAG... ........T. .......... AG.A..A.A.
b solanum  ..A..G..TT .G..T.A.C. G...A.G..G A.......TG TG..C....A TG.T.C....
a antirrh  G.A..G..TT .A........ G..TAAG... A.......C. ....C..T.. AG.G..A.A.
d antirrh  ..AA.A...T .G..T.A.C. G..T..G..G A.G.....TG TG..C....A .G.A.CT...
e antirrh  G.C.....TT .G.....A.C. G.....C..T A.G....... ....C...AA TG.TC.G..G
a spinaci  G.A..G...T .A........ G..TAAG... ..CC...... .......... TG.A.GATAC
```

```
bb daucus ..TA.G..TT .G..TCA.C. ...AT.G..T A.G.....TG TG.....T.A TG...TAT.G
uk daucus GT.A.A..T. .G..TCAA.. ...TAA.AGT A.G.....TG TA.......A TG.T.CATT.
gingko2   ..T.AA..T. .T..T..... G..TA....G A.......C. A......TT. TGTA.CT...
pseudotsu ..T..C..T. .G..T...C. G...A.G..G ........C. .C......AT GG.A.CT..C
b spinaci ..TA.G..TT .G.AT.A.C. ...T..G..T A.G......G TG.......A TGTT..T...

equisetum GTTCGAGTTA TCCAAGACAA GCAATTGGGT CAACCTATTA GCCTTGCCGG ATCCACTTTG
gingko    ...AAG.... .T.....TG. .G.....A.G .....CC.CT .TT.G.TA.. T..A..AC.A
a ceratop A..AAG..GT A......TG. .A.TC.TCCG AT.GA.T.G. CA...TGT.. ...A.T.C.T
b ceratop A.AGAGA... .T.....TTC A.GC...AA. .....CC..T ......T.AA T......C.T
aa cerato ..CAA...GC .......TG. .A.....TCA TTTGA.C... C.T.GTGT.. T..T...C..
a avena   A.AAAG..C. .TG.G.CTG. .GC.C.CCCC TTTGA..... ....ATGT.. T..AG.AC.C
a hordeum C.AAAG..C. ..GCC..TG. .GC.C.CCCC TTTGA..... ..T.GTGT.. T..AT.AC.C
b hordeum ..GA.G..C. ....G...CC TGC.A..CCA AGC.TGTGCT TAG...GGTC GA.TCT.CGC
c hordeum ..GAAGC.C. .T..G..TGG CA.CC..TCA ..G.....C. ....CTGT.. C..A..CA..
c calamov ..GAAGA.CC .T..G..TG. CAGCC.A.C. ..G...C.C. ....CTGT.. ......GC.C
a oryza   A.CAAGA... ..G....TG. .TCGC.CCAC TTGGA..... ..T.ATGT.. T..A..AC..
b oryza   ..GA.G..C. ....G..TCC TGC.C.AACA ..G..GC.GT ..T.G.TT.. G.....GC..
a zea     ..GAAGA... .TG....TG. .GC.C.CTCC ATTGA..... ..T.GTGT.. T..A...C.T
a panicum ..GAAGA... .TG....TGC .A..C.CCCC .TT.A.G... ..T.GTGT.. T..A...C.T
a lemna   ..GA.G..C? ?????????? ???????.CC GG.GAGG.G. ....CTG... G..G...C..
c lemna   .....GA.A. ....G..... .AGGC..TCC ..G..GT.G. .T..CTGT.. ...G..GC.T
c elodea  ....AC..A. .......... .AGG...... .....AT.G. .T...TGT.. C......C.A
a muscari ..GA.GA..G ....G...G. .A.GC..CCC TTCGAC..C. C.T.CTG... C......C.C
b muscari ..GA.G..GG ........G. .AGC....C. ..G..CC.GT .T..C.TG.. T..G..C...
c muscari ...AA...G. .A..T..... .A.G...CCA ..GT..C.G. .T..ATG... ...G......
b billber ..GA.T..C. .......TG. .ACT...ATG ..G...C.CT .T..C.TG.. G..T...C.T
a aquile  ..AA....CC .......TG. .A..C.ACCC TTCGAGT.G. C.T.GTGT.. T..A..CC..
d aquile  ....AT..G. .T.....TG. AGC.C.TATG ......C.AT ....A.T... C..A..AC..
e aquile  ...G.TA..G .T..GAGTGG AG.....AAA .G....T.GT .TT.G.TTAA T..A...C.T
a urtica  ..CAAG..GT ....G...G. ..CGC.TCCC TT.GACC.C. CG..CTG... T.....GC..
e urtica  ..C..G..GG .T...AGTG. AG.TC.AAAA .....C..GT .TT.G.TAAA C.....G..A
d quercus ...G.G..G. .T..G..TG. TAGGC.TATG ..G...T.GT ..T.G.TT.. T..A..GC.T
e quercus ..AA.GA.C. .T...GAGTG. AG..C.AAAG ..G...C..T ..T.G.T.AA T..A...C.T
a dianthu ..GA.G..G. ........G. .A..C.CTC. GTTGA.C.G. .TT.ATG... T..T...C.T
b dianthu ..CT.T..CG .G..G..TG. ..GGC..CAG ..G..GC..T .TT.G.TG.. G..T..GC..
c dianthu ..GAAG..GG .T....GAG. .G.C...TC. ......C.G. .T..C.GT.. T..A.....A
aa dianth ..CA.G..GT ........G. .A.GC.TTCG TCTGA.TGG. CTT.ATG... G..A...C.T
a cucurb  T.AAA...AC .......TG. .A....ACAG TTTGA.C.A. CTT.ATGT.. T..A.....A
a arabido .C.A.G..GC .T.....TG. AA.GC.TTCC TTTGACC... C.T.GTGT.. C.....CC.T
b arabido ....TT..GG ....G...G. TAGGC.AAC. ..GT....GT ..T.G.TT.. T..T...C.T
c arabido ...AA...CG .T.....T.. .AGTC.CTCA ..G..A..A. .T...T.T.. ...T......
d arabido .....T..GG .T......G. TAGGC.CACG ..GTT...AT ..T.G.TG.. T..G......
e arabido ...AAG...G .T..GAGTG. .G..C.CAAG AG...AC..T .TT.A.TTAA T..T...C.A
a daucus  ..GAAG..GC .A..G..TG. TA..C.TCCC TTCGAAC.G. CGT.GTGT.. C..A..CC.C
b daucus  ....CC..A. ....G..TG. TTC.C.TATG .....AT.GT .TT.G.TT.. T..A...C.T
a hebest  ..GAAG...C .T.....TG. AA..C.CCCA TTTGA.T.G. CTT.GTGT.. T..A..C..A
e hebest  ...AAG..C. .T..GAGTG. AG....AA.G ......C..T ..T.G.T.AA C..A..CC.T
a myrospe ..GAAG...C .T.....TG. AA..C.CCCA TTTGA.T.A. CT..GTGT.. T..A..C..A
b myrospe .....G..A. .T..G...G. AAGGC..A.G ..G...C.GT ..T.G.TT.. ...A..CC.T
e myrospe ...AAG..C. .T..GAGTG. AG....AA.G AG....C..T ..T.G.T.AA C..A..CC.T
a pisum   ..GAAG...C .T......G. AA..C.CCCA TTTGA.T.G. CT..GTG... T..G..C...
a lycoper ..GAAG..AG .......TG. .A.GC.TCCA TTTGA.T.A. CAT.GTG... C..T...C.T
b lycoper ..G..G.... CT..G..TG. ATC.C..ATG ..G...T.AT .T..A.TT.. T.....AC.T
d lycoper ..GA.G.... .T..G..TG. ATC.C.CATG ..G...T.AT .TT.A.TT.. C..A..AC.T
e lycoper ...AAG...G .T..GAGTG. .G..C.AAAG ..G.....CT ..T.G.TTAA T..G...C.T
a solanum ..GAAG..AG .......TG. .A.GC.TCCA TTTGACT.A. CAT.GTG... C..T...C.T
b solanum ..G..G.... CT..G..TG. ATC.C..ATG ..G...T.AT .T..A.TT.. T.....AC.T
a antirrh ..AAAG..G. .T.....TG. .A...C.TTC. GTTGA.C... C.T.ATGT.. C..G..AC.C
d antirrh ...A.G..CT .......TG. AGC...AATG ......T.GT .T..G.TA.. T..A..AC.C
e antirrh ...AAGA..G .....ACTG. AG..C.AAAG ..G...C..T ..T.G.TTAA C..T..CC.C
a spinaci ..G.AG..G. .......TG. .A..C.TCC. TTTGA.C.G. CT..ATGT.. C..G...C.T
bb daucus ....CC..A. ....G..TG. TTT.C.TATG ..T..CT.GT .TT.G.TT.. T..A...C.A
uk daucus ..C.CT..G. .T..G..TG. ATCTC.T.AC A.....T.AT .TT.A.GT.. ...A..CC.T
```

```
gingko2    A..AAG.... .T.....TG. .G...G...G .....CC.CT .TT.G.TA.. T..A..AC.A
pseudotsu  ..GAAGC... ....GTC.G. .G.....ATG ......C.AT .TT.A.TG.. T..A..GC.T
b spinaci  ..GTCC..G. .T..G..TG. A.GTC.ACAG ..G...C..T ..T.A.TT.. T.....AC.T

equisetum  AGAGCACCCC ATGGTTGCCA TGCACAATAT ATGGCGAATA TGGGTTCTAT CGCTTCCTTA
gingko     C.G..G.... ....G..T.. .........C .....A.... .......G.. T......C.T
a ceratop  .....T..T. ..A....... ..TC..G... ...CTT..C. .......AG. A.....T..G
b ceratop  C..T....T. .......T.. CATG...... .....T.... ....G..... T.....A...
aa cerato  .....C..A. ..A....T.. CTT...G..C ....A...C. ..AA...G.. T..A...C..
a avena    ..G.....A. .CA....T.. CCTT..G... ....A...C. ..AAC..G.. T..A...C.T
a hordeum  ..G...G.A. .CA.C..T.. CCTT..G... ....A...C. ..AAC..G.. T..A...C.T
b hordeum  TCCC...ATG GGT.CCATGC GCAGT.CATG GCAAAC.TGG G.TCGAT.GC AT..CT.G.C
c hordeum  ..G....... ....C..... ...C..G..C .....C..C. ....C..C.. ...G..GC..
c calamov  C.T..T..T. .......... ...G.....C .....A.... ....G..GG. A..G..GC.T
a oryza    ..G.....A. .CA....T.. .CTT..G... ....A...C. ..AAC..G.. T..A...C.T
b oryza    C.TT.G..G. .......... ..GC..G... ........C. ........C. T..A..TC.T
a zea      ........A. ..A.C..T.. CCTTA.G... ....A...C. ..AAC..G.. T..A...C.T
a panicum  ..G.....A. .CACG....G CCTT..G... ...AA...C. ..AAG..GG. T..G...C..
a lemna    C....T.... .C..C..... ...C..G..C ....A...C. ....C...G. G.....TC.G
c lemna    .....G.... ....A..... ...T..G... ...AA...C. ....C..C.. T.....TC.G
c elodea   .....T.... .......T.. C..T.....C .....C.... ....G..C.. T......C.T
a muscari  C.G..C.... .CA.C..... CCTC..G..C ....A...C. ..AAC..C.. ...C...C.C
b muscari  C....T.... ....C..... ...T..G..C ........C. ....C..GG. TAG...GC.T
c muscari  ..G..T..T. ....C..T.. ...C..G..C .....T.... .........G. T..C...C.T
b billber  C.C..C..T. ....A..... ...T..G..C .....C.... ....A..G.. T.....AC.T
a aquile   .....C..A. .CA.C..... .TT...G... ....A...C. ....A...C. T..C..TC.G
d aquile   C.T..C.... ....G..T.. ...T..G..C .....C.... ....A..G.. T..A..A...
e aquile   C.T..T..T. .......T.. CA....G... .....A..C. ....A..... TCG...GC.T
a urtica   C.G..C..G. .CA.C..T.. CCTT..G..C ........C. ..AAC..G.. ...G..GC.G
e urtica   C.T..C..G. .C..A..T.. ..TG.....C ...T.C..C. ....G..... G..C..G..G
d quercus  C.T..T.... ....C..T.. ...T..G... .....T.... ....C..A.. T..C..A..G
e quercus  C.TT.C..A. ...AA..T.. C..C...... ....AT.... ....C..... T..C..A..G
a dianthu  .......T. ....A..T.. ...T..G..C ...A...... ..AAC..C.. ..GG..T..G
b dianthu  ..G..G.... .C........ .T.T..G... .....T.... ....A..G.. T.....G..G
c dianthu  .....G..G. ....G..... .......G.. .....A..C. ....C..A.. ...G..TC.G
aa dianth  ..G..G..T. .C..C..T.. C..G.....C ....A..... ..AA...... T.GG..A..G
a cucurb   .....T..A. .CA....... .TT...G... ....AA..C. ..AAC..... A..C..T..G
a arabido  ........G. .CA.C..... .TTG..G..C .....C..C. ....A..A.. T..A..TC.G
b arabido  ..G..T..T. ........T.. CT.T..G... .....T..C. ....A..... T..G..T...
c arabido  .....T..T. .......T.. C.....G... ...AGT.... ....A..AG. G..G..TC.T
d arabido  C.....T..T. ....C..T.. ...T.....C ...A.T..C. ....C..... T..G..G...
e arabido  .....T..T. ....C..... .A.G..G... .......... ....C...G. A.....TC.T
a daucus   .....T..AT ..A....... .TT......C ....AA..C. ..AA...... T.....TC..
b daucus   ..G..T..T. .......... ...T..G..C .....T..C. ....A..A.. T..G..A..G
a hebest   ..G..T..T. ..A....... .TTG.....C .....A..C. ....A..A.. T......C.G
e hebest   ..GT....A. ...AA..T.. CAAC.GG... .....T..C. ....C..A.. T..C..TC.G
a myrospe  ..G..T..T. ..A.C..... .TTG.....C ........C. ....A..A.. T......C.G
b myrospe  C.T..C..A. .......T.. ...C..G... ...CAA..G. ....C..A.. T.....G..G
e myrospe  ..GT.T..A. .......T.. CA....G..C .....T..C. ....C..A.. T.....TC.G
a pisum    .....T..A. ..A....... .TTG..G..C .....T..C. ....A..A.. T.....G..G
a lycoper  ..G..C..T. .CTAC..... .TT...G... ....A...C. ..AA...A.. T..A..AC.T
b lycoper  .......... .......... ......G..C .....A.... ....G..... ...C..A...
d lycoper  ..G..T..T. .......T.. CC.T..G... ....TC.... ....AAA.G. T..G..G...
e lycoper  ...T....T. ...AA..... .AGCAGTA.. .....C..C. ....A..... AT.G..T..G
a solanum  ..G..C..T. .CTAC..... .TT...G... ....A...C. ..AA...A.. T..A..AC.T
b solanum  .......T. .......... C.....G..C .....A.... ....G..... T..C..A...
a antirrh  ..G..C..A. ..AC...... .TTG..G..C ....A...C. ...AC..A.. T..G..G..G
d antirrh  ..G..T..T. .C........ C..C..G..C ....AA.... ....G..G.. A.....T..G
e antirrh  C.T..G.... ....C..T.. .ATG..G... ....TC..C. ....C..A.. T.....A...
a spinaci  ..G..C..A. .......T.. ...T.....C ....A..... ..AAC..... T.GA..T..G
bb daucus  ..G..T..T. .......... ...T..G..C .....T..C. ....A..AG. T..G..A..G
uk daucus  C.G..T..T. ....C..... ...T.....C .....C..C. .......A.. T.GCA.A...
gingko2    CAGT.G.... ....G..T.G ....TC...C .....A.... .......A.. T........T
pseudotsu  C.G..G.... ....G..... C...C....C .....C..C. ........CG. T.....GC.T
```

