Plant Photoreceptors: Phylogenetic Overview

Lariguet, Patricia; Dunand, Christophe

How to cite

LARIGUET, Patricia, DUNAND, Christophe. Plant Photoreceptors: Phylogenetic Overview. In: Journal of Molecular Evolution, 2005, vol. 61, n° 4, p. 559–569. doi: 10.1007/s00239-004-0294-2

This publication URL: https://archive-ouverte.unige.ch/unige:118845
Publication DOI: 10.1007/s00239-004-0294-2

J Mol Evol (2005) 61:559–569
DOI: 10.1007/s00239-004-0294-2

JOURNAL OF MOLECULAR EVOLUTION

Plant Photoreceptors: Phylogenetic Overview

Patricia Lariguet, Christophe Dunand

Department of Plant Biology, University of Geneva, Quai Ernest Ansermet 30, CH-1211 Geneva 4, Switzerland

Received: 5 October 2004 / Accepted: 4 June 2005 [*Reviewing Editor:* Dr. Rüdiger Cerff]

Abstract. Plants possess photoreceptors to perceive light which controls most aspects of their lives. Three photoreceptor families are well characterized: cryptochromes (crys), phototropins (phots), and phytochromes (phys). Two putative families have been identified more recently: Zeitlupes (ZTLs) and UV-B photoreceptors (ULI). Using *Arabidopsis thaliana* and *Oryza sativa* photoreceptor sequences as references, we have searched for photoreceptor encoding genes in the major phyla of plant kingdom. For each photoreceptor family, using a phylogenetic tree based on the alignment of conserved amino acid sequences, we have tried to trace back the evolution and the emergence of the diverse photoreceptor ancestral sequences.The green alga *Chlamydomonas* contains one cry and one phot sequence, probably close to the corresponding ancestral sequences, and no phy-related sequence. The putative UV-B photoreceptors seem to be restricted to the Brassicacae. Except for mosses and ferns, which contain divergent photoreceptor numbers, the composition of the diverse photoreceptor families is conserved between species. A high conservation of the residues within domains is observed in each photoreceptor family. The complete phylogenic analysis of the photoreceptor families in plants has confirmed the existence of crucial evolutionary nodes between the major phyla. For each photoreceptor class, a major duplication occurred before the separation between Mono- and Eudicotyledons. This allowed postulating on the putative ancestral function of the photoreceptors.

Correspondence to: Christophe Dunand; email: christophe.dunand@bota.unige.ch

Key words: Cryptochromes — Phototropins — Phylogeny — Phytochromes — UV-B photoreceptors — Zeitlupes

Introduction

Because plants are sessile in nature, their survival and reproduction depend on their ability to adapt their development in response to environmental changes such as light, water availability, temperature, and pathogens. In particular, land plants need to deal with various light environments depending on the degree of shading by soil, foliage, or clouds, time of day, and time of year. Plants are able to perceive mainly the UV, blue (B), green, red (R), and far-red (FR) wavebands through different photoreceptor families (Fig. 1): cryptochromes (crys), phototropins (phots), phytochromes (phys), zeitlupes (ZLPs), and a putative UV-B photoreceptor (Gyula et al. 2003; Somers et al. 2000; Suesslin and Frohnmeyer 2003).

Crys flavoproteins are receptors perceiving B and UV-A radiations (320–520 nm). They consist of three members, cry1, cry2 (Lin 2000), and CryDASH (cry3 in *Arabidopsis*) (Brudler et al. 2003). Crys flavoproteins are responsible for photomorphogenesis, flowering, and clock resetting in plants (Lin and Shalitin 2003). They have sequence homologies with light-dependent DNA repair photolyases without DNA photolyase activity (Sancar 2003). Photolyases and crys are characterized by a flavin adenine dinucleotide (FAD) chromophore and contain either deazaflavin or pterin as light-harvesting chromophore

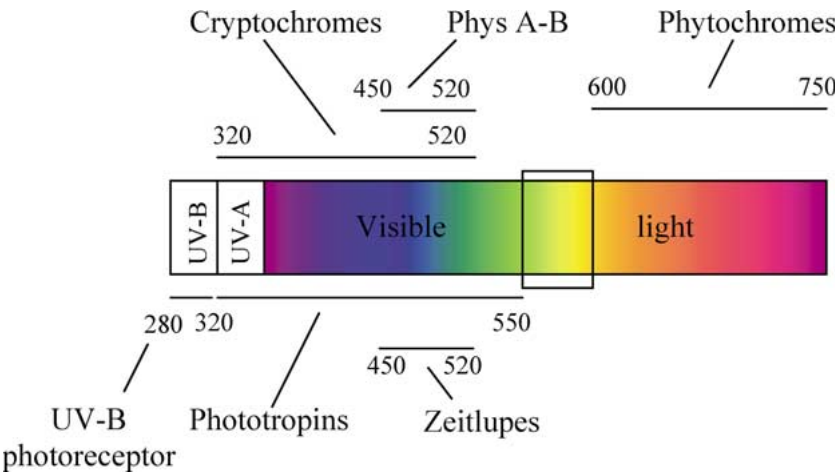


Fig. 1. The plant photoreceptor absorption zones. The wavelength ranges perceived by the different known and putative *Arabidopsis* photoreceptors have been are visualized. Note that 550–660 nm (yellow) does not seem to be sensed by any known photoreceptor, while 450–520 nm is absorbed by crys, phots, phys, and ZTL. Modified from Sullivan and Deng (2003).

(Sancar 1994) (Fig. 2A). In addition, plant crys flavoproteins contain an extra C-Term extension, which can interact with COP1, a E3 ubiquitin ligase (Wang et al. 2001; Yang et al. 2001). This extension is not present in the CryDASH family. Crys homologues are apparently present in several phyla including animals, yeasts, plants, and bacteria (Daiyasu et al. 2004; Sancar 2003). In animals they function as circadian clock components, possibly within the clock oscillator in Vertebrates and in the clock input in Drosophila (Panda et al. 2002). Phylogenetic analysis of the crys/photolyase family revealed five evolution branches (Lin 2000b; Brudler et al 2003). However, those studies included only a few plants, leading to limited knowledge concerning the evolution of crys in plants.