```
b spinaci .....T..T. .......... .T.T..G... .....T.... ....G..C.. T.....T...

equisetum GTTATGGCCG TTGTTGTTAA TGACAATGAA AAATTGTGGG GTCTCATTGT TTGTCACCAT
gingko    ........T. .GA.AA.C.. ..GA.C...T ..GC.T.... ..T.GG.... G..C..T..C
a ceratop ........T. .G.....C.. ...AGGG... ..GC...... ....TG.A.. ...C..TA..
b ceratop ........TA ..A.AA.C.. ..GA.....C ..GC.C.... .G..GG.... ...C.....C
aa cerato ..C.....A. .G..CA.A.. ...GGGC..G .G...A.... ....TG.G.. A..C..TA.C
a avena   ..C.....T. ....G..... ...G...... ...C.A.... .C...C.... ...C......
a hordeum ..C.....T. ....G..... ...G.G.... .T.C.A.... .C..TG.... ...C......
b hordeum A.GGCA.TGA .CA.CAG..G C.GTGGG..G ..G....... .GT.AG.G.. G..C.....C
c hordeum ..G...T.G. .CAC...A.. C..GG....C ..GC.C.... .G..GG.G.. ...C..T..C
c calamov ..G...T.A. .CAC...A.. ...GG....G ..GC...... ....TG.G.. C..C..T..C
a oryza   ..C.....T. ....G..... ...G.....G ...C.A.... .A...C.... ...C......
b oryza   ........A. .GA.CA...G .AGTGG..GG ..G....... .GT.GG.A.. A..C.....C
a zea     ..C.....T. ....G..C.. ...A...... .GGC...... .......... ...C......
a panicum ..A..C..T. ....G..A.. C..AG...C. ..GC.A.... .C......A. ...C......
a lemna   ..G.....G. .G..G..C.. C...GGA..G CGGC.C.... .G...G.G.. ...C......
c lemna   ..C...T... .CACAA.C.. C...G.A.G. ..GC.C.... .AT.GG.G.. ...C.....C
c elodea  ..A...T.T. .G..GA.C.. ...GG.CATC ...C...... .GT.GG.C.. G..C..T...
a muscari ..C....... .C..C..... C..GG.GA.G CGTC.C.... .C...G.G.. G..C.....C
b muscari .CA...T... .GT.ATCA.C G.GAGCGAG. ..GC...... .C...G.C.. G..C..T..C
c muscari ..G...T.G. ..ACCA.... ...TG....T ..GC...... .CT.GG.G.. ...C..T..C
b billber ..C.....T. ..A.CA.C.. ..GA.G...T ..G....... .G...G.G.. ...C.....C
a aquile  ........G. .G..CA.... ...TGGC..T G.GGAAGA.. .GACT.GGC. ...GGGT.T.
d aquile  .CCC....A. .G..CA.C.. ..GG......T ..G....... .G..GG.... G.........
e aquile  ........A. .......C.. .TGGCG...T TCTC.A.... .GT.AG.A.. ...C..T...
a urtica  ..C.....G. .G..G..C.. C..GGGC... CGCC.C.... .AT.GG.C.. G..C...A.C
e urtica  .CGC...... .C..GA.A.. CC.A.CC..C CG.C.A.... .A...G.C.. G..C.....C
d quercus .CA.....G. ..A..A.C.. ..GA.....T .GG..A.... ..T.GG.G.. ......T..C
e quercus ..A.....AA ..A.CA.C.. ..CT.....T ..GC.T.... .AT.GG.A.. G..C.....C
a dianthu ..CC.....AA .C..G..... ...TG.G..T CGGC.T.... ..T.GG.G.. A..C..T...
b dianthu ..C.....G. .GA.CA.... C.GG.....C GC........ ..T.GG.... G.....T...
c dianthu ..G.....T. .GACGA.A.. ...TG.A..G ..G....... .AT.AG.... .........C
aa dianth ..A.....T. .C..G..... C..TG.G..C CGTC.T.... ..T.AG.C.. C..C..T...
a cucurb  ........A. ....G..... ...AGGG..T .G...A.... .CT.GG.A.. A......TA..
a arabido ........G. ....A..... C..GG.A..T .G.C.A.... ..T.AG.G.. .......A..
b arabido .CA.....G. ..A.AA.C.. ..GA...... .GGC.T.... ..T.GG.... ...C..T..C
c arabido ..C...T.T. .AAC.A.C.. ..GT.G...T C.C..A.... .CT.GG.G.. ......T..C
d arabido .C......A. ..A.AA.A.. ..GA..C... .GGC.T.... ..T.AG.... ...C.GT..C
e arabido .CAC.C..AA ....A..A.. A.G...A..T ..GC.T.... .AT.AG.... .G....T...
a daucus  ..A.....A. .C...A.C.. ...TTCG..T ...C.T.... .AT.AG.G.. G..C...A.C
b daucus  .C......A. ..A....C.. ..CG.C..GT .G.C.A.... .AT.AG.... ...C..T..C
a hebest  ........A. ....A..C.. .....GC..T CG.C.C.... ..T.GG.A.. ...C..TA.C
e hebest  ..G.....A. ..A.A..C.. .AGA.....T .GGC.T.... ..T.GC.A.. A.....T..C
a myrospe ...C....A. ....G..C.. .....G...T .G.C.A.... ..T.AG.A.. ...C..TA.C
b myrospe ..A.....A. .CA.AA.C.. C.GG.....T ..GC...... .C..GG.... ...C..T...
e myrospe ..A.....A. ..ACA..C.. ..GA.....T .C.C.T.... ..T.GC.A.. ......T..C
a pisum   ........A. ....C..C.. .....GC..T .G.C.T.... ..T.GG.A.. ......TA.C
a lycoper ..A.....A. ....G..C.. ....GGG..T .GGC.A.... .C..GG.... ...C...A.C
b lycoper ACAC....A. ..A..A.C.. ..GA......T .GGC...... .CT.GG.... .G.A.....C
d lycoper ACAC.T..T. .....A.A.. ..GA.....T .G...A.... ..T.GG.... .G.G.....C
e lycoper ........A. ..C..A.C.. .AG.GG...T TC.C.T.... ..T.G..... C.....T..C
a solanum ..A.....A. ....G..C.. ....GGG..T .GGC.A.... .CT.GG.... ..CC...A.C
b solanum ACAC....A. ..A..A.C.. C.GA.....T .GGC.A.... .CT.GG.... .G.A.....C
a antirrh ..A.....A. ....A..C.. C..GGGA..T .GGC.C.... .AT.AG.C.. ...C..TA.C
d antirrh ACAT....T. .CA.CA.... ..GA.....T .GGC.A.... .CT.GG.A.. .G.C..T...
e antirrh .CG.....G. ..C..AGC.. CA.TGG...C ..GC.T.... .GT.GG.A.. A..C.....C
a spinaci ........T. .......... ...TG.A..T .GGC.A.... ..T.AG.... C..C..T..C
bb daucus ...G....A. ..A.GTCA.T GAGACTG.TG .G.C.A.... .AT.AG.... ...C..T..C
uk daucus ..G.....T. ..A.C..CC. ..CT.....T .....A.... ....GG.C.C ...C..T..C
gingko2   ........T. .GA.AA.C.. ..GA.T...T ..GC.T.... ..T.GG.... G.TC..T..C
pseudotsu ........T. .GA..A.... ..GG.....T ..GC.C.... ..T.GG.... G..C......
b spinaci ........A. .GA.AA.C.. ..GA.G...T .G........ ..T.GG.... A..C..T...
```

```
equisetum ACAAGCCCAA GAGCGGTTCC CTACCCCTTC
gingko    ..CTC...TC .T...TCC.. ..T....C.T
a ceratop ..TTCA..G. ..TTT..C.. ..T....C..
b ceratop ...TCTG.TC .TTACC.C.T T.TT..A..A
aa cerato ...TCT.... .GTTT..... ..T....C..
a avena   GAG.....C. ..TAT..C.. T.TT..GC.G
a hordeum GAG.....C. ..TAT..C.. ..........
b hordeum ...TCT...C .GTGCA.C.. ..T....C..
c hordeum ..G.....G. .GTTC..C.T ACC..TTAG.
c calamov ..T.....G. .GTTT..C.. T?????????
a oryza   GAG.....C. ..TAT..... T.T...A..G
b oryza   ...TCT...C .GTGCA.C.. T.T...AC.A
a zea     GAG.....C. ..TAT..C.. T.T...AC.G
a panicum GGG..T..T. ..TAT..C.. ..T....C..
a lemna   GG.TC...C. .GTTC..C.. ..TT..TC.G
c lemna   .GC.....T. ..TTCA.C.. T.C...TCG.
c elodea  .....TT.G. .GTTT..C.. ..T....C..
a muscari GAG.....G. .GTTC..C.. ..TT...C..
b muscari ..TTCG.AG. .GTGCA.C.. ..T....C..
c muscari ..T..TT.C. .GTTT..C.. ..T....C..
b billber ..TTCT..TC ..TGTA.... .CTT...C.T
a aquile  GTTGT.TGTC ATAACAC.A. ACCAAGG..T
d aquile  ...TCTG.CC .CTGCA.C.. ..TT...C..
e aquile  ..TTCA..TC .GTAT..... .........A
a urtica  ..GTC...G. .GTTC..C.. ..T....C..
e urtica  .GCTCA..GC .GTAT..A.. .......C..
d quercus ..TTCTG.T. .GTGCA.... ..T....C..
e quercus ..CTC.A.CC .CTAT..C.. .......C..
a dianthu ..GTCT..TC .GTTT..... ..T....C..
b dianthu ..GTCT..TC .TT.TA.... T.TT..TC.T
c dianthu ......T.G. .GTTCA.... A..T..TC.A
aa dianth ...TCT..TC .GTTT..... .CTT...C..
a cucurb  T....T..C. ..TTT..... G.TT..TC.T
a arabido ..G.CT..G. .GTTT..... A.TT..TC..
b arabido ..TTCTT.TC .CTGCA.A.. G.TT..GC.A
c arabido G....T..T. ..TTT..... G.TT..A..A
d arabido ...TCAG.TC .TTGCA.A.. T.TT..T..G
e arabido TGTTCT..T. ..TAC..... A.T...G..A
a daucus  ..T..T.... .GTTT..G.. .......C.T
b daucus  ..GTCTG.T. ..AGCA.C.. .......C..
a hebest  ..T.CT..C. .GTTT..... .......C.T
e hebest  ..TTC..TGC .TTAT..G.. .......C..
a myrospe ..T.CT..T. .GTTT..... ..T....C..
b myrospe ..TTCTG.TC ..TGCA.... ..T....C.T
e myrospe ..TTCA...C .CTAT..G.. ..T.......
a pisum   ..T.CT.... .GTTT..... T.TT..TC.A
a lycoper ..G.C..... .GTTC..... T.....AC.G
b lycoper ..TTCTGT.C .GT.CA.... T.TG..TC.T
d lycoper .GTTCAG.TC .GTTCA.... T.T....C.T
e lycoper ..TTCT..TC .GTAT..G.. T.T...AC.T
a solanum ..G.C..... .GTTC.CG.. ..T...AC.G
b solanum ..TTCTGTTC .GT.CA.... T.T...TC.T
a antirrh ..G.CT.... ..TTT..C.. .......C..
d antirrh ..TTCTG.TC ..TGCA.... .CTA...C..
e antirrh ..CTCA...C .CTAC..C.. .......C..
a spinaci ...TCT.... .GTTT..... .CTT...C.T
bb daucus .TGTCTG.T. ..AGCA.C.. ...T...C..
uk daucus .A.TCTG.TC .GT.CC.... A.T...A???
gingko2   .TCTC...CT .T..A..C.. ..T....C.T
pseudotsu ..CTC...CG CG..T..... .......C..
b spinaci ...TCT..GC .GT.CA.... T.T...G..G
```

Appendix E

Alignment Of Amino Acid Sequences Of Poaceae

LABELS:
SM = *Selaginella*, MS = *Muscari*, FI = *Flagellaria*, JA = *Joinvillea*, TC = *Thamnochortus*, AP = *Aristida*, AS = *Avena*, BB = *Bambusa*, BG = *Bouteloua*, BI = *Bromus*, CL = *Calamovilfa*, DU = *Danthonia*, EC = *Eragrostis*, HV = *Hordeum*, OS = *Oryza*, PA = *Phragmites*, PC = *Panicum*, PP = *Poa*, PS = *Pennisetum*, SV = *Stipa*, ZM = *Zea*, A, B, C, & U = *PHYA*, *PHYB*, *PHYC*, & *PHYU*

```
SM  QASRFLFMKNRVRMLCDCSAPPVKITQDKELRQPISLAGSTLRAPHGCHAQYMGN
MSB QASRFLFKQNRVRMIADCHAVPVRVVQDESLAQPLCLVGSTLRAPHGCHAQYMAN
FIB QASRFLIKQNRVRMIADCHAPSVRVIQEPSLPQPLCLVGSTLRSPHACHAQYMAN
JAB QASRFLFRQNRVRMIADCHAAXVRVIQDPALPQPLCLVGSTXRAPHGCHAQYMAN
BGB QASRFLFRQNRVRMIADCHATPVRVIQDPGLSQSLCLVGSTLRAPHECHAQYMAN
ECB QASRFLFRQNRVRMIADCHATPVRVIQDPGLSQPLCLVGSTLRAPHGCHAQYMAN
DUB QASRFLFRQNRVRMIADCHATAVSVIQDPGLQQALCLVGSTLRAPHGCHAQYMAN
BBB QASRFLFRQNRVRMIADCHAAPVRVIQDPTLPQPLCLVGSTLRAPHGCHAQYMAN
BIB QASRFLFRQNRVRMIADCHAAPVRVIQDPAMPQPLCLVGSTLRSPHGCHAQYMAN
HVB QASRFLFRQNRVRMIADCHAAAVRVIQDPAMPQALCLVGSTLRSPHGCHAQYMAN
PPB QASRFLFRQNRVRMIADCHAAPVRVIQDPAMPQPLCLVGSTLRSPHGCHAQYMAN
SVB QASRFLFRQNRVRMIADCHAAPVRVIQDPAMQQPLCLVGSTLRSPHGCHAQYMAN
OSB QASRFLFRQNRVRMIADCHAAPVRVIQDPALTQPLCLVGSTLRSPHGCHGQYMAN

MSC QASRFLFMKNKVRMICDCSAPPVKVIHDKKLPQSLSLCGSTLRAPHGCHAQYMAN
FIC QASRFLFMKNKVRMICDCSAPPVKVVQDDRLPQPLSLCGSTLRAPHGCHAQYMAN
JAC QASRFLFMKNKVRMICECSAPPVKVVQDDRLAQPLSLCGSTLRAPHGCHAQYMAN
TCC QASRFLFMKNKVRMICDCAAPPVKVVQGERLAQPLSLCGSTLRAPHGCHAQYMAN
APC QASRFLFMKNKVRMICDCSATPVKIIQDDGLAQPLSLCGSTLRAPHGCHAQYMAN
DUC QASRFLFMKNKVRMICDCSATPVKIIQDDSLAQPLSLCGSTLRAPHGCHAQYMAN
PAC QASRFLFMKNKVRMICDCSATPVKIIQDDSLSQPLSLCGSTLRAPHGCHAQYMAN
PSC QASRFLFMKNKVRMICDCSAVPVQIIQDDTLAQPLSLCGSTLRAPHGCHAQYMAN
ZMC QASRFLFMKNKVRMICDFSATPVLIIQDGSLAQPLSLCGSTLRASHGCHAQYMAN
ECC QASRFLFMKNKVRMICDCSVKPVKIIQDDSLAQPLSLCGSTLRAPHGCHAQYMAN
CLC QASRFLFMKNKVRMICDCSAKPVKILQDDSLAQPLSLCGSTLRAPHGCHAQYMAN
BGC QASRFLFMKNKVRMICDCSAKPVKIIQDDSLAQLSSLCGSTLRAPHGCHAQYMAN
BBC QASRFLFMKNKVRMICDSSATPVKLIQDDNLAQPVSLCGSTLRAPHGCHAQYMAN
OSC QASRSLFMKNKVRMICDCSATPVKIIQDDSLTQPISICGSTLRAPHGCHAQYMAS
SVC QASRFLFMKNKVRMICDCAATPVKLIQDDNLSQPISLSGSTMRAPHGCHAQYMAN
PPC QASRFLFMKNKVRMICDCAATPVKLIQDDNLSQPISLCGSTMRAPHGCHAQYMAN
BIC QASRFLFMKNKVRMICDCASTPVKLIQGDNLSQPISLCGSTMRAPHGCHAQYMAN
HVC QASRFLFMKNKVRMICDCAASPVKLIQDGNLSQPISLCGSTMRAPHGCHAQYMAN

MSA QAARFLFMKSKVRMICDCRAKPVRIVQDEKLPFDITFCGSTLRAPHSCHLQYMEN
FIA QAARFLFMKNKVRMICDCRAKPVKIYQDESLPFDISLCGSTLRAPHSCHLQYMEN
PCU QAARFLFMKNKVRIISDCHARPVKIIEDAKLPLHVSLCGSTLRAPHTCRLQYMKN
PCA QAARFLFMKNKVRMICDCRARSVKILEDEALSIDISLCGSTLRAPHSCHLQYMEN
APA QAARFLFMKNKVRMICDCRARSVKIIEDEALSIDISLCGSTLRAPHSCHLQYMEN
PAA QAARFLFMKNKVRMICDCRARSVKIIEDEGLSIDISLCGSTLRAPHSCHLQYMEN
PSA QAARFLFMKNKVRMICDCRARSVKIIEDEALSIDISLCGSTLRAPHSCHLQYMEN
ZMA QAARFLFMKNKVRMICDCRARSVKIIEDEALSIDISLCGSTLRAPHSCHLKYMEN
DUA QAARFLFMKNKVRMICDCRARSVKIIEDEAISTDISLCGSTLRAPHSCHLQYMEN
BGA QAARFLFMKNKVRMICDCRARSVKIIEDEAASIDISLCGSTLRAPHSCHLQYMEN
CLA QAARFVFMKNKVRMICDCRARSVKIIEDEAVSINISLCGSTLRAPNSCHLQYMEN
ECA QAARFLFMKNKVRMICDCRARSVKVIKDEAASIDVSLCGSTLRAPHSCHLQYMEN
OSA QAARFLFMKNKVRMICDCRARSIKIIEDESLHLDISLCGSTLRAPHSCHLQYMEN
PPA QAARFLFMKNKVRMICDCRARSIKVIEAEALPFEISLCGSALRAPHNCHLQYMEN
SVA QAARFLFMKNKVRMICDCRARSIKIIEDETLPFDISLCGSALRAPHSCHLQYMEN
HVA QAATFLFMKNKVRMICDVRARTLKVIADEALPFDISLCGSSLRAAHSCHLQYMEN
ASA QAARLLFMKNKVRMICDCRARSIKVIEAEALPFDISLCGSALRAPHSCHLQYMEN
SM  MGSVASLVMAMIINDNDEPSGGGGGRG------QHKGRRLWALVVCHHTSPRSVP
MSB MGSVSSLAMSVIINGSEEDSTRN-------------TMKLWGLVVCHHTSQRCIP
FIB MGSIASLVMAVIISSSSDDSSSAVS----------SGTKLWGLVVCHHTSPRCIP
JAB MGSIASLVMAVIISHGADDDNVPRTG-TS------SAMKLWGLVVCHHTSPRCIP
```

```
BGB MGSIASLVMAVIISSGGDDEQTSRSG-IS------SAMKLWGLVVCHHTSPRFIP
ECB MGSIASLVMAVIISSGGDDEQTSRSG-IS------SAMKLWGLVVCHHTSPRFIP
DUB MGSIASLVMAVIISSGSDDEQMAQGG-IS------SAMKLWGLVVCHHTSPRFIP
BBB MGSIASLVMAVIISSGGDDEHNMTRGGVP------SAMKLWGLVVCHHTSPRCIP
BIB MGSIASLVMAVIISSGGEDEHNMTRGVIP------SAMKLWGLVVCHHTSPRCIP
HVB MGSIASLVMAVIISSGGEDEHNMTRGVIP------SAMKLWGLVVCHHTSPRCIP
PPB MGSIASLVMAVIISSGGEDEHSMGRGAVP------SAMKLWGLVVCHHTSPRCIP
SVB MGSIASLVMAVIISGGGDDEHNMGRGAIP------SAMKLWGLVVCHHTSPRCIP
OSB MGSIASLVMAVIISSGGDDDHNIARGSIP------SAMKLWGLVVCHHTSPRCIP

MSC MGSVASLVMSVTIN---DDDDEPGTD--------QKGRKLWGLVVCHHTSSRFVP
FIC MGSTASLVMSVTVN--EEGDDDTGSD-QQQQQQQRGRKLWGLVVCHHTSPRFVP
JAC MGSIASLVMSVTIN--EDDDDDTGSD------QQQKGRKLWGLVVCHHTCPRFVP
TCC MGSIASLVMSVTVN--EDDGDDTGND------QQQKGQKLWGLVVCHHTSPRFVP
APC MGSVASLVMSVTVNEDDEEDGDTGSD------QQPKGRKLWGLVVCHHTSPRFVP
DUC MGSIASLVMSVTINEDEEENGDTGSD------QQPKGRKLWGLVVCHHTSSRFVP
PAC MGSVASLVMSVTINEDEEEDRDTGSD------QQPKGRKLWGLVVCHHTSPRFVP
PSC MGSVAPLVMSVTINEDDEEDGDTGSD------QQPKGRKLWGLVVCHHTSPRFVP
ZMC MGSVASLVMSVTINDDEEEDGDTDSD------QQPKGRKLWGLVVCHHTSPRFVP
ECC MGSVASLVMSVTVNEDEDDDGDTGSD------QQPKGRKLWGLVVCHHTSPRFVP
CLC MGSVASLVMSVTVNEDEDDDAETRSD------QQPKGRKLWGLVVCHHTSPRFVP
BGC MGSVASLVMSVTVNEDEEDEADNGSD------QQPRGRKLWGLVVCHHTSPRFVP
BBC MGSVASLVMSITINEDEEEDGNTLCD------QQPKGRKLWGLVVCHHTSPRFVP
OSC MGSVASLVMSVTINEDEDDDGDTGSD------QQPKGRKLWGLMVCHHTSPRFVP
SVC MGSVASLVMSITINEDEEEDGDTGSD------QQPKGRKLWGLVVCHHTSPRFVP
PPC MGSIASLVMSITINEDGEEDGETGSD------QQPKSRKLWGLVVCHHTSPRFVP
BIC MGSIASLVMSITINEDEDEDGDTGSD------QQPKGRKLWGLVVCHHTSPRFVP
HVC MGSIASLVMSVTVNEDDDEDGDTGSD------QQPKGRKLWGLVVCHHTSPRFVP
MSA MNSIASLVMAVVVNEEAEEAERDPEAE--QS--QQPKRKRLWGVVCHHESPRFVP

FIA MNSIASLVMAVVVNERGEEDEPEPG----QP-QQQKRKRLWGLVVCHNESPRFVP
PCU MKSVASLVIAVVVNEDAEDVEVVIK----QQTQHHQKKKLWGLIICHHGSPRYVP
PCA MNSIASLVMAVVVNENEEDDEPNPE----QP-QQQQKKRLWGLLVCHHESPRYVP
APA MKSIASLVMAVVVNENEEGDEAEPG----QPPQQQQKKKLWGLLVCHHESPRYVP
PAA MNSIASLVMAVVVNENEEDDEAEPG----QPSQQLKKKKLWGLIVCHHESPRYVP
PSA MNSIASLVMAVVVNENEEDDEPEPE----QPPQQQKKKRLWGLIVCHHESPRYVP
ZMA MNSIASLVMAVVVNENEEDDEPEPE----QPPQQQKKKRLWGLIVCHHESPRYVP
DUA MNSIASLVMAVVVNENEEDDEAEPE----QPPQQQKKKRLWGLIVCHHESPRYVP
BGA MNSIASLVMAVVVNEDEEDDEVEPE----QPSQQQKKKKLWGLVVCHHESPRYVP
CLA MNLIASLVMVVVVNENEEDDEAETE----QPPQQQKKRKLWGLIVCHHESPRYVP
ECA MNSIASLVMAVVVNENEEGDEPEPE----QPPQQQKKKKLWGLIVCHHESPRYVP
OSA MNSIASLVMAVVVNENEDDDEVGAD----QPAQQQKRKKLWGLLVCHHESPRYVP
PPA MNSIASLVMAVVVNENDEDDEVESD----QPTEQQKKKKLWGLLVCHHESPRYVP
SVA MNSIASLVMAVVVNENEEDDEVEAE----QPAQQQKRKRLWGLLVCHHESPRYVP
HVA MNSIASLVMAVVVNESEEDDEAESE----QPAQQQKKKILWGLVVCHHESPRYVP
ASA MNSIASLVMAVVVNENEEDDEAESE----QPAQQQKKKKLWGLLVCHHESPRYVP
```

Appendix F

Phytochrome Nucleotide Data From Poaceae

PHYTOCHROME DNA DATA FROM POACEAE, ALL GENES IN SEQUENCE OF PHYC, PHYB, PHYA

Muscari	TCGAGGTTTC	TTTTCATGAA	GAACAAGGTC	AGGATGATAT	GTGACTGCTC
Aristida	..C.......	....T.....	......A..G	C.........	.C..T.....
Avena	??????????	??????????	??????????	??????????	??????????
Bambusa	..C.......	.G..T.....	......A..G	C.........	....T.CG..
Bouteloua	..C..A....	.A..T.....	......A..G	C.........	....T.....
Bromus	..C.......	.G..T.....	......A..G	C.........	....T..TG.
Calamovilfa	..C..A....	.G..T.....	A.....A..G	C.........	....T.....
Danthonia	..C.....C.	.G..T.....	......A..G	C.A.......	.C..T.....
Eragrostis	..C..A....	.G..T.....	A..T..A..G	....	....T.....
Flagellaria	..CC......	.G..T.....	.........G	C.......C.	.C..T.....
Hordeum	..C.......	.G........	...T..A..G	C.........	....T..TG.
Joinvillea	..C.......	.C..T.....	......A..A	C.A.....C.	.C..A.....
Oryza	..C......T	.G........	......A..G	C.........	.C..T.CG..
Phragmites	..C.......	.G..T.....	...T..A..G	C.........	.C..T.....
Panicum	??????????	??????????	??????????	??????????	??????????
Poa	..C.......	.G........	......A..A	....	....T..TG.
Pennisetum	..C.......	.G..T.....	......A..G	C.........	.C..T.....
Stipa	..C.......	.G..T.....	......A..G	C.........	....T..TG.
Thamnochortus	....	.C........	......A..G	........C.	.C..T...G.
Zea	..C.......	.G..T.....	......A..G	C.........	....T.T...

Muscari	TGCTCCTCCG	GTTAAAGTGA	TACATGACAA	GAAGTTGCCA	CAGTCTCTGA
Aristida	...CA....T	..G..GA.C.	.T..G..TG.	TGGCC.AG..	..AC....C.
Avena	??????????	??????????	??????????	??????????	??????????
Bambusa	C...A....T	..G..GC.C.	.T..G..TG.	C..CC.AG..	...C..G.C.
Bouteloua	...CAAA..T	..G..GA.C.	.T..G..TG.	C.GCC.AG.G	...CTCTCT.
Bromus	.T..A....T	..G..GC.C.	.T..G.GTG.	C..CC.AT..	...C..A.C.
Calamovilfa	C..CAAA...	..G..GA.CC	.T..G..TG.	C.GCC.AG.T	...C....C.
Danthonia	...CA....T	..G..GA.C.	.T..G..TG.	C.GCC.AG..	...C....C.
Eragrostis	..TGAAG..T	..G..GA.C.	.T..G..TG.	C.GCC.AG..	...C....C.
Flagellaria	G........T	..C..G..TG	.C..G...G.	C.G.C....T	...C....C.
Hordeum	....T.A..T	..G..GC.C.	.T..G..TGG	C..CC..T..	...C..A.C.
Joinvillea	C........A	.....G..TG	.C..G..TG.	C.G.C.AG.C	...C....C.
Oryza	C..AA.G..T	..G..GA.C.	.C..A..TG.	C.GCC.AA..	..AC..A.A.
Phragmites	....A....T	..G..GA.C.	.T..G..TG.	C.GCC.AT..	...C....C.
Panicum	??????????	??????????	??????????	??????????	??????????
Poa	...CA....T	..G..GC.C.	.T..G..TG.	C..CC.TT..	...C..A.C.
Pennisetum	...CGT...A	..GC.GA.C.	.T..G..TG.	T.CCC.AG..	...C....C.
Stipa	....A....T	..G...C.C.	.T..G..TG.	C..CC.AT..	...C..A.C.
Thamnochortus	....	........AG	.T..G.GTG.	.CGTC.TG.T	...C......
Zea	...CA....A	..GCTGA.C.	.T..G..TGG	C.GCC.TG..	...C.C..C.

Muscari	GTCTATGCGG	ATCGACTTTG	AGGGCTCCTC	ATGGCTGTCA	TGCCCAGTAC
Aristida	.C..C.....	T..A..AC.T	C.......C.	....T..C..	...A......
Avena	??????????	??????????	??????????	??????????	??????????
Bambusa	.C..C..T..	C.....CC.C	.....A..C.	....T..C..	...G......
Bouteloua	.C..C..T..	C..T..AC.C	C.......C.	....T..C..	...G..A...
Bromus	.C..C..T..	C.....CA..	.....A..C.	....T..C..	C.........
Calamovilfa	.C..C..T..	...C..GC.C	C.T.......	....T..C..	...G..A...
Danthonia	....C.....	C..C..GC.C	C.......C.	....T..C..	....