The phots family is composed in *Arabidopsis* of two members, namely, phot1 and phot2 (Briggs and Christie 2002), and perceives UV-A, UV-B, and green lights (320–550 nm). Their main functions are phototropism, chloroplast relocation, and stomata opening (Briggs and Christie 2002; Kinoshita et al. 2001). Phots are flavoproteins containing two LOV domains at the N-terminal and a serine/threonine kinase domain at the C-terminal part (Fig. 2B) (Huala et al. 1997; Sakai et al. 2001). They represent a class of receptor kinases that appear to be exclusive to plants, although no phylogenetic analysis is available to our knowledge (Briggs and Christie 2002).

Arabidopsis has five distinct phys, phyA- to -E, which are broad-range sensing photoreceptors with maximum absorption for R and FR lights (600–750 nm). PhyA and phyB can also act as B light photoreceptors (Lagarias et al. 1997; Lariguet and Fankhauser 2004b; Moller et al. 2002). They have different but overlapping functions in photomorphogenesis, shade avoidance, clock resetting, and gravitropism inhibition (Lariguet and Fankhauser 2004a; Moller et al. 2002). The phytochrome apoproteins contain a chromophore binding site in the N-terminal part and two histidine kinase-related domains (HKRD1 and -2). The HKRD1 domain contains two PAS domains with a putative function for protein–protein interaction (Fig. 2C). Several phylogenetic analyses were performed concerning phys (Alba et al. 2000a; Lamparter 2004; Mathews et al. 2003; Mathews and Sharrock; Montgomery and Lagarias 2002). They have shown that phy-like genes are present not only in all green plants but also in certain bacteria and fungi (Lamparter 2004; Mathews and Sharrock 1997). The plant phy family results from several gene duplications, with the first one preceding the origin of seed plants (Mathews and Sharrock 1997). An exhaustive phylogenetic analysis was performed in Gymnosperms (Schmidt and Schneider-Poetsch 2002). But the use of very small phys Gymnosperm sequences did not allow a global analysis with the phy sequences of other phyla. No global phylogenetic analyses including Gymnosperm, Angiosperm, and older phy sequences such as Bryophytes and Pteridophytes have been reported. This global approach would allow visualizing the hot spot of the phy sequences evolution.

The ZTL family consists of three members in *Arabidopsis*—ZTL; flavin-binding, kelch repeat, F-box1 (FKF1); and LOV kelch protein 2 (LKP2)—and is predominantly involved in the control of the circadian period in plants (Imaizumi et al. 2003; Mas et al. 2003; Schultz et al. 2001; Somers et al. 2000). They possess a PAS-like LOV domain, an F-box domain, and six kelch repeats (Fig. 2D) (Somers et al. 2000). The PAS region is highly similar to the LOV domain of phot1, the F-box is involved in substrate recognition for degradation, and the kelch repeats are implicated in protein–protein interactions. Therefore the ZTL proteins could be involved in the degradation of circadian clock components (Han et al. 2004). The similarity of the ZTL PAS domain to the phots LOV domain and the strong light depen-