```
Eragrostis    .C..C..T.. T.....AC.C C.......C. ....T..C.. ...G..A...
Flagellaria   .C..C..... C..T..C... ........C. .......C.. C..G......
Hordeum       .C..C..T.. C..A..CA.. .....A..C. .......C.. ..........
Joinvillea    .C..C..... C..T..C... ........C. ....T..C.. ...A......
Oryza         .CA....... ...T...C.C .....A..C. ....T..C.. ...A......
Phragmites    .C........ C..C..AC.C C.......C. ....T..C.. ...G......
Panicum       ?????????? ?????????? ?????????? ?????????? ??????????
Poa           .C..C..T.. C.....CA.. .....A..C. ....T..C.. ......A...
Pennisetum    .C.....T.. T..C..GC.C C.......C. ....T..C.. C..G..A...
Stipa         .C..C.CT.. C.....CA.. .....A..C. ....T..C.. ..........
Thamnochortus .C..C..... C..T...... .......... ....T..C.. C..A..A...
Zea           .C..C..T.. T..T..CC.C ......T.C. ....T..C.. ...A......

Muscari       ATGGCTAATA TGGGTTCTGT TGCCTCCCTT GTGATGTCGG TTACCATTAA
Aristida      .....A..C. .......... ...T..G... ........C. .C..TG.A..
Avena         ?????????? ?????????? ?????????? ?????????? ??????????
Bambusa       .....G..C. ....C..C.. ...G..A... .........A .C..T..A..
Bouteloua     .....A..C. ....G..C.. ...A..G... ........A. .C..TG.A..
Bromus        .....C..C. ....C..CA. C..A..G..A .........A .C..T..A..
Calamovilfa   .....A.... ....G..G.. A..G..G... ........A. .C..TG.A..
Danthonia     .....A..C. .......AA. ...A..G... ........A. .C..T..A..
Eragrostis    .....A.... ....G..C.. ...G..G... ........A. .C..TG.A..
Flagellaria   .....A.... ....C..CAC C..A..G... ..C.....A. .C...G.A..
Hordeum       .....C..C. ....C..CA. C..G..G..A .......... .C..TG.A..
Joinvillea    .....A.... ....C..CA. ...A..G... ........A. .C.....A..
Oryza         .....A.GC. ....C..C.. ...A..A... .......... .C..T..A..
Phragmites    .....A..C. .......C.. ...A..A... .......... .C..T..A..
Panicum       ?????????? ?????????? ?????????? ?????????? ??????????
Poa           .....C..C. ....C..CA. C..G..A..G .........A .C..T..A..
Pennisetum    .....A..C. ....C..... ...GC.G... ........A. .A..T..A..
Stipa         .....G..C. ....C..... ...G..G... .........A .C..T..A..
Thamnochortus .....A.... ....C..CA. ...A...... ........A. .C..GG.A..
Zea           .....A..C. .......... ...A..G... ........A. .C..T..A..

Muscari       TGATGATGAT GATGAA---- -----CCTGG TACTGAT--- ----------
Aristida      C..A.....C ..G..GGACG GTGACA.... ..G...C... .........C
Avena         ?????????? ?????????? ?????????? ?????????? ??????????
Bambusa       ...G.....G ..G..GGATG GGAACA..CT GTG...C... .........C
Bouteloua     ...G.....G ..G..CGAGG CTGACAAC.. G.GC..C... .........C
Bromus        C..G..C..G .....GGATG GAGACA.... G.G...C... .........C
Calamovilfa   ...G.....G .....TGATG CAGAGA.CA. G.G...C... .........C
Danthonia     C..G.....G ..G...AATG GAGACA.... G.G...C... .........C
Eragrostis    ...G..C..G .....CGATG GGGATA.C.. G.G...C... .........C
Flagellaria   C..G..A.GC ..C..T.... ..GATA.C.. G.G...CCAG CAGCAGCAGC
Hordeum       C..G.....C .....GGATG GAGACA.... G.G...C... .........C
Joinvillea    ...G.....C .....T.... ..GATA.... G.G...C... .........C
Oryza         ...G.....G .....TGATG GAGACA.... G.G...C... .........C
Phragmites    C..G.....G ..G..GGACA GGGACA.C.. G.G...C... .........C
Panicum       ?????????? ?????????? ?????????? ?????????? ??????????
Poa           C..G....GG ..G..GGATG GAGAAA.C.. G.G...C... .........C
Pennisetum    C..A...... ..G..GGATG GGGATA.C.. G.G...C... .........C
Stipa         ...G.....G ..G..GGACG GGGACA.... G.G...C... .........C
```

```
Thamnochortus ...G...... .GC..T.... ..GATA.C.. G.A...C... .........C
Zea           C..C.....G ..G...GATG GGGATA.C.A C.G...C... .........C

Muscari       ---------- -CAGAAGGGG AGAAAGCTGT GGGGCTTGGT GGTTTGCCAT
Aristida      AGCAGCCA.. .---..A..A ..G....... ....GC.... ...G......
Avena         ?????????? ?????????? ?????????? ?????????? ??????????
Bambusa       AGCAGCCG.. .---..A..C ..G....... ....GC.... ...C......
Bouteloua     AGCAGCCG.. .---.G...C ..G....... ....GC.T.. ...C......
Bromus        AGCAGCCC.. .---..A..C ..G....... ....GC.... ..........
Calamovilfa   AGCAGCCA.. .---.....C ..G....... ....TC.T.. ...C......
Danthonia     AACAGCCG.. .---..A... ..G.....C. ....GC.... ...C......
Eragrostis    AGCAGCCA.. .---..A..C ..G....... ....GC.T.. ......T...
Flagellaria   AGCAGCAG.. ....CGA..C ..G.....C. ....GC.... C..G.....C
Hordeum       AGCAGCCC.. .---..A..C ..G.....C. ....GC.... ..........
Joinvillea    AGCAGCAG.. .---..A..C ..G.....T. ....A..... ...C......
Oryza         AGCAGCCG.. .---..A... ..G....... ....G...A. ...C......
Phragmites    AGCAGCCC.. .---..A..C ..G.....T. ....GC.... ...C......
Panicum       ?????????? ?????????? ?????????? ?????????? ??????????
Poa           AGCAGCCA.. .---..AA.C ..G.....A. ....GC.... ..........
Pennisetum    AGCAGCCC.. .---..A..C ..G....... ....GC.... ...C......
Stipa         AGCAGCCG.. .---..A..A ..G..A.... .....C.... ..........
Thamnochortus AGCAGCAG.. .---.....C CAG.....T. ....A..... T..C......
Zea           AACAACCG.. .---..A..C ..G....... ....GC.... C..C......

Muscari       CACACTAGTT CCAGGTTTGT CCCCTCGCGG TTCCTGTTCA AGCAGAATAG
Aristida      .....G..CC .G........ ...T?????? ?????????? ??????????
Avena         ?????????? ?????????? ?????????? ?????????? ??????????
Bambusa       .....A..CC .G.....C.. .........C .......... G......CC.
Bouteloua     .....A..CC .G.....C.. ...T.....C .........C G......CC.
Bromus        .....C..CC .G.....C.. ...T..A..C .........C G...A..CC.
Calamovilfa   ........CC .G........ ...T?????? ?????????? ??????????
Danthonia     .....G..C. .G........ A.....A..C .........C G......CC.
Eragrostis    .....A..CC .G.....C.. ......A..C .........C GT.....CC.
Flagellaria   .....G...C .G........ ....AGT... ..T..CA.T. ........C.
Hordeum       .....G..CC .G.....C.. ......A..C .........C G...A..CC.
Joinvillea    .....AT..C .G........ ...T.....T ..T..T...C G......CC.
Oryza         .....A..CC .G........ ...T..A..C .........C G......CC.
Phragmites    ........CC .A.....C.. ....?????? ?????????? ??????????
Panicum       ?????????? ?????????? ?????????? ?????????? ??????????
Poa           .....C..CC .G.....C.. .........C .........C G...A..CC.
Pennisetum    .....G..CC .G.....C.. T...?????? ?????????? ??????????
Stipa         ..T..G...C .G.....C.. ...T..A..C .........C G...A..CC.
Thamnochortus ..T.....CC .G........ T...?????? ?????????? ??????????
Zea           ..T..A..CC .G........ ....?????? ?????????? ??????????

Muscari       GGTGAGGATG ATCGCCGACT GCCACGCGGT GCCGGTGAGG GTGGTCCAAG
Aristida      ?????????? ?????????? ?????????? ?????????? ??????????
Avena         ?????????? ?????????? ?????????? ?????????? ??????????
Bambusa       C...C..... ..T..T..T. .......C.C .......... ..CA....G.
Bouteloua     A...C..... ..T.....T. .......CAC ...A...... ..CA....G.
Bromus        T...C..... ..T.....T. .......A.C T......... ..CA....G.
Calamovilfa   ?????????? ?????????? ?????????? ?????????? ??????????
```

```
Danthonia     T...C..... ..T..T..T. ....T..CAC TG.A..T..C ..CA.A..G.
Eragrostis    A...C..... ..T..T..T. ....T..CAC A......... ..CA....G.
Flagellaria   ....C.C... ........T. .......CCC .T.....C.C ..AA......
Hordeum       T...C..... ..T..T..T. .......A.C TG........ ..CA....G.
Joinvillea    C...C..... ..T.....T. .......A.C C???...C.C ..TA....G.
Oryza         T...C..... ..T..T..T. ....T..T.C .......... ..CA....G.
Phragmites    ?????????? ?????????? ?????????? ?????????? ??????????
Panicum       ?????????? ?????????? ?????????? ?????????? ??????????
Poa           T...C.C... ..T..T..T. .......A.C T......... ..CA....G.
Pennisetum    ?????????? ?????????? ?????????? ?????????? ??????????
Stipa         T..AC.A... ..T..T..T. ....T..C.C .........A ..CA....G.
Thamnochortus ?????????? ?????????? ?????????? ?????????? ??????????
Zea           ?????????? ?????????? ?????????? ?????????? ??????????

Muscari       ACGAGAGCTT GGCTCAGCCC CTGTGTCTCG TGGGTTCGAC CTTGCGAGCT
Aristida      ?????????? ?????????? ?????????? ?????????? ??????????
Avena         ?????????? ?????????? ?????????? ?????????? ??????????
Bambusa       ..CCC.CGC. .C.A.....G .....C..G. .T..G..C.. AC....T..G
Bouteloua     ..CCTG.TC. AT.G...T.A .....CT.G. .T.....T.. TC....C..A
Bromus        ..CCTGCAA. .C.A.....G .....CT.A. .T..G..... TC....CT.C
Calamovilfa   ?????????? ?????????? ?????????? ?????????? ??????????
Danthonia     .TCCTG.GC. .CAG...G.G T.....T.G. .C..C..C.. GC....T..C
Eragrostis    ..CCTG.TC. .T.G.....T ..C..CT.G. .T..C..T.. TC.......G
Flagellaria   .ACCCTCGC. .C.G.....G .......... .C..C..C.. GC.....T.G
Hordeum       ..CCTGCAA. .C.A...G.G .....CT.A. .T..G..... TC.T..CT.C
Joinvillea    ..CCCGCAC. CC.G.....G ..T.....G. ....C..... G?.?..T..A
Oryza         .TCCTGCAC. AA.A.....G .....CT.G. .T..G..C.. GC....TT.G
Phragmites    ?????????? ?????????? ?????????? ?????????? ??????????
Panicum       ?????????? ?????????? ?????????? ?????????? ??????????
Poa           ..CCTGCAA. .C.C.....T ..C..CT.G. .T..G..... TC....CT.C
Pennisetum    ?????????? ?????????? ?????????? ?????????? ??????????
Stipa         ..CCTGCAA. .CAG.....A .....CT.G. .T..G..... GC....CT.C
Thamnochortus ?????????? ?????????? ?????????? ?????????? ??????????
Zea           ?????????? ?????????? ?????????? ?????????? ??????????

Muscari       CCCCATGGCT GCCATGCTCA GTACATGGCG AACATGGGCT CGGTTGCTTC
Aristida      ?????????? ?????????? ?????????? ?????????? ??????????
Avena         ?????????? ?????????? ?????????? ?????????? ??????????
Bambusa       ..A..C..G. ....C..A.. .......... ........G. .CA....A..
Bouteloua     ..G..C.AG. ....C..A.. .........C ........G. ..A.......
Bromus        ..A.....G. .......G.. .........A ........G. .CA....A..
Calamovilfa   ?????????? ?????????? ?????????? ?????????? ??????????
Danthonia     ..A..C..G. .T..C..A.. .......... .......... .CA.C..A..
Eragrostis    ..G..C..G. .......A.. .........A ........G. ..A....A..
Flagellaria   ..G..C.CG. ....C..G.. .........C .......... .CA.C..G..
Hordeum       ..A.....G. .......G.. .........A ........G. ..A....A..
Joinvillea    ..G..C..A. .T.....G.. .........A .......... .CA....G..
Oryza         ..G.....T. ......GC.. ...T...... ........T. .CA....A..
Phragmites    ?????????? ?????????? ?????????? ?????????? ??????????
Panicum       ?????????? ?????????? ?????????? ?????????? ??????????
Poa           ..A..C..G. .......G.. .........A ........G. .CA....A..
Pennisetum    ?????????? ?????????? ?????????? ?????????? ??????????
```

```
Stipa          ..A..C..G. .......G.. .........A ........G. .CA....A..
Thamnochortus ?????????? ?????????? ?????????? ?????????? ??????????
Zea            ?????????? ?????????? ?????????? ?????????? ??????????

Muscari        GCTTGCAATG TCCGTGATTA TCAACGGGAG CGAAGAGGAC GGGACTAGGA
Aristida       ?????????? ?????????? ?????????? ?????????? ??????????
Avena          ?????????? ?????????? ?????????? ?????????? ??????????
Bambusa        ...G.TT... G.G.....C. .T.GTA.TG. T.GG..T... .A.CAC.AC.
Bouteloua      A....TT... G.G..A.... .T.GTA.TG. T.GT..T..T .A.CAA.CAT
Bromus         T..C.TC... G.A.....C. ...GTA.TG. T.GG.....T .A.CAC.AC.
Calamovilfa    ?????????? ?????????? ?????????? ?????????? ??????????
Danthonia      .....TT... G.G.....C. .T.GTA.TG. TAGT..T..T .A.CAA.T.G
Eragrostis     .....TT... G.G.....A. .T.G.A.TG. T.GT..T..T .A.CAA.CAT
Flagellaria    ...C.TC... G....C..C. .T.G.A.C.. .TCG..C... A.C.GC----
Hordeum        T..C.TC... G.A.....C. ...GTA.CG. T.GG.....T .A.CAC.AC.
Joinvillea     ...C.TC... G.G.....C. .T.G.CATG. ..CG..T..T .AC.A.GT.C
Oryza          T....TT... G.A.....C. .T.GTA.TG. T.GG..T..T .ATCA..AC.
Phragmites     ?????????? ?????????? ?????????? ?????????? ??????????
Panicum        ?????????? ?????????? ?????????? ?????????? ??????????
Poa            T..C.TC... G.A.....C. ...GTA.TG. T.GG.....T .A.CAC..C.
Pennisetum     ?????????? ?????????? ?????????? ?????????? ??????????
Stipa          T....TT... G.A.....C. ...GT..TG. T.GG..T... .A.CAC.AC.
Thamnochortus ?????????? ?????????? ?????????? ?????????? ??????????
Zea            ?????????? ?????????? ?????????? ?????????? ??????????

Muscari        AT-------- ---------- ---------- ---------- -ACGATGAAG
Aristida       ?????????? ?????????? ?????????? ?????????? ??????????
Avena          ?????????? ?????????? ?????????? ?????????? ??????????
Bambusa        TGACCCGGGG AGGCGTCCCG .......... ........TC GG.A......
Bouteloua      CAAGGAGCGG C...ATCTCG .......... ........TC GG.A......
Bromus         TGACGCGAGG CGTCATCCCG .......... ........TC GG........
Calamovilfa    ?????????? ?????????? ?????????? ?????????? ??????????
Danthonia      CACAGGGTGG C...ATCTCC .......... ........TC GG.......A
Eragrostis     CAAGAAGCGG T...ATCTCG .......... ........TC GG.A......
Flagellaria    --......AG CGCCGTCTCC .......... ........TC GGG...C...
Hordeum        TACGCGGGGC GTCAATCCCG .......... ........TC GG.A......
Joinvillea     CACGGACTGG C...ACCTCA .......... ........TC GG.A......
Oryza          T.GCACGGGG CAGCATCCCG .......... ........TC GG........
Phragmites     ?????????? ?????????? ?????????? ?????????? ??????????
Panicum        ?????????? ?????????? ?????????? ?????????? ??????????
Poa            TGGGGCGTGG CGCTGTCCCG .......... ........TC GG........
Pennisetum     ?????????? ?????????? ?????????? ?????????? ??????????
Stipa          TGGGACGGGG CGCCATCCCG .......... ........TC GG.A......
Thamnochortus ?????????? ?????????? ?????????? ?????????? ??????????
Zea            ?????????? ?????????? ?????????? ?????????? ??????????

Muscari        CTGTGGGGCC TCGTCGTGTG CCATCACACT TCGCAGAGGT GCATCCCTGC
Aristida       ?????????? ?????????? ?????????? ?????????? ????????..
Avena          ?????????? ?????????? ?????????? ?????????? ????????..
Bambusa        T.A.....AT .G..G..... ...C..T..A ..T.CCC... .......C??
Bouteloua      T.......GT .G..G..... ...C.....A ....CAC... TT.....C..
Bromus         ........G. .G..G..... ...C...... ..T.CAC... ........??
```

Calamovilfa	??????????	??????????	??????????	??????????	????????..
Danthonia	T.......GT	.G..G.....	...C..T..A	....CAC...	TT.....C..
Eragrostis	T.......AT	.G..G.....	...C.....A	....CAC...	TT........
Flagellaria	....	.......C..	...C.....G	....C.C...	.......C..
Hordeum	T.......GT	.A..G.....	...C.....A	..T.CAC...	.......C..
Joinvillea	........G.	.G..G.....	...C.....A	..A.CAC...	.......C??
Oryza	T.......GT	.G..A..A..	...C.....A	..T.CAC...	....
Phragmites	??????????	??????????	??????????	??????????	????????..
Panicum	??????????	??????????	??????????	??????????	????????..
Poa	T.A.....GT	.G..G.....	...C.....A	..T.CAC...	.......C..
Pennisetum	??????????	??????????	??????????	??????????	????????..
Stipa	T.......AT	.G..G.....	...C.....A	....C.C...	.......C..
Thamnochortus	??????????	??????????	??????????	??????????	??????????
Zea	??????????	??????????	??????????	??????????	????????..

Muscari	GCGGTTCCTC	TTCATGAAGA	GCAAGGTCCG	GATGATCTGC	GACTGCCGCG
Aristida	CA....T..T	....	A...A..A..	A.....T..T	..T..T..T.
Avena	CA..C.T..T	....	A...A..A..	......T..T	..T.....T.
Bambusa	??????????	??????????	??????????	??????????	??????????
Bouteloua	CA.A..T..T	....	A...A...A.	A.........	..T..T..T.
Bromus	??????????	??????????	??????????	??????????	??????????
Calamovilfa	CA....TG.T	....	A...A.....	A........T	..T..T....
Danthonia	CA....T..T	....	A...A.....	.........T	..T..T..T.
Eragrostis	AA....T..G	....	A...A.....	A........T	..T.....A.
Flagellaria	CA.A..T..T	....	A.........	.........T	........T.
Hordeum	TACC..T..G	....	A...A..G..	......T...	..TGT.....
Joinvillea	??????????	??????????	??????????	??????????	??????????
Oryza	CA....T..T	....	A...A.....	......T..T	..T.....T.
Phragmites	CA....T..T	....	A...A.....	A........T	..T..T..T.
Panicum	AA.......T	....	A...A..A..	A........T	..T..T..T.
Poa	AA....T..T	....	A...A..A..	......T..T	..T.....T.
Pennisetum	CA....T..T	..T.......	A...A..A..	A.........	..T..T..T.
Stipa	CA....T..T	........A.	A...A..T..	......T..T	..T.....T.
Thamnochortus	??????????	??????????	??????????	??????????	??????????
Zea	CA....T...	....	A...A...A.	A........T	..T.....T.

Muscari	CCAAGCCGGT	GAGGATTGTC	CAGGACGAGA	AGCTGCCCTT	CGACATCACC
Aristida	.G.GAT.C..	..A....A.T	G.A..T...G	CA..TT..A.	T..T..T.G.
Avena	.G.GAT.CA.	A.A.G.CA.T	G...CT...G	CA..C.....	T..T..T.G.
Bambusa	??????????	??????????	??????????	??????????	??????????
Bouteloua	.A.GAT.T..	A.A...AA.T	G.A..T...G	CAGCCT..A.	T.......G.
Bromus	??????????	??????????	??????????	??????????	??????????
Calamovilfa	.A.GAT.T..	A.A....A.T	G.A..T..AG	CAG.CT..A.	TA....T.G.
Danthonia	.A.GAT.C..	..A....A.T	G.A..T...G	CAA.CT..AC	T.....T.G.
Eragrostis	.A.GAT.C..	..A.G..A.T	A.A..T...G	CAGCCT.TA.	T..TG.T.G.
Flagellaria	.A..A..A..	..A....TA.	..A......T	C...T..T..	...T....G.
Hordeum	.A.GAA.CC.	A.A.G.CA..	GCC..T...G	CA..C.....	T..T..T.G.
Joinvillea	??????????	??????????	??????????	??????????	??????????
Oryza	.A.GAT.TA.	C.A....A..	G.A..T...T	C...C.A...	G..T..T.G.
Phragmites	.A.GAT.C..	..A....A.T	G.A..T...G	GA..CT..A.	T..T..T.G.
Panicum	.A.GAT.T..	..A....C.T	G.A..T..AG	CA..CT..A.	T..T..T.G.
Poa	.G.GAT.CA.	A.A.G.GA..	G.A.CT...G	CA..C.....	T..A..T.G.

```
Pennisetum    .A.GAT.T.. ..A....A.T G.A......G CA..CT..A. T..T..T.G.
Stipa         .A.GAT.CA. A.A...CA.T G.A..T.... CA..C..... T..T..T.G.
Thamnochortus ?????????? ?????????? ?????????? ?????????? ??????????
Zea           .A.GAT.C.. ..A....A.T G.A..T...G CA..CT..A. T..T..T.G.

Muscari       TTCTGCGGCT CCACTCTCCG GGCCCCCCAC AGCTGCCACC TCCAGTACAT
Aristida      ..A..T.... .A.....TA. A..A..A... ..T..T.... .T.....T..
Avena         C.A..T..T. .AG.A...A. ...A..A... ..T..T.... .T.....T..
Bambusa       ?????????? ?????????? ?????????? ?????????? ??????????
Bouteloua     ..G..T..T. .A.....GA. ...A..A..T ..T..T.... .T.....T..
Bromus        ?????????? ?????????? ?????????? ?????????? ??????????
Calamovilfa   ..G..T..T. .A.....GA. ...A..AA.T ..T..T..T. .T.....T..
Danthonia     ..G..T..T. .A.....TA. A..A..A..T .....T.... .T.....T..
Eragrostis    ..G..T..T. .A.....GA. ...A..G..T ..T..T.... .T.....T..
Flagellaria   ..A..T..T. .G..G...A. A..A..A..T ..T..T...T .G.....T..
Hordeum       ..G..T..T. .AT.A...A. ...AG.A... .....T.... .T.....T..
Joinvillea    ?????????? ?????????? ?????????? ?????????? ??????????
Oryza         ..A..T..T. .A..A..GA. ...A..A... ..T..T..T. .T.....T..
Phragmites    ..G..T..T. .A.....TA. A..G..A... ..T..T.... .T.....T..
Panicum       ..G..T..T. .A.....TA. A..A..A... ..T..T.... .T.....T..
Poa           ..G..T..T. .AG.A..AA. ...A..A... .AT..T.... .T.....T..
Pennisetum    ..G..T..T. .A.....TA. A..A..A... ..T..T.... .T.....T..
Stipa         ..G..T..T. .AG.A...A. ...A..A... ..T..T.... .T.....T..
Thamnochortus ?????????? ?????????? ?????????? ?????????? ??????????
Zea           ..G..T..T. .A.....TA. A..A..A..T .....T.... .TA....T..

Muscari       GGAGAACATG AACTCCATCG CCTCCCTCGT CATGGCCGTC GTCGTTAACG
Aristida      .......... ..G..G..T. .A..T..T.. ......T..T ..G.......
Avena         .......... .....G..T. .A.....T.. ......T..T ..G.....T.
Bambusa       ?????????? ?????????? ?????????? ???????..T ..T..G..T.
Bouteloua     ......T... .....G..T. .A.....T.. ......T..T ..A.....T.
Bromus        ?????????? ?????????? ?????????? ?????????? ??????????
Calamovilfa   ......T... ....TG..T. .A.....T.. G....TT..T ..A.....T.
Danthonia     .......... .....A..T. .A.....T.. ......T..T ..G.....T.
Eragrostis    ......T... .....T..T. .A.....T.. ......A..T ..A.....T.
Flagellaria   .......... .....A..T. .A.....T.. ......T..T ..G..C....
Hordeum       .......... .....G..T. .A.....T.. ......T..T ..G.....T.
Joinvillea    ?????????? ?????????? ?????????? ?????????? ??????????
Oryza         .......... .....G..T. .A.....T.. ......T..T ..G.....T.
Phragmites    .......... .....G..T. .A.....T.. ......T..T ..G.....T.
Panicum       .......... .....A..T. .A.....T.. ......T..T ..G..C..T.
Poa           .......... .....G..T. .A.....T.. ......T..T ..G.....T.
Pennisetum    .......... .....A..T. .A.....T.. ......T..T ..G..C..T.
Stipa         .......... .....A..T. .A.....T.. ......T..G ..G.....T.
Thamnochortus ?????????? ?????????? ?????????? ?????????? ??????????
Zea           .......... .....G..T. .A.....T.. ......T..T ..G..C..T.