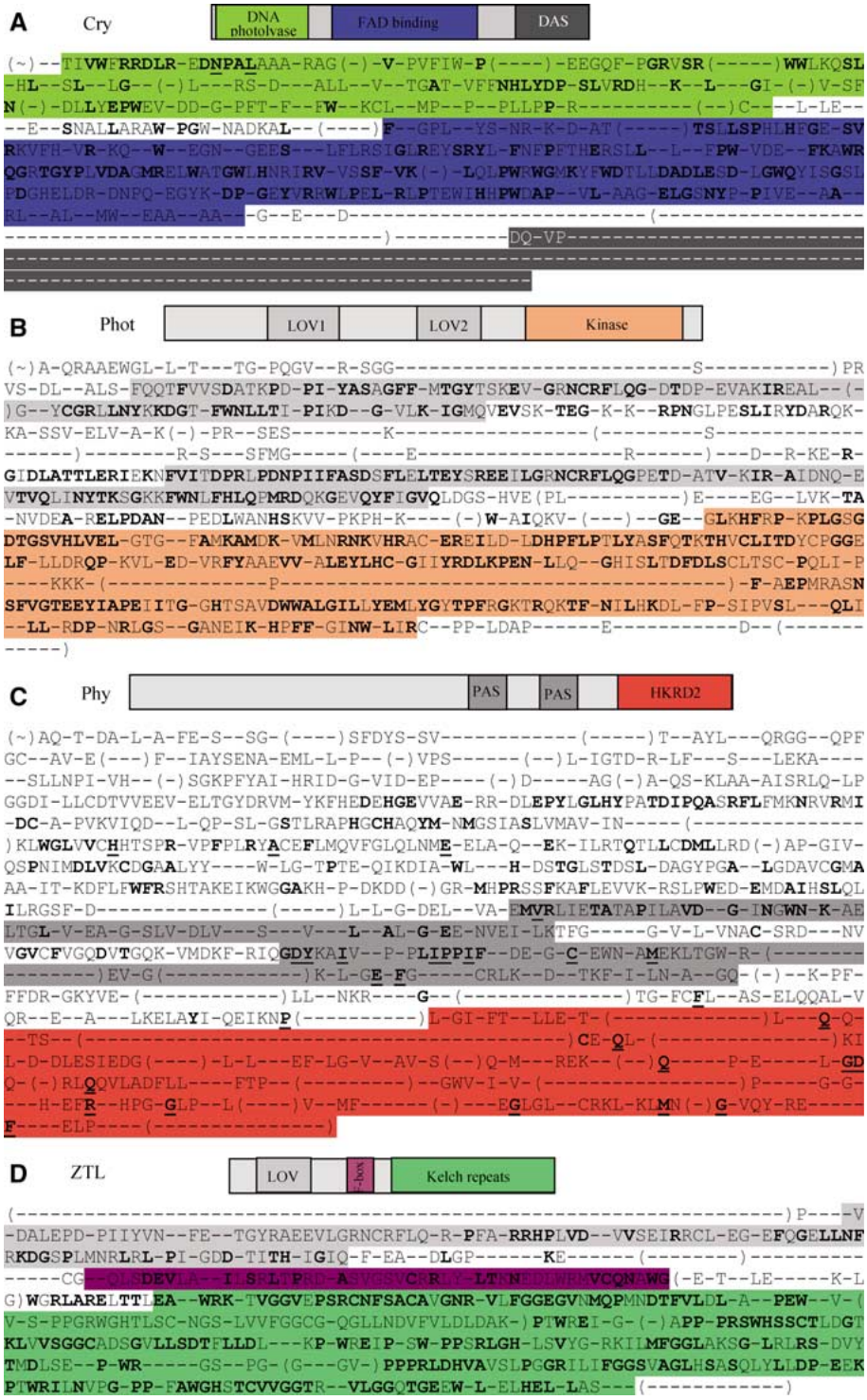


Fig. 2. Crys, phots, phys, and ZTLs. Primary structure, domains, and consensus sequences. (A) Cryptochromes; (B) phototropins; (C) phytochromes; (D) zeitlupes. Each domain is represented on the schematic structure and on the consensus sequences. The 60% consensus sequences were obtained from the complete sequences. The 100% consensus is represented in boldface. The residues underlined are also conserved in the non–green plant organisms. The sequences in brackets correspond to regions that are variable or not present in all the genes.

dence of their functions suggest that ZTL may be a novel class of blue light receptors (Imaizumi et al. 2003; Somers et al. 2000). ZTL sequences were found in *Arabidopsis thaliana*, *Adiantum capillus-veneris*, and *Neurospora crassa*, but with very low amino acid identity (Somers et al. 2000).

The UV-B region (280–320 nm) is not entirely absorbed by the ozone layer. In plants, a portion of this radiation not only triggers multiple cellular damages, but also could act as an informational signal. The above-mentioned photoreceptors are apparently not the UV-B sensors (Suesslin and Frohnmeyer 2003). Recently, ULI3 has been identified as a component of the UV-B signaling pathway. ULI3 encodes for an unknown protein containing a putative heme and diacylglycerol binding sites (Suesslin and Frohnmeyer 2003). To our knowledge, no data are available in the litera-

ture concerning the repartition of those putative UV-B sensors in the tree of life.

Numerous publications report independent studies of the different plant photoreceptors or the situation in a single specie. Until now, the presence of the different plant photoreceptors in the various species and their evolution have not been described in an exhaustive manner. The availability of an increasing number of plant sequences gave us the opportunity to search for photoreceptor encoding sequences in the major phyla of the plant kingdom. The elucidation of the individual photoreceptor function is often difficult due to their redundant and overlapping action. For each family of photoreceptors, using a phylogenetic tree based on the alignment of conserved amino acid sequences, we have traced back the evolution and the emergence of the diverse photoreceptors from an ancestral sequence. We have also tried to correlate the presence or the absence of particular photoreceptors with specific roles in the plant.

Our global phylogenetic analysis revealed that crys are highly divergent in ferns, phots are specific to green plants, and *Chlamydomonas* possesses only one cry and one phot, probably close to the origin of crys and phots plant sequences but no phys. ZLPs were found in Pteridophytes and Angiosperms. The putative UV-B photoreceptor ULI3 is actually restricted to Brassicacea. The composition of the nonmoss and nonfern photoreceptor families is conserved between species. The existence of a crucial evolution node suggested by partial crys analysis was confirmed: the cry1 and the cry2 branches were separated before the divergence between the Mono- and the Dicotyledons. The same appeared to be true for phots, where the phot1 and phot2 branches separated before the divergence between the two clades. Finally, the time of occurrence of the evolutionally nodes of phys was specified.

Materials and Methods

Retrieval of the Photoreceptor Sequences

Arabidopsis photoreceptor protein sequences from the different photoreceptor families have been used as a starting point for the data mining. Each of the different amino acid sequences of *Arabidopsis* was submitted against the whole genomic rice database with a tblastn search on the Rice Genome Project (RGP) Web site (http://rgp.dna.affrc.go.jp/).

In order to find photoreceptor sequences in other organisms, protein sequences from rice or *Arabidopsis* were used as input sequences in TBLASTN searches in different databases. Photoreceptors were sought at the NCBI Web site (www.ncbi.nih.gov/BLAST) and in several specialized databases such as the PEP EST database (www.moss.leeds.ac.uk), the PlantGDB database (www.plantgdb.org/), the DOE Joint Genome Institute (JGI) Web site (genome.jgi-psf.org/), MaizeGDB (www.maizegdb.org/), the Solanaceae genomic network (soldb.cit.cornell.edu), and the Plant Genome network (http://pgn.cornell.edu/). Nonannotated sequences were analyzed for the presence of the gene with different programs such as FGenesh (http://www.softberry.com/berry.phtml) and GenScan (http://genes.mit.edu/GENSCAN.html). The corresponding coding sequence (CDS) was translated with the "translate" tool on Expasy (http://us.expasy.org/tools/dna.html) and controlled for specific motifs.

The photoreceptor sequences of the following organisms have been used for the comprehensive phylogenetic analysis. *Adiantum capillus-veneris, Ac; Arabidopsis thaliana, At; Agrobacterium tumefaciens,* At; *Avena Sativa, As; Ceratopteris richardii,* Cr; *Ceratodon purpureus, Cp; Chlamydomonas reinhardtii, Cr; Dryopteris filix-mas, Df; Gloeobacter violaceus, Gv; Hypolepis punctata, Hp; Ipomoea nil, In; Lycopersicon esculantum, Le; Marchantia polymorpha, Mp; Mesembryanthemum crystallinum, Mc; Mougeotia scalaris, Ms; Neurospora crassa, Nc; Nicotiana plumbaginifolia, Np; Nicotiana tabaccum, Nt; Anabaena (Nostoc), Nos; Onoclea sensibilis,* Os; *Oryza sativa, Os; Physcomitrella patens, Pp; Pinus sylvestris,* Ps; *Pisum sativum, Ps; Populus balsamifera, Pb; Rhyzobium leguminosarum, Rl; Sorghum bicolor, Sb; Solanum tuberosum, St; Spinacia oleracea, So; Stellaria longipes, Sl; Tricicum aestivum, Ta; Vicia faba, Vf; Xanthomonas axonopodis, Xa; and Xenopus laevis, Xl.* All corresponding accession numbers are summarized in Fig. 3 and Supplemental Table.