Muscari       AGGAGGCCGA GGAGGCGGAG CGGGACCCCG AGGCGGAG-- ----CAATCC
Aristida      .AA.T.AA.. ..GC.AT..A GCC..G..T. GA------.. .......C.A
Avena         ..A.T.AA.. ...T.AT..A GCT..GT.T. .A------.. .......C.A
Bambusa       ..A.T.AA.. ...T.AT..A GTC..GG.T. .A------.. .......C.A
Bouteloua     ....T.AA.. ...T.AT..A GTC..G..T. .A------.. ......GC.A
```

```
Bromus        ?????????? ?????????? ?????????? ?????????? ??????????
Calamovilfa   ..A.T.AA.. ...T.AC..A GC...GA.T. .A------.. .......C.A
Danthonia     .AA.T.AA.. ...T.AT..A GCC..G..T. .A------.. ......GC.A
Eragrostis    ..A.T.AA.. ..GT.AT... .CC..G..T. .A------.. .......C.A
Flagellaria   ..AGA.GA.. ...A.AT... .C...G..T. G.------.. ......GC.G
Hordeum       ..AGT.AA.. ...T.AC... GCC..GT... .A------.. .......C.A
Joinvillea    ?????????? ?????????? ?????????? ?????????? ??????????
Oryza         ..A.T.AG.. T..T.AT..A GTT.GGG.T. .T------.. .......C.T
Phragmites    .AA.T.AA.. ...T.AT..A GCC..A..A. GA------.. .......C.A
Panicum       .AA.T.AA.. ...T.AT..A .CCA.T..T. .A------.. .......C.A
Poa           ..A.T.A... ...T.AT..A GTC..GT.T. .C------.. .......C.A
Pennisetum    .AA.T.AA.. A..T.AT..A .CC..G..T. ..------.. .......C.A
Stipa         ..A.T.AA.. ...T.AT..A GTC..GG.T. .A------.. .......C.A
Thamnochortus ?????????? ?????????? ?????????? ?????????? ??????????
Zea           .AA.T.AA.. ...T.AT..A .CC..G..T. .A------.. .......C.A

Muscari       ---CAGCAGC CGAAGAGAAA GCGTCTCTGG GGCCTCGTGG TGTGCCACCA
Aristida      CCA..A.... A.C...AG.. .AAA..G... ..T..AC.T. .T........
Avena         GCA....... A...A.AG.. .AAA..A... ......C.T. .T........
Bambusa       GCA..A.... A......G.. AAAG..A... .....TA.T. .T........
Bouteloua     TCA.....A. A.....AG.. .AAG..A... ..T.....T. .C........
Bromus        ?????????? ?????????? ?????????? ?????????? ??????????
Calamovilfa   CCA..A.... A...A.A..G AAAA..G... ..T...A.T. .T........
Danthonia     CCA..A.... A.....AG.. .A.G..G... ..T...A.T. .T........
Eragrostis    CCG..A.... AA....AG.. .AAG..G... ..T...A.T. .T........
Flagellaria   .......... A......G.. .A.AT.A... .....T..T. ........A.
Hordeum       GCA....... A.....AG.. .ATA..A... .....T..T. .T........
Joinvillea    ?????????? ?????????? ?????????? ?????????? ??????????
Oryza         GCA..A.... A......G.. .AAA..A... ..A...C.T. .T........
Phragmites    TCA..A.... T.....AG.. .AAA..G... ..T...A.T. .T........
Panicum       .......... A.C...AG.. .A.G..G... ..T...C.T. .T..T..T..
Poa           ACAG...... A.....AG.. .AAA..G... ......C... .T........
Pennisetum    CCA....... A.....AG.. .A.G..G... ..T...A.T. .T........
Stipa         GCA....... A......G.. AA.A..G... ......C.T. .T........
Thamnochortus ?????????? ?????????? ?????????? ?????????? ??????????
Zea           CCA..A.... A.....AG.. .A.G..G... ..T...A.T. .T........

Muscari       CGAGAGCCCG AGGTTCGTCC CC
Aristida      T........C ..A.AT.... .T
Avena         T........C ..A.AT.... .T
Bambusa       T........C ..A.AT..A. ..
Bouteloua     T........A ..A.AT.... ..
Bromus        ?????????? ?????????? ??
Calamovilfa   T.....T..A ....AT.... ..
Danthonia     T........C ..A.AT..G. ..
Eragrostis    T.....T..A ..A.AT..G. ..
Flagellaria   T.....T..A C.......T. ..
Hordeum       T........C ..A.AT.... ..
Joinvillea    ?????????? ?????????? ??
Oryza         T........C ..A.AT..T. .T
Phragmites    T........C ..A.AT..T. ..
Panicum       T........C ..A.AT.... ..
```

```
Poa           T..A.....C ..A.AT.... ..
Pennisetum    T........C ..A.AT.... .T
Stipa         T........C ....AT.... .A
Thamnochortus ?????????? ?????????? ??
Zea           T........C ..A.AT.... .T
```

Appendix G

Additive Binary Coding Matrices For Comparison Of Different Phylogenies Of The Grass Family (Poaceae)

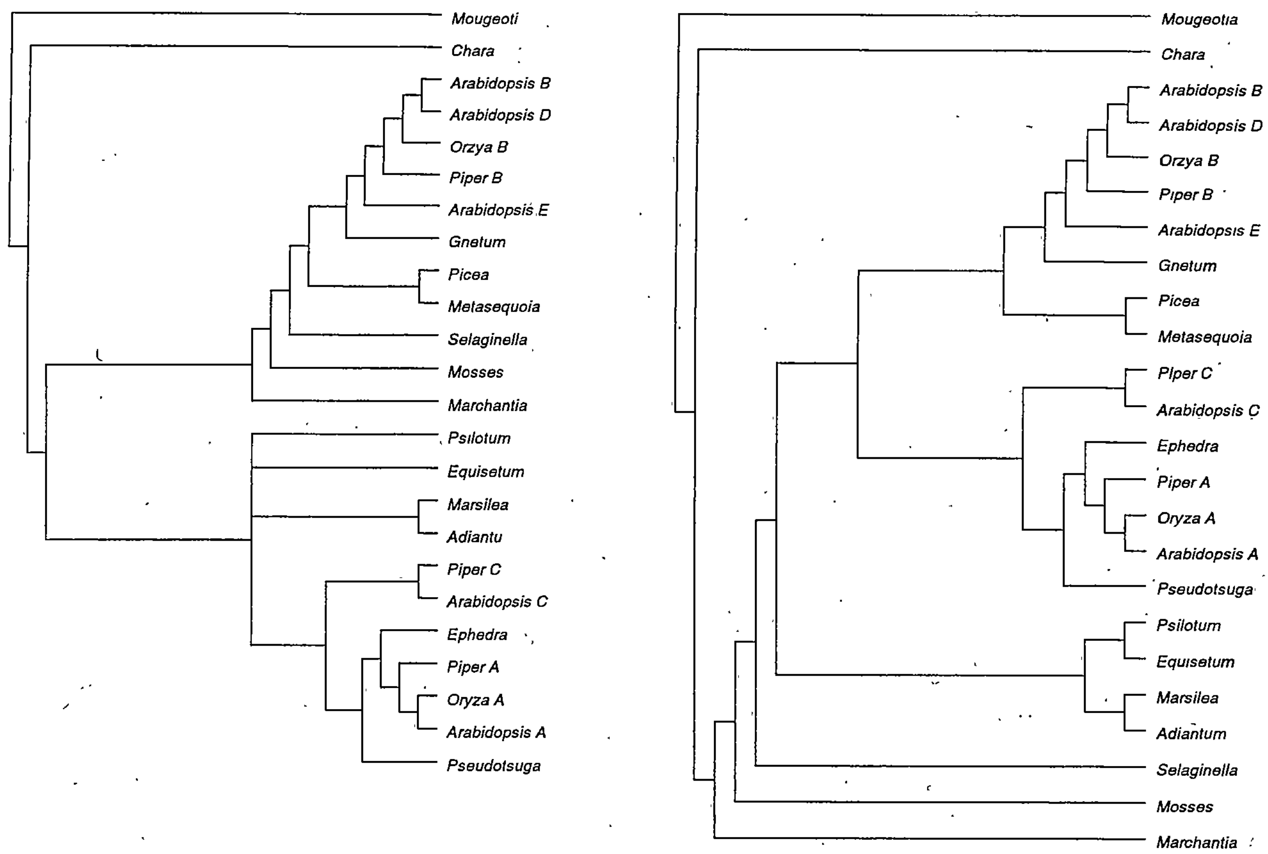
Labels: ja = *Joinvillea*, os = *Oryza*, bb = *Bambusa*, hv = *Hordeum*, bi = *Bromus*, as = *Avena*, pp = *Poa*, sv = *Stipa*, ec = *Eragrostis*, bg = *Bouteloua*, du = *Danthonia*, ap = *Aristida*, pa = *Phragmites*, zm = *Zea*, ps = *Pennisetum*, pc = *Panicum*

```
   I  Phytochrome tree                 II  rbcL tree
ja 1000000000000000000000000000000 100000000000000000000000000000
os 0100000000000000000001000000011 010000000000000001000000000001
bb 0010000000000000000011000000011 001000000000000001000000000001
hv 0001000000000000101111000000011 000100000000000001111000000001
bi 0000100000000000101111000000011 000010000000000001111000000001
as 0000010000000000011111000000011 000001000000000001110100000001
pp 0000001000000000011111000000011 000000100000000001110100000001
sv 0000000100000000000111000000011 000000010000000001100000000001
ec 0000000010000000000000100000111 000000001000000001000011110001
bg 0000000001000000000000100000111 000000000100000001000011110001
du 0000000000100000000000000011111 000000000010000001000011100001
ap 0000000000010000000000010001111 000000000001000001000011000001
pa 0000000000001000000000010001111 000000000000100001000010001001
zm 0000000000000100000000000111111 000000000000010001000010001101
ps 0000000000000010000000001111111 000000000000001001000010001111
pc 0000000000000001000000001111111 000000000000000101000010001111

   III  Morphology tree            IV  cpDNA tree
ja 100000000000000000000000000 1000000000000000000000000000
os 010000000000000010111000001 0100000000000000000010000110
bb 001000000000000010111000001 0010000000000000000010000110
hv 000100000000000011000000001 0001000000000000000000111100
bi 000010000000000011000000001 0000100000000000000000111100
as 000001000000000011000000001 0000010000000000000001011100
pp 000000100000000011000000001 0000001000000000000001011100
sv 000000010000000011000000001 0000000100000000000000001100
ec 000000001000000010110111001 0000000010000000011110000101
bg 000000000100000010110111001 ????????????????????????????
du 000000000010000010110100101 0000000000100000011110000101
ap 000000000001000010100000001 0000000000010000000110000100
pa 000000000000100010110110001 0000000000001000001110000101
zm 000000000000010010110100111 0000000000000100101110000100
ps 000000000000001010110100111 0000000000000010101110000100
pc 000000000000000110110100111 ????????????????????????????
```

Appendix H

Modified Phytochrome Trees Used In Tree Mapping Experiments



The phytochrome phylogenies that resulted from modification of Figure 10 to agree with branching order of taxa of the green plant phylogeny (Figure 12), but with most nonangiosperms retained in separate clades (left), or just gymnosperm taxa retained in separate clades (right).

Appendix I

Distance Matrices Used For Calculation Of Absolute And Relative Evolutionary Rates Of Phytochrome Sequences

Data: full length phytochrome sequences
Distance: Jukes-Cantor corrected proportion of nonsynonymous differences.
No. of Codons in Subset: 1131 of 1131
Genetic Code: "Universal"
Gap Sites and Missing Information Data: All such sites were removed only in the pairwise comparisons

OTU Labels
1.. Selaginella
2.. Adiantum
3.. Arabidopsis PHYB
4.. Arabidopsis PHYD
5.. Arabidopsis PHYE
6.. Solanum PHYB
7.. Oryza PHYB
8.. Arabidopsis PHYA
9.. Cucurbita PHYA
10.. Pisum PHYA
11.. Solanum PHYA
12.. Avena PHYA
13.. Oryza PHYA
14.. Zea PHYA
15.. Arabidopsis PHYC

Distances in the upper-right matrix
Standard Errors in lower-left matrix
'*' indicates an invalid distance value

OTUs	1	2	3	4	5	6	7
1		0.2235	0.2884	0.3016	0.3619	0.2574	0.2942
2	0.0107		0.3661	0.3772	0.4474	0.3381	0.3546
3	0.0126	0.0148		0.0944	0.3396	0.1433	0.1896
4	0.0129	0.0152	0.0064		0.3588	0.1698	0.2131
5	0.0148	0.0174	0.0141	0.0147		0.3326	0.3640
6	0.0117	0.0141	0.0081	0.0090	0.0139		0.1747
7	0.0127	0.0145	0.0096	0.0103	0.0148	0.0092	
8	0.0138	0.0151	0.0162	0.0161	0.0175	0.0156	0.0165
9	0.0138	0.0149	0.0158	0.0158	0.0175	0.0154	0.0162
10	0.0136	0.0145	0.0161	0.0160	0.0176	0.0155	0.0163
11	0.0142	0.0155	0.0164	0.0162	0.0180	0.0158	0.0167
12	0.0146	0.0158	0.0169	0.0173	0.0183	0.0164	0.0174
13	0.0146	0.0160	0.0167	0.0172	0.0183	0.0163	0.0173
14	0.0148	0.0163	0.0168	0.0175	0.0186	0.0164	0.0175
15	0.0150	0.0165	0.0159	0.0164	0.0184	0.0164	0.0171

OTUs	8	9	10	11	12	13	14	15
1	0.3332	0.3330	0.3259	0.3461	0.3582	0.3597	0.3669	0.3717
2	0.3744	0.3711	0.3547	0.3898	0.3999	0.4052	0.4163	0.4233
3	0.4157	0.4019	0.4121	0.4221	0.4377	0.4328	0.4362	0.4052
4	0.4131	0.4031	0.4083	0.4155	0.4516	0.4497	0.4581	0.4209
5	0.4536	0.4535	0.4559	0.4683	0.4771	0.4789	0.4868	0.4824
6	0.3931	0.3881	0.3902	0.3999	0.4190	0.4155	0.4190	0.4184
7	0.4259	0.4142	0.4170	0.4318	0.4531	0.4529	0.4572	0.4426
8		0.1291	0.1302	0.1375	0.2642	0.2724	0.2660	0.4118
9	0.0076		0.1414	0.1398	0.2566	0.2586	0.2605	0.4240
10	0.0077	0.0080		0.1361	0.2540	0.2538	0.2512	0.4198
11	0.0079	0.0080	0.0078		0.2572	0.2617	0.2637	0.4197
12	0.0118	0.0115	0.0115	0.0116		0.0543	0.0623	0.4441
13	0.0120	0.0116	0.0115	0.0117	0.0047		0.0610	0.4463
14	0.0118	0.0117	0.0114	0.0118	0.0051	0.0051		0.4408
15	0.0161	0.0164	0.0163	0.0163	0.0170	0.0171	0.0170	

Phytochrome nucleotide data corresponding to sites 1 to 2388 of Appendix A.
Distance: Jukes-Cantor corrected proportion of nonsynonymous differences.
No. of Codons in Subset: 796 of 796
Genetic Code: "Universal"
Gap Sites and Missing Information Data: All such sites were removed only in the pairwise comparisons

OTU Labels
1.. Ceratodon
2.. Selaginella
3.. Arabidopsis PHYB
4.. Arabidopsis PHYD
5.. Arabidopsis PHYE
6.. Solanum PHYB
7.. Oryza PHYB
8.. Arabidopsis PHYA
9.. Cucurbita PHYA
10.. Pisum PHYA
11.. Solanum PHYA
12.. Avena PHYA
13.. Oryza PHYA
14.. Zea PHYA
15.. Arabidopsis PHYC

Distances in the upper-right matrix
Standard Errors in lower-left matrix
'*' indicates an invalid distance value

OTUs	1	2	3	4	5	6	7
1		0.1842	0.3382	0.3395	0.3885	0.3096	0.3150
2	0.0113		0.2724	0.2814	0.3223	0.2378	0.2732
3	0.0167	0.0144		0.0839	0.2863	0.1276	0.1766
4	0.0168	0.0147	0.0072		0.3050	0.1482	0.1978
5	0.0187	0.0164	0.0150	0.0157		0.2768	0.3110
6	0.0158	0.0133	0.0091	0.0099	0.0147		0.1650
7	0.0160	0.0145	0.0110	0.0117	0.0159	0.0106	
8	0.0167	0.0154	0.0173	0.0172	0.0185	0.0167	0.0177
9	0.0171	0.0152	0.0173	0.0172	0.0183	0.0167	0.0177
10	0.0166	0.0148	0.0171	0.0169	0.0186	0.0165	0.0175
11	0.0174	0.0157	0.0176	0.0173	0.0192	0.0171	0.0183
12	0.0175	0.0156	0.0182	0.0186	0.0190	0.0175	0.0186
13	0.0172	0.0152	0.0177	0.0182	0.0189	0.0172	0.0181
14	0.0175	0.0155	0.0177	0.0182	0.0187	0.0172	0.0184
15	0.0179	0.0158	0.0174	0.0180	0.0197	0.0176	0.0187

OTUs	8	9	10	11	12	13	14	15
1	0.3365	0.3491	0.3354	0.3573	0.3599	0.3536	0.3597	0.3713
2	0.2996	0.2932	0.2840	0.3103	0.3071	0.2943	0.3033	0.3111
3	0.3565	0.3582	0.3516	0.3655	0.3810	0.3697	0.3683	0.3577
4	0.3537	0.3559	0.3461	0.3583	0.3929	0.3818	0.3840	0.3747
5	0.3848	0.3783	0.3868	0.4024	0.3967	0.3928	0.3884	0.4157
6	0.3378	0.3395	0.3316	0.3495	0.3609	0.3526	0.3516	0.3625
7	0.3671	0.3661	0.3614	0.3842	0.3901	0.3785	0.3847	0.3929
8		0.1094	0.0960	0.1188	0.2080	0.2087	0.2030	0.3542
9	0.0083		0.1024	0.1134	0.2034	0.2022	0.2054	0.3734
10	0.0077	0.0080		0.1144	0.2086	0.2065	0.2022	0.3584
11	0.0087	0.0084	0.0085		0.2185	0.2199	0.2189	0.3654
12	0.0121	0.0119	0.0121	0.0124		0.0385	0.0463	0.3705
13	0.0121	0.0118	0.0120	0.0125	0.0047		0.0447	0.3681
14	0.0119	0.0120	0.0118	0.0124	0.0052	0.0051		0.3653
15	0.0172	0.0179	0.0173	0.0176	0.0178	0.0177	0.0176	

Average distance: 0.2974 ñ 0.0076

Phytochrome nucleotide data corresponding to sites 298 to 491 of Appendix A.
Distance: Jukes-Cantor corrected proportion of nonsynonymous differences.
No. of Codons in Subset: 195 of 195
Genetic Code: "Universal"
Gap Sites and Missing Information Data: All such sites were removed only in the pairwise comparisons

OTU Labels
1.. Mougeotia
2.. Ceratodon
3.. Selaginella
4.. Arabidopsis PHYB
5.. Arabidopsis PHYD
6.. Arabidopsis PHYE
7.. Solanum PHYB
8.. Oryza PHYB
9.. Arabidopsis PHYA
10.. Cucurbita PHYA
11.. Pisum PHYA
12.. Solanum PHYA
13.. Avena PHYA
14.. Oryza PHYA
15.. Zea PHYA
16.. Arabidopsis PHYC

Distances in the upper-right matrix
Standard Errors in lower-left matrix
'*' indicates an invalid distance value

OTUs	1	2	3	4	5	6	7	8
1		0.1995	0.1835	0.2604	0.2753	0.2951	0.2601	0.2723
2	0.0243		0.1268	0.2623	0.2531	0.2724	0.2366	0.1868
3	0.0230	0.0185		0.2328	0.2242	0.2070	0.1908	0.2173
4	0.0284	0.0284	0.0264		0.0581	0.1866	0.0864	0.1558
5	0.0294	0.0278	0.0258	0.0119		0.2003	0.0949	0.1612
6	0.0314	0.0299	0.0253	0.0235	0.0245		0.1662	0.2283
7	0.0286	0.0267	0.0235	0.0148	0.0156	0.0220		0.1437
8	0.0293	0.0231	0.0253	0.0205	0.0209	0.0267	0.0198	
9	0.0319	0.0309	0.0294	0.0292	0.0301	0.0307	0.0284	0.0317
10	0.0316	0.0332	0.0300	0.0335	0.0344	0.0333	0.0306	0.0339
11	0.0294	0.0301	0.0274	0.0284	0.0303	0.0300	0.0282	0.0314
12	0.0299	0.0326	0.0293	0.0322	0.0330	0.0312	0.0306	0.0327
13	0.0343	0.0332	0.0316	0.0360	0.0361	0.0330	0.0350	0.0363
14	0.0331	0.0319	0.0295	0.0332	0.0336	0.0333	0.0331	0.0349
15	0.0339	0.0329	0.0316	0.0341	0.0342	0.0345	0.0331	0.0353
16	0.0278	0.0276	0.0242	0.0302	0.0296	0.0321	0.0318	0.0305

OTUs	9	10	11	12	13	14	15	16
1	0.3136	0.3088	0.2767	0.2832	0.3500	0.3334	0.3447	0.2510
2	0.2988	0.3332	0.2878	0.3228	0.3322	0.3138	0.3291	0.2508
3	0.2779	0.2866	0.2489	0.2755	0.3111	0.2812	0.3123	0.2018
4	0.2785	0.3426	0.2672	0.3217	0.3779	0.3386	0.3507	0.2916
5	0.2915	0.3549	0.2953	0.3339	0.3801	0.3437	0.3536	0.2826
6	0.2856	0.3228	0.2760	0.2923	0.3184	0.3225	0.3406	0.3053
7	0.2625	0.2968	0.2605	0.2940	0.3588	0.3328	0.3325	0.3116
8	0.3146	0.3455	0.3096	0.3275	0.3817	0.3618	0.3684	0.2939
9		0.0899	0.0728	0.0880	0.1983	0.1977	0.1930	0.3084
10	0.0149		0.0810	0.0836	0.1763	0.1686	0.1698	0.3340
11	0.0133	0.0141		0.0807	0.1903	0.1838	0.1794	0.3091
12	0.0147	0.0143	0.0140		0.1897	0.1875	0.1859	0.3214
13	0.0234	0.0218	0.0228	0.0228		0.0523	0.0593	0.3282
14	0.0233	0.0212	0.0223	0.0226	0.0110		0.0604	0.3091
15	0.0229	0.0213	0.0220	0.0225	0.0118	0.0118		0.3265
16	0.0314	0.0330	0.0314	0.0322	0.0327	0.0313	0.0325	

Phytochrome nucleotide data from grasses.
Distance: Jukes-Cantor corrected proportion of nonsynonymous differences.
No. of Codons in Subset: 108 of 108
Genetic Code: "Universal"
Gap Sites and Missing Information Data: All such sites were removed only in the pairwise comparisons

OTU Labels
1.. Selaginella
2.. Muscari PHYB
3.. Flagellaria PHYB
4.. Joinvillea PHYB
5.. Stipa PHYB
6.. Poa PHYB
7.. Oryza PHYB
8.. Hordeum PHYB
9.. Eragrostis PHYB
10.. Bouteloua PHYB
11.. Danthonia PHYB
12.. Bambusa PHYB
13.. Bromus PHYB
14.. Piper PHYC
15.. Muscari PHYC
16.. Aristida PHYC
17.. Bambusa PHYC
18.. Bromus PHYC
19.. Calamovilfa PHYC
20.. Danthonia PHYC
21.. Eragrostis PHYC
22.. Flagellaria PHYC
23.. Thamnochortus PHYC
24.. Hordeum PHYC
25.. Joinvillea PHYC
26.. Phragmites PHYC
27.. Pennisetum PHYC
28.. Poa PHYC
29.. Zea PHYC
30.. Bouteloua PHYC
31.. Oryza PHYC
32.. Stipa PHYC
33.. Muscari PHYA
34.. Zea PHYA
35.. Poa PHYA
36.. Panicum PHYU
37.. Phragmites PHYA
38.. Oryza PHYA
39.. Hordeum PHYA
40.. Flagellaria PHYA
41.. Danthonia PHYA
42.. Calamovilfa PHYA
43.. Bouteloua PHYA
44.. Eragrostis PHYA
45.. Panicum PHYA
46.. Avena PHYA
47.. Aristida PHYA
48.. Stipa PHYA
49.. Pennisetum PHYA

Distances in the upper-right matrix
Standard Errors in lower-left matrix
'*' indicates an invalid distance value

OTUs	1	2	3	4	5	6	7	8
1		0.2872	0.3067	0.2985	0.2993	0.2916	0.3067	0.3236
2	0.0426		0.1786	0.1623	0.1630	0.1569	0.1609	0.1717
3	0.0445	0.0319		0.1108	0.1423	0.1325	0.1433	0.1549
4	0.0435	0.0302	0.0241		0.0921	0.0970	0.0883	0.0972
5	0.0428	0.0298	0.0273	0.0213		0.0242	0.0470	0.0446
6	0.0420	0.0291	0.0262	0.0219	0.0103		0.0516	0.0399
7	0.0436	0.0296	0.0275	0.0209	0.0146	0.0153		0.0622
8	0.0452	0.0307	0.0287	0.0219	0.0142	0.0134	0.0169	