Comprehensive Phylogenetic Analysis of Photoreceptor Sequences Identified in Plants

Among the high number of available photoreceptor sequences in the databases, only two sequences for each photoreceptor in Mono- and Eudicotyledons phyla were used to build the different trees. Complete sequences were used in most cases when the available sequences allowed it. The very short sequence of PsPhyO was aligned only with the *Arabidopsis* phys and the two other *Pinus* sequences. In the global phys tree, the branch length of PsPhyO is approximate, but not the position. Photoreceptor protein sequences present in various plants were aligned using Clustal W (Thompson et al. 1994). The alignment was further inspected and visually adjusted and realigned with Clustal X. Due to the large number of sequences analyzed; only a distance analysis was performed. The distance tree was constructed with the NEIGHBOR option of the PHYLIP 3.6a3 package (Felsenstein 1993) under the JTT substitution frequency matrix, and 1000 bootstrap replicates were carried out. The maximum likelihood tree was inferred with the PHYML algorithm, under the JTT substitution frequency matrix, by using the BIONJ starting tree (Guindon and Gascuel 2003). Njplot software was used to visualize phylogenetic trees, and BioEdit software to obtain the different consensus sequences. The consensus sequences represented in Fig. 2 correspond to the complete sequences of plants; 100% and 60% conserved residues are represented.

Results

Cryptochromes

Crys have homologous isoforms in all kingdoms (Cashmore et al. 1999). The crys/photolyase protein family is divided into five isolated and independent phylogenetic branches: the classes I and II CPD photolyases, the CryDASH, the plant crys, and, finally, the (6-4) photolyases and animals crys together (Brudler et al. 2003). Very few data were available in plants. The evolutional situation in plants has been investigated by looking at sequences of a large

	Procaryote	Animal	Algae	Bryophyte		Pterydophyte	Gymnosperm	Angiosperm			
				Liverworth	Moss			Monocotyledon		Dicotyledon	
			Cr					*Sb*	*Os*	*Le*	*At*
Cry	yes	yes	1	?	1 (*Pp*)	5 (*Ac*)	1 (*Ps*)	1	2	3	2
Cry3	yes	no	no	-	-	-	-		1		1
Pho	no	no	1	?	?	1 (*Ac,Cr*))	2 (*Ps*)	2	2	2	2
Phy-Phot	no	no	no			1 (*Ac, Df, Os, Hp*)	-	-	no	no	no
Phy	Bacterio-phy	no	no	1 (*Mp*)	2 (*Cp*), 4 (*Pp*)		2 (*Ps*)	3	3	5	5
ZTL	no	no	no	-	-	1 (*Cr*)		2 or 3	3	2 or 3	3
ULI3	no	no	no	-	-	-	-	-	25 %	29 %	2

Fig. 3. Summary of the photoreceptors composition in the different phyla. Gray boxes correspond to completed genome projects. Question marks represent the absence of photoreceptor sequence, probably due to the low number of EST. Minus marks correspond to the absence of photoreceptor sequence. Plus marks are used for photoreceptor sequences that are present in various numbers according to the species. The cells surrounded by bold lines correspond to a group of organisms in which the emergence and the evolution of a photoreceptor family are well defined. For the initials of the species refer to Materials and Methods.

number of representative plant organisms. Except for the class I CPD photolyases, which do not exist in plants, the four other divisions were confirmed (Fig. 4). The class II CPD photolyases branch forms an independent group with any duplication suggesting that it is closer to an ancestral sequence. This ancestral branch led to the emergence of three independent phylogenetic groups: CryDASH, (6-4) photolyases, and plant cryptochromes (step 1 and step 1bis; Fig. 4). Interestingly the green alga *Chlamydomonas reinhardtii* contained only a single cry sequence (CrCPH; shaded gray in Fig. 4). The *Chlamydomona*s cry sequence has been situated between the bacteriocry sequence from *Agrobacterium tumefaciens* (At-Cry) and the other plant cry sequences. Therefore the origin of the plant cryptochromes could be related to the unicellular green alga cry sequence (Fig. 4). The duplication of this single ancestral sequence led to cry1 and cry2 sequences, found, respectively, in two independent branches (step 2; Fig. 4). Mono- and Eudicotyledon cry sequences were present in both cry1 and cry2 branches. This indicated that the duplication event that resulted in the separation of these cry1/cry2 branches occurred after the origin of the seed plants but before the Mono- and Eudicotyledon separation (step 3; Fig. 4). This idea was confirmed by the existence of two independent cry sequences in *Amborella tricopoda*, a basal Angiosperm. Unfortunately these sequences were too short to be incorporated in the phylogentic analysis. Surprisingly, five cry isoforms resulting from four duplication events were present in the fern *Adiantum* (AcCry1 to AcCry5) and constituted an independent group in the cry1 branch. Apart from the special cases of *Adiantum* and of *Lycopersicon*, which contain two cry1 sequences (LeCryc and -b), two copies of cry sequences were detected in all examined species. The separation between the two clades (Mono- and Eudicotyledons) is well respected.

The existence of homologies in Prokaryotes for cry1/2 and CryDASH, respectively, supported the idea that cry1/2 and CryDASH resulted from one duplication event in an ancestral organism and evolved separately after the divergence of Prokaryotes and Eukaryotes (Brudler et al. 2003) (Fig. 4). However, the possibility of convergent evolution within the cry family remains. This idea was also supported by the molecular structure: the DNA photolyase and FAD binding domains were highly conserved through the various species and between the different cry sequences, whereas the C-Term DAS domain, which was absent in CryDASH sequences, was poorly conserved within the cry1/2 branch (Fig. 2A). Therefore, the conserved region (DNA photolyase and FAD binding domains) of the crys probably corresponds to the ancestral sequence. The gene structure also differed between CryDASH and cry1/2, which contained, respectively, 11 and 3 introns in *Arabidopsis*. These divergences could be directly the result of the evolution of cry1/2. This hypothesis was supported by the absence of CryDASH sequence in *Chlamydomonas* and the low level of identities between the plant and the animal sequences (27% between *Arabidopsis* and human). Despite the significance of the acronym DASH (Drosophila, *Arabidopsis*, Synechocystis, Human) (Brudler et al. 2003), cryDASH sequences were not found in insects or in Mammals.

Phototropins

Contrary to crys that exist in various phyla, phots sequences were found only in green organisms

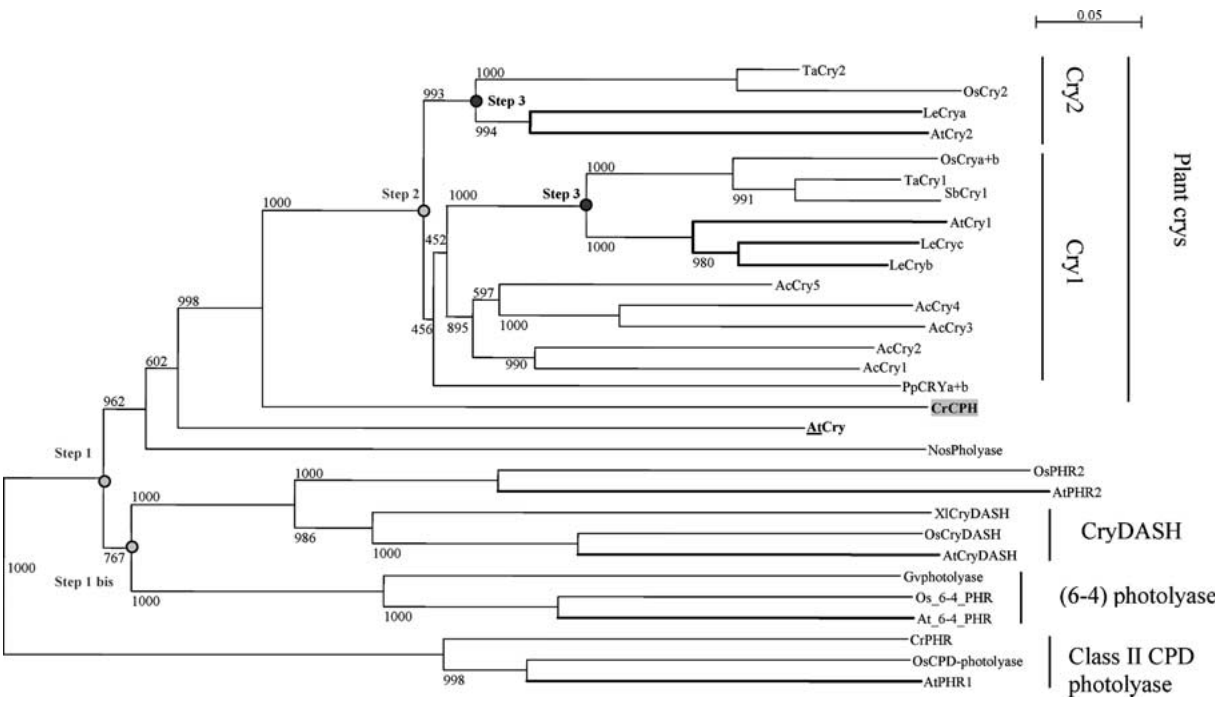


Fig. 4. Phylogenetic relationships of the cryptochromes from plants based on protein sequence alignments. Several bacterio-cryptochrome protein sequences were used to root the tree. The bold lines represent cry sequences from Eudicotyledons. The main nodes corresponding to putative duplications are symbolized with a gray circle and a step number. All branches are drawn to scale and the scale bar represents 0.05 substitution per nucleotide.

(Fig. 5). To our knowledge, no phot phylogenetic analysis has been reported in plants. Our analysis showed that *Chlamydomonas* contained only one phot sequence, called CrPhot1. CrPhot1 was anterior to the plant phot ancestors and independent of the protein kinase-like proteins (used to root the phot branch). Thus CrPhot1 could be related to the phot ancestral sequence from which the plant phots evolved (step 1; Fig. 5). No phot-encoding sequence has been found in liverworts and mosses, probably due to the low level of expression of phots (for example, four ESTs in *Arabidopsis*). One major duplication event seemed to have occurred since the origin of the seed plants, which generated the Phot1 and Phot2 branches (step 2; Fig. 5). Phy–Phot chimeric sequences were found in Pteridophytes (Kawai et al. 2003). They formed an independent branch from the seed plant phots and have evolved from one single phot sequence after the appearance of the seed plants (step 3; Fig. 5). Both phot1 and phot2 branches contained Mono- and Eudicotyledon phot sequences. This indicated that the Mono- and Eudicotyledon separation occurred after the duplication event that led to the separation of the phot1/phot2 branches (step 4; Fig. 5). A more recent duplication led to the existence of two Phot1 sequences in the Eudicotyledonous *Pisum* sativum and *Vicia faba*.

At the molecular level, the sequences of the two LOV domains and kinase domain were highly conserved between the isoforms and the species (Fig. 2B). Supporting the importance of these domains, the interdomain sequences were poorly conserved.

Phytochromes

Phys are present in plants, prokaryotes, and fungi (Montgomery and Lagarias 2002). The ancestral phytochrome-like protein could have originated both the bacteriophytochromes and the phys of photosynthetic organisms (step 1; Fig. 6). Surprisingly, no phy-like protein sequence has been found in *Chlamydomonas*. From the emergence of the terrestrial plants to the higher plants, the phy family has followed several evolutions. As for crys, Bryophytes and Pteridophytes phys sequences constituted an independent branch versus phys of higher plants (step 2; Fig. 6). Phylogenetic analyses of Gymno- and Angiosperms phys protein sequences revealed four major duplication events. Analysis of complete or significant sequences indicated that *Pinus* contained three phys as previously shown in a study dedicated to the Gymnosperm phys and using short sequences (Schmidt and Schneider-Poetsch 2002). The global phylogenetic analysis in the major plant phyla allowed positioning of the first two duplications. PsphyN and PsphyO from *Pinus* were related to phyA and phyC, respectively, and PsPhyP to the

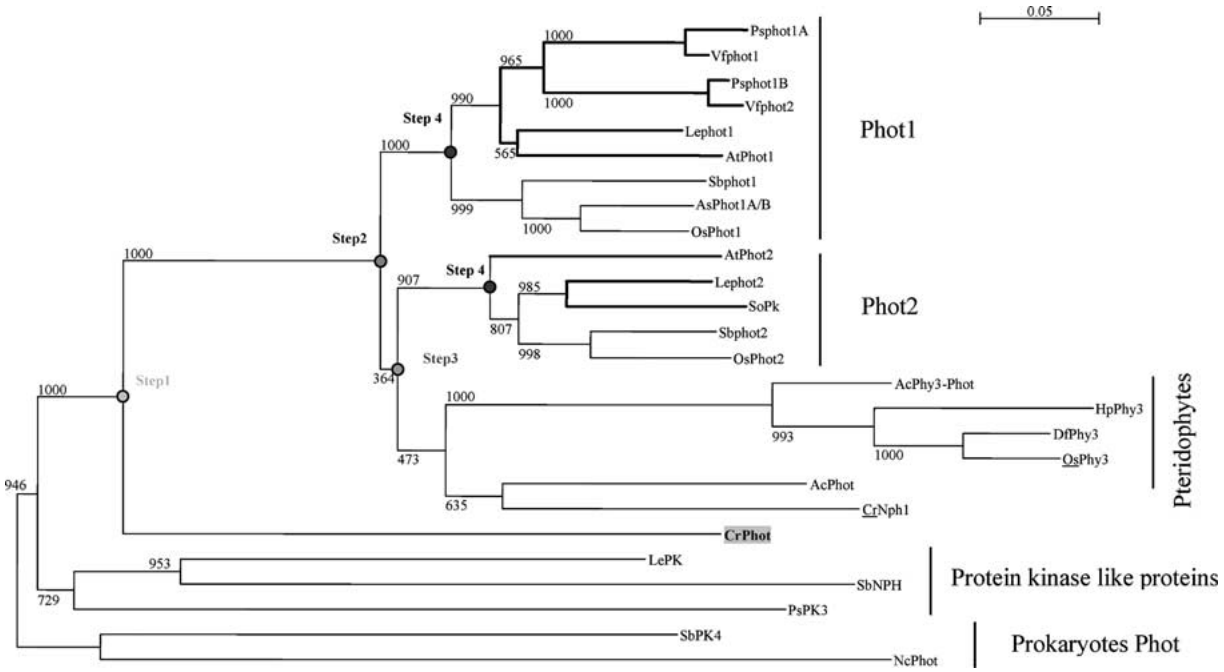


Fig. 5. Phylogenetic relationships of the phototropins from plants based on protein sequence alignments. Phot sequences from prokaryotes were used to root the tree. The bold lines represent the phot sequences from Eudicotyledons. The main nodes corresponding to putative duplications are symbolized by a gray circle and a step number. All branches are drawn to scale and the scale bar represents 0.05 substitution per nucleotide.

phyB/D/E branch. Therefore, the first two duplications arose before the separation between Gymnosperms and Angiosperms. The first one generated the phyA/C and phyB/D/E lines (step 3; Fig. 6) and the second one led to the separation between phyA to phyC (step 4; Fig. 6).

In Angiosperms, two independent events separated phyB/D from phyE (step 5; Fig. 6) and phyB from phyD (step 6; Fig. 6). Only one sequence of the phyB/D/E subfamily was found in the Monocotyledons. This suggested that the two latest duplication events within the phyB/D/E branch likely occurred after the separation of the two clades and only in the eudicotyledonous plants. A special evolution could be observed for the last duplication between PhyB and phyD. PhyB- and phyD-related sequences were closer between paralogues than between orthologues (supplemental data) (Pratt 1995). Two hypotheses could explain this phenomenon. First, the last duplication could have occurred independently in the various eudicotyledonous species. Second, and more probably, the phyB/D separation occurred before the emergence of the major Eudicotyledon families and an evolutional convergence within each species led to the actual phylogenetic situation.

The level of phys sequence conservation, even within domains, is much lower than that of the phots and crys photoreceptors (Fig. 2C). This low conservation rate could be due to the higher number of duplications (steps 1 to 6; Fig. 6).

Zeitlupe

Concerning plants, previous studies revealed the presence of ZTLs only in *Adiantum capillus-veneris* and in *Arabidopsis thaliana* (Somers et al. 2000). The analysis we conducted showed that ZTL sequences were actually present in Pteridophytes and Angiosperms (Fig. 7). No ZTL homologue was found in prokaryotic organisms, in *Chlamydomonas*, in Bryophytes, in Gymnosperms, or in animals. The small size of the EST databases and the low expression level of the ZTL gene (10 ESTs in *Arabidopsis*) might explain the absence of ZTL homologues in Gymnosperm and Bryophyte species. Interestingly, the ZTL sequence contained homologies with the two phot LOV domains. Three homologues have been found in the Mono- and Eudicotyledons, resulting from two duplication events (Fig. 7). Incomplete sequences from *Sorghum bicolor* and *Lycopersicum* were not represented but confirmed the previous idea. The first duplication occurred before the separation of Mono- and Eudicotyledons and led to the separation of the ZTL1 and ZTL2 branches (step 1; Fig. 7). The second appeared more recently, after the separation between Mono- and Eudicotyledons (step 2; Fig. 7). The ZTL sequence from the fern *Ceratopteris*, named CrZTL, appeared in the AtFBX2b/c branch with a low bootstrap support. The weakness of its position and the fact that only one sequence of a Pteridophyte has been found suggested that CrZTL could be close

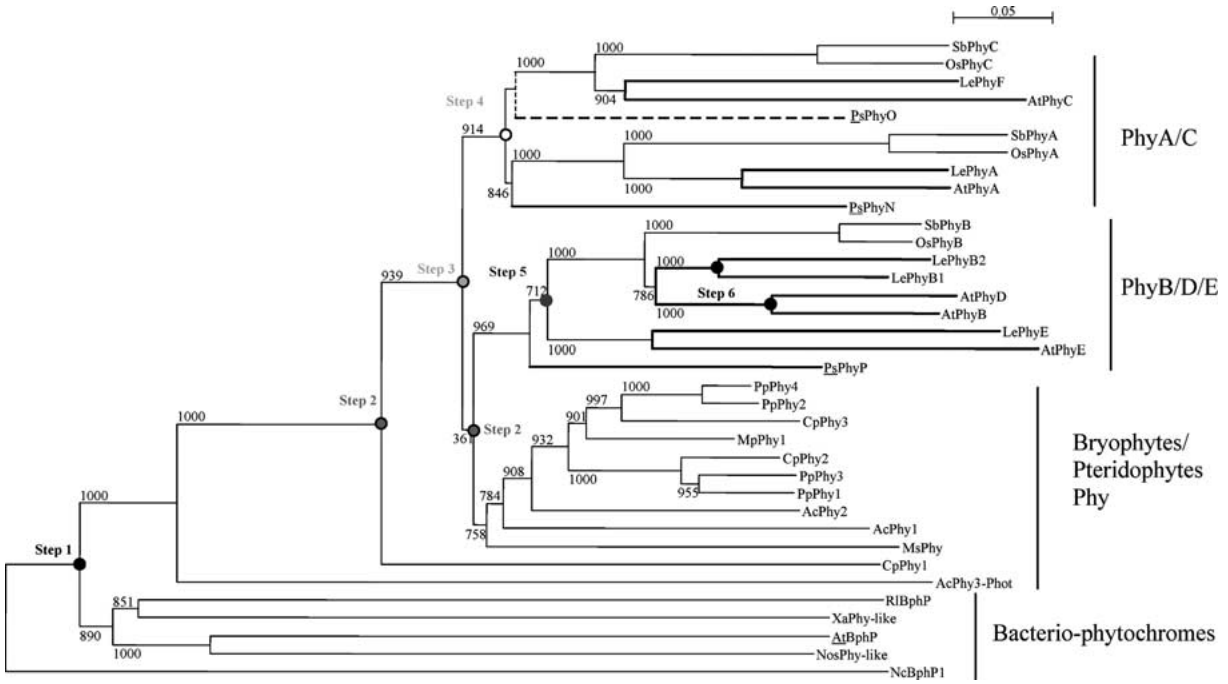


Fig. 6. Phylogenetic relationships of the phytochromes from plants based on protein sequence alignments. Several bacteriophytochrome protein sequences were used to root the tree. The gray and the bold lines represent the phy sequences from Gymnosperms and Eudicotyledons, respectively. The approximate length of the PsPhyO branch is represented by a dashed line. The main nodes corresponding to putative duplications are symbolized by a gray circle and a step number. All branches are drawn to scale and the scale bar represents 0.05 substitution per nucleotide.

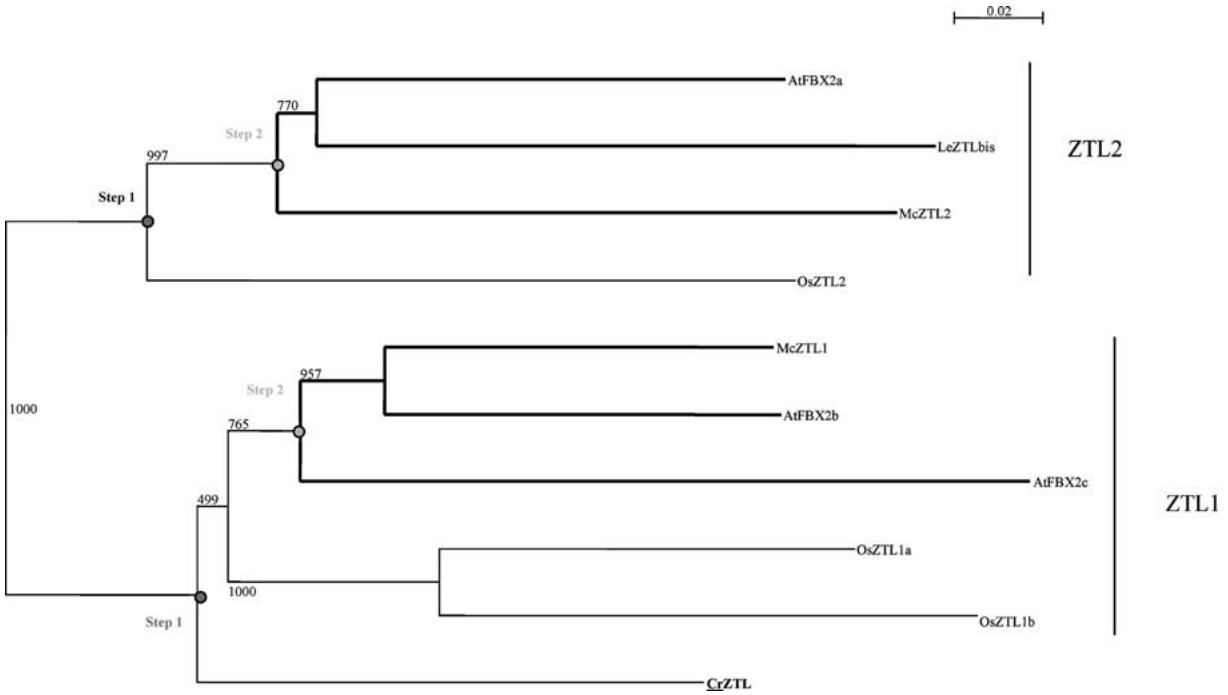


Fig. 7. Phylogenetic relationships of the zeitlupes from plants based on protein sequence alignments. The bold lines represent the ZTL sequences from Eudicotyledons. The main nodes corresponding to putative duplications are symbolized by a gray circle and a step number. All branches are drawn to scale and the scale bar represents 0.02 substitution per nucleotide.

to the ancestral sequence. The observation that ZTLs appear to be present only in Pteridophytes and Angiosperms suggests that ZTLs arose later during plant evolution and that they have arisen twice (in Pteridophytes and then in Angiosperms) or that they have been lost in Gymnosperms.

Supporting the idea of a recent sequence appearance during the plant evolution, the three ZTL domains were highly conserved (Fig. 2D).

Putative UV-B Photoreceptor

AtULI3, a putative receptor of UV-B perception in *Arabidopsis* (Suesslin and Frohnmeyer 2003), had one close paralogue and one orthologue in *Brassica oleracea* (data not shown). No close homologous sequence has been found in either animals or *Chlamydomonas*. The most homologous genes found in the other green plants showed only 20 to 25% homologies for restricted parts of the sequence. This suggested that this putative UV-B photoreceptor is Brassicaceae-specific.

Maximum-Likelihood Tree

Analyses with maximum-likelihood (ML) trees (data not shown) have been performed for each plant photoreceptor family and supported that the grouping and evolution observed with the distance trees are not the result of long branch attraction.

Discussion

Monocotyledons (rice) and Eudicotyledons (*Arabidopsis*) diverged from a common ancestor, some 150 MY ago (Wikstrom et al. 2001). It can be assumed that, at their origin, the first Monocotyledon and Eudicotyledon had a similar number of photoreceptor genes. A differential evolution leading to today's *Arabidopsis* and rice yields to a variation in the gene numbers and sequences. The evolution rate varies between the phyla but also between the photoreceptor families and probably between the different domains. Indeed the N-Term domain (putative output domain) of the phy is evolving quickly (Alba et al. 2000b; Matsushita et al. 2003).

Stability in Photoreceptor Copy Number

Compared to the high duplication rate of the peroxidase genes (Passardi et al. 2004), the photoreceptor families showed a slow evolution in copy number since the emergence of the green plants. The large increase in peroxidase copy number could result from conservation of duplicated peroxidase genes through evolution. The elevated peroxidase copies could be related to the large range of physiological processes in which the proteins are implicated (Hiraga et al. 2001). On the contrary, the size conservation of the photoreceptor families could be associated with the major functions of the photoreceptors in light perception. Indeed, the photoreceptors, which contained complex output domains, need mainly to detect light quantity, quality, direction, and periodicity and to transmit the information.

In general, Angiosperms possess two crys, three or five phys, and two phots, but special cases related to specific duplication events were occasionally observed. *Lycopersicon* contained three cry-encoding sequences; *Vicia faba* and *Pisum sativum* both possess two phot1 homologous sequences. These atypical situations could be due to differential duplication rates between species or related to specific biotopes.

Phytochrome Duplications

The photoreceptor families have followed a different evolution despite the conservation of gene number. Specific appearance and evolution of certain photoreceptor families could be related to particular phyla or to a modification of the biotope (Fig. 3). For example, in the phy family the separation between phyB/D/E and phyA/C occurred before the Monocot/Eudicotyledon divergence, whereas the separation between phyB/D and phyE and between phyB from phyD occurred after the Monocot/Eudicotyledon separation.

Within the different photoreceptor families, the homologies were higher between orthologues than between paralogues. This rule was not respected in the case of phyB and phyD. Indeed phyB- and phyD-related sequences were closer between paralogues within each Eudicotyledon group. This situation could be the result of a convergent evolution that led to the actual phylogenetic image of the Eudicotyledon family evolution (Wikstrom et al. 2001). The result of exhaustive phyB/D/E phylogenetic analysis (supplementary data) showed, as expected, that Caryophyllales were anterior to Rosids and Asterids. In opposition to the usual phylogenetic analysis (Wikstrom et al. 2001), in our analysis, the Asterids looked posterior to the Rosids. This observation was strongly supported by high bootstraps. The discrepancies concerning the relative position of the major Eudicotyledon groups in global evolution could be addressed with a systematic research of the phyB- and phyD-related sequences and their phylogenetic analysis.

Photoreceptor Evolution and Functions

Many data concerning photoreceptor functions have been accumulated for *Arabidopsis*. They can be used as a reference to understand the function of the photoreceptors in relation to their evolution.

Phys and crys encoding sequences were found in many organisms with variations of the sequence and

an increase in copy numbers. Regarding their major role in circadian clock entrainment (Chen et al. 2004), these two protein families are probably important for a normal growth in these organisms. In *Arabidopsis*, phys and crys are also implicated in other light-regulated processes such as flowering, de-etiolation, and gravitropism inhibition (Briggs et al. 2001; Lariguet and Fankhauser 2004a; Mockler et al. 2003). These specific functions could be achieved thanks to particular phys or crys isoforms or to specific interactions of these photoreceptors with privileged partners. For example, the phy binding proteins Phytochrome Kinase Substrate (PKS) are only found in superior Angiosperms, and not in basal Angiosperms such as Amborellacea and Gnetacea (supplementary data) or in "lower" organisms. Other partners, such as COP1, HY5, and DET1, are present in all plant organisms with high homology and in animals with only a lower level (30 to 40%) (data not shown).

PhyA, -B, and -E have a role in *Arabidopsis* seed germination (Sullivan and Deng 2003). PhyA and -B have existed since the Gymnosperms, while phyE is detected only in the Eudicotyledons. Therefore, it can be speculated that phyE could have a specific