9	0.0439	0.0295	0.0289	0.0193	0.0186	0.0185	0.0185	0.0199
10	0.0452	0.0303	0.0290	0.0206	0.0201	0.0201	0.0201	0.0209
11	0.0451	0.0332	0.0288	0.0218	0.0209	0.0230	0.0220	0.0198
12	0.0433	0.0275	0.0258	0.0210	0.0136	0.0128	0.0135	0.0171
13	0.0444	0.0287	0.0268	0.0217	0.0109	0.0099	0.0139	0.0136
14	0.0363	0.0323	0.0438	0.0404	0.0391	0.0401	0.0403	0.0413
15	0.0316	0.0374	0.0429	0.0392	0.0401	0.0410	0.0410	0.0426
16	0.0311	0.0366	0.0440	0.0399	0.0400	0.0398	0.0395	0.0414
17	0.0343	0.0371	0.0458	0.0450	0.0444	0.0438	0.0432	0.0442
18	0.0344	0.0400	0.0453	0.0425	0.0432	0.0439	0.0427	0.0463
19	0.0364	0.0342	0.0448	0.0400	0.0428	0.0437	0.0420	0.0449
20	0.0330	0.0367	0.0428	0.0408	0.0413	0.0410	0.0408	0.0427
21	0.0350	0.0359	0.0448	0.0410	0.0412	0.0431	0.0407	0.0442
22	0.0368	0.0360	0.0446	0.0408	0.0407	0.0412	0.0414	0.0433
23	0.0371	0.0361	0.0445	0.0417	0.0428	0.0436	0.0426	0.0455
24	0.0340	0.0393	0.0445	0.0412	0.0423	0.0430	0.0421	0.0453
25	0.0325	0.0335	0.0402	0.0382	0.0387	0.0394	0.0385	0.0417
26	0.0327	0.0350	0.0432	0.0397	0.0401	0.0398	0.0396	0.0415
27	0.0329	0.0358	0.0435	0.0398	0.0400	0.0398	0.0395	0.0414
28	0.0346	0.0383	0.0464	0.0422	0.0432	0.0431	0.0427	0.0449
29	0.0346	0.0377	0.0434	0.0414	0.0430	0.0424	0.0415	0.0438
30	0.0359	0.0349	0.0471	0.0442	0.0424	0.0426	0.0428	0.0442
31	0.0348	0.0396	0.0476	0.0436	0.0425	0.0432	0.0417	0.0449
32	0.0315	0.0372	0.0456	0.0423	0.0432	0.0426	0.0426	0.0449
33	0.0460	0.0482	0.0596	0.0560	0.0588	0.0562	0.0575	0.0556
34	0.0454	0.0509	0.0568	0.0551	0.0576	0.0567	0.0541	0.0545
35	0.0463	0.0524	0.0554	0.0541	0.0582	0.0566	0.0556	0.0557
36	0.0504	0.0542	0.0563	0.0552	0.0567	0.0560	0.0551	0.0570
37	0.0437	0.0492	0.0552	0.0535	0.0565	0.0545	0.0525	0.0529
38	0.0426	0.0511	0.0522	0.0522	0.0540	0.0543	0.0530	0.0555
39	0.0504	0.0547	0.0604	0.0575	0.0600	0.0583	0.0575	0.0575
40	0.0440	0.0505	0.0555	0.0550	0.0557	0.0536	0.0537	0.0535
41	0.0445	0.0514	0.0566	0.0545	0.0567	0.0550	0.0543	0.0536
42	0.0467	0.0518	0.0586	0.0578	0.0603	0.0584	0.0573	0.0581
43	0.0454	0.0494	0.0565	0.0513	0.0545	0.0536	0.0523	0.0515
44	0.0471	0.0494	0.0568	0.0559	0.0572	0.0559	0.0546	0.0541
45	0.0456	0.0490	0.0565	0.0553	0.0575	0.0558	0.0529	0.0537
46	0.0473	0.0513	0.0553	0.0553	0.0578	0.0562	0.0553	0.0553
47	0.0439	0.0512	0.0560	0.0538	0.0548	0.0528	0.0527	0.0518
48	0.0435	0.0520	0.0554	0.0553	0.0578	0.0561	0.0554	0.0562
49	0.0442	0.0503	0.0553	0.0544	0.0570	0.0553	0.0534	0.0538

OTUs	9	10	11	12	13	14	15	16
1	0.3073	0.3216	0.3213	0.3037	0.3148	0.2229	0.1829	0.1802
2	0.1590	0.1664	0.1957	0.1419	0.1521	0.1830	0.2326	0.2290
3	0.1553	0.1565	0.1552	0.1279	0.1371	0.2956	0.2872	0.3019
4	0.0765	0.0864	0.0968	0.0891	0.0949	0.2617	0.2515	0.2652
5	0.0729	0.0849	0.0917	0.0410	0.0265	0.2559	0.2657	0.2711
6	0.0727	0.0847	0.1089	0.0363	0.0220	0.2667	0.2756	0.2691
7	0.0723	0.0844	0.1003	0.0403	0.0425	0.2684	0.2746	0.2650
8	0.0834	0.0906	0.0827	0.0634	0.0409	0.2786	0.2918	0.2858
9		0.0088	0.0543	0.0715	0.0682	0.2218	0.2429	0.2483
10	0.0063		0.0543	0.0763	0.0803	0.2381	0.2467	0.2582
11	0.0158	0.0158		0.0945	0.0969	0.2728	0.2766	0.2998
12	0.0184	0.0190	0.0213		0.0266	0.2613	0.2570	0.2689

13	0.0179	0.0196	0.0216	0.0109		0.2598	0.2841	0.2784
14	0.0360	0.0376	0.0409	0.0397	0.0395		0.1385	0.1421
15	0.0379	0.0382	0.0411	0.0393	0.0419	0.0274		0.0816
16	0.0379	0.0388	0.0427	0.0398	0.0408	0.0277	0.0200	
17	0.0410	0.0421	0.0453	0.0417	0.0434	0.0329	0.0228	0.0157
18	0.0395	0.0410	0.0463	0.0418	0.0452	0.0324	0.0221	0.0170
19	0.0377	0.0386	0.0438	0.0402	0.0429	0.0299	0.0240	0.0139
20	0.0386	0.0396	0.0440	0.0400	0.0413	0.0286	0.0195	0.0106
21	0.0386	0.0395	0.0442	0.0399	0.0428	0.0298	0.0224	0.0110
22	0.0387	0.0403	0.0440	0.0391	0.0416	0.0275	0.0200	0.0178
23	0.0394	0.0410	0.0448	0.0417	0.0435	0.0241	0.0218	0.0177
24	0.0389	0.0407	0.0454	0.0412	0.0436	0.0291	0.0206	0.0150
25	0.0360	0.0375	0.0415	0.0370	0.0397	0.0253	0.0184	0.0160
26	0.0366	0.0375	0.0425	0.0394	0.0401	0.0292	0.0195	0.0101
27	0.0394	0.0405	0.0445	0.0393	0.0407	0.0305	0.0206	0.0115
28	0.0383	0.0398	0.0451	0.0418	0.0439	0.0307	0.0232	0.0169
29	0.0396	0.0412	0.0444	0.0404	0.0427	0.0318	0.0233	0.0139
30	0.0409	0.0417	0.0459	0.0411	0.0428	0.0298	0.0254	0.0150
31	0.0390	0.0408	0.0446	0.0419	0.0442	0.0327	0.0238	0.0156
32	0.0394	0.0410	0.0462	0.0418	0.0438	0.0306	0.0214	0.0149
33	0.0547	0.0560	0.0547	0.0550	0.0561	0.0445	0.0419	0.0406
34	0.0530	0.0549	0.0560	0.0550	0.0556	0.0455	0.0440	0.0385
35	0.0551	0.0550	0.0544	0.0555	0.0561	0.0450	0.0421	0.0405
36	0.0537	0.0554	0.0593	0.0525	0.0547	0.0499	0.0450	0.0411
37	0.0496	0.0515	0.0524	0.0539	0.0545	0.0435	0.0434	0.0362
38	0.0531	0.0551	0.0522	0.0530	0.0559	0.0445	0.0392	0.0377
39	0.0571	0.0592	0.0586	0.0573	0.0579	0.0476	0.0462	0.0436
40	0.0511	0.0529	0.0533	0.0525	0.0539	0.0440	0.0427	0.0376
41	0.0535	0.0551	0.0566	0.0556	0.0547	0.0460	0.0439	0.0380
42	0.0549	0.0569	0.0602	0.0575	0.0573	0.0467	0.0448	0.0396
43	0.0516	0.0535	0.0531	0.0528	0.0526	0.0443	0.0413	0.0370
44	0.0538	0.0554	0.0565	0.0555	0.0552	0.0449	0.0422	0.0396
45	0.0530	0.0550	0.0570	0.0549	0.0555	0.0430	0.0442	0.0380
46	0.0549	0.0570	0.0565	0.0551	0.0558	0.0460	0.0434	0.0399
47	0.0501	0.0520	0.0539	0.0523	0.0537	0.0445	0.0433	0.0372
48	0.0566	0.0586	0.0567	0.0545	0.0566	0.0457	0.0416	0.0387
49	0.0523	0.0543	0.0553	0.0544	0.0550	0.0449	0.0435	0.0380

OTUs	17	18	19	20	21	22	23	24
1	0.2118	0.2138	0.2330	0.1987	0.2188	0.2363	0.2364	0.2091
2	0.2347	0.2648	0.2058	0.2303	0.2227	0.2195	0.2209	0.2574
3	0.3210	0.3164	0.3108	0.2912	0.3113	0.3093	0.3025	0.3080
4	0.3172	0.2918	0.2662	0.2740	0.2768	0.2721	0.2785	0.2788
5	0.3166	0.3052	0.3012	0.2852	0.2847	0.2794	0.2963	0.2956
6	0.3110	0.3129	0.3111	0.2831	0.3053	0.2850	0.3052	0.3033
7	0.3035	0.2989	0.2916	0.2790	0.2785	0.2866	0.2932	0.2926
8	0.3152	0.3375	0.3232	0.3001	0.3161	0.3073	0.3243	0.3276
9	0.2807	0.2656	0.2473	0.2568	0.2564	0.2565	0.2605	0.2596
10	0.2931	0.2821	0.2572	0.2668	0.2663	0.2729	0.2772	0.2792
11	0.3274	0.3384	0.3118	0.3143	0.3171	0.3122	0.3173	0.3296
12	0.2879	0.2899	0.2731	0.2713	0.2708	0.2627	0.2840	0.2837
13	0.3065	0.3252	0.3011	0.2850	0.3008	0.2888	0.3028	0.3087
14	0.1903	0.1866	0.1629	0.1501	0.1613	0.1461	0.1113	0.1556
15	0.1037	0.0981	0.1137	0.0781	0.1007	0.0826	0.0957	0.0863
16	0.0550	0.0640	0.0436	0.0259	0.0280	0.0682	0.0677	0.0503
17		0.0616	0.0779	0.0549	0.0662	0.1147	0.1189	0.0756
18	0.0166		0.0847	0.0547	0.0668	0.0943	0.0888	0.0301
19	0.0189	0.0197		0.0502	0.0214	0.0799	0.0913	0.0777
20	0.0157	0.0156	0.0149		0.0368	0.0727	0.0770	0.0593
21	0.0173	0.0173	0.0096	0.0127		0.0702	0.0816	0.0592
22	0.0237	0.0212	0.0194	0.0184	0.0181		0.0584	0.0774
23	0.0241	0.0204	0.0208	0.0189	0.0195	0.0164		0.0769
24	0.0186	0.0114	0.0188	0.0163	0.0163	0.0190	0.0189	
25	0.0211	0.0192	0.0186	0.0156	0.0173	0.0130	0.0144	0.0179

26	0.0150	0.0149	0.0142	0.0096	0.0119	0.0188	0.0199	0.0156
27	0.0171	0.0183	0.0164	0.0131	0.0144	0.0207	0.0212	0.0171
28	0.0166	0.0105	0.0197	0.0159	0.0184	0.0205	0.0218	0.0126
29	0.0173	0.0198	0.0188	0.0142	0.0160	0.0230	0.0233	0.0192
30	0.0198	0.0209	0.0138	0.0163	0.0131	0.0197	0.0222	0.0200
31	0.0166	0.0172	0.0169	0.0149	0.0141	0.0208	0.0230	0.0178
32	0.0146	0.0096	0.0191	0.0149	0.0166	0.0215	0.0216	0.0123
33	0.0418	0.0439	0.0410	0.0413	0.0416	0.0404	0.0422	0.0415
34	0.0410	0.0424	0.0407	0.0396	0.0395	0.0419	0.0412	0.0409
35	0.0444	0.0447	0.0440	0.0427	0.0426	0.0414	0.0403	0.0426
36	0.0426	0.0455	0.0409	0.0424	0.0408	0.0421	0.0438	0.0433
37	0.0396	0.0406	0.0381	0.0374	0.0372	0.0401	0.0399	0.0392
38	0.0409	0.0394	0.0383	0.0381	0.0367	0.0375	0.0376	0.0380
39	0.0457	0.0476	0.0460	0.0447	0.0445	0.0456	0.0455	0.0461
40	0.0393	0.0414	0.0386	0.0390	0.0385	0.0389	0.0423	0.0402
41	0.0409	0.0416	0.0399	0.0389	0.0388	0.0414	0.0404	0.0405
42	0.0435	0.0434	0.0407	0.0407	0.0408	0.0426	0.0431	0.0420
43	0.0401	0.0406	0.0392	0.0382	0.0380	0.0401	0.0395	0.0391
44	0.0409	0.0432	0.0421	0.0407	0.0408	0.0414	0.0407	0.0417
45	0.0405	0.0419	0.0390	0.0391	0.0390	0.0402	0.0403	0.0404
46	0.0427	0.0431	0.0422	0.0409	0.0408	0.0414	0.0407	0.0423
47	0.0401	0.0411	0.0389	0.0384	0.0377	0.0403	0.0410	0.0396
48	0.0416	0.0418	0.0401	0.0384	0.0382	0.0399	0.0403	0.0401
49	0.0405	0.0419	0.0402	0.0391	0.0390	0.0414	0.0406	0.0404

OTUs	25	26	27	28	29	30	31	32
1	0.1910	0.1959	0.1975	0.2159	0.2145	0.2280	0.2177	0.1849
2	0.1959	0.2140	0.2209	0.2476	0.2404	0.2125	0.2613	0.2359
3	0.2597	0.2948	0.2976	0.3290	0.2966	0.3347	0.3414	0.3198
4	0.2436	0.2634	0.2632	0.2893	0.2808	0.3093	0.3051	0.2894
5	0.2546	0.2725	0.2714	0.3060	0.3029	0.2968	0.2993	0.3051
6	0.2621	0.2705	0.2694	0.3060	0.2974	0.2990	0.3070	0.2996
7	0.2515	0.2663	0.2652	0.2997	0.2867	0.3004	0.2898	0.2987
8	0.2855	0.2872	0.2861	0.3246	0.3114	0.3153	0.3245	0.3237
9	0.2257	0.2354	0.2641	0.2541	0.2665	0.2803	0.2612	0.2655
10	0.2417	0.2452	0.2753	0.2703	0.2830	0.2895	0.2807	0.2820
11	0.2833	0.2979	0.3190	0.3267	0.3180	0.3343	0.3220	0.3383
12	0.2367	0.2651	0.2628	0.2908	0.2747	0.2827	0.2917	0.2898
13	0.2639	0.2723	0.2776	0.3125	0.2994	0.3000	0.3157	0.3116
14	0.1220	0.1560	0.1670	0.1699	0.1802	0.1613	0.1896	0.1689
15	0.0705	0.0782	0.0859	0.1080	0.1078	0.1260	0.1131	0.0930
16	0.0559	0.0237	0.0304	0.0637	0.0436	0.0503	0.0546	0.0503
17	0.0939	0.0503	0.0642	0.0614	0.0664	0.0850	0.0614	0.0480
18	0.0791	0.0502	0.0733	0.0256	0.0848	0.0942	0.0657	0.0213
19	0.0744	0.0457	0.0595	0.0844	0.0771	0.0434	0.0635	0.0800
20	0.0534	0.0214	0.0392	0.0567	0.0458	0.0593	0.0499	0.0501
21	0.0649	0.0323	0.0466	0.0750	0.0570	0.0389	0.0454	0.0614
22	0.0376	0.0752	0.0899	0.0891	0.1095	0.0823	0.0915	0.0968
23	0.0464	0.0842	0.0942	0.1006	0.1120	0.1028	0.1104	0.0985
24	0.0697	0.0547	0.0648	0.0366	0.0802	0.0871	0.0704	0.0345
25		0.0604	0.0676	0.0835	0.0889	0.0888	0.0788	0.0814
26	0.0166		0.0347	0.0500	0.0413	0.0563	0.0410	0.0367
27	0.0177	0.0123		0.0707	0.0504	0.0687	0.0645	0.0594
28	0.0197	0.0149	0.0179		0.0845	0.0938	0.0770	0.0234
29	0.0205	0.0135	0.0150	0.0197		0.0802	0.0682	0.0715
30	0.0204	0.0159	0.0177	0.0208	0.0192		0.0735	0.0800
31	0.0191	0.0134	0.0170	0.0187	0.0175	0.0182		0.0611
32	0.0195	0.0127	0.0163	0.0100	0.0180	0.0191	0.0165	
33	0.0425	0.0422	0.0424	0.0407	0.0434	0.0410	0.0443	0.0420
34	0.0429	0.0395	0.0408	0.0406	0.0424	0.0401	0.0403	0.0405
35	0.0425	0.0432	0.0427	0.0431	0.0459	0.0432	0.0439	0.0433
36	0.0444	0.0414	0.0426	0.0437	0.0431	0.0408	0.0426	0.0425
37	0.0404	0.0372	0.0396	0.0389	0.0404	0.0378	0.0380	0.0388
38	0.0387	0.0383	0.0390	0.0390	0.0410	0.0390	0.0381	0.0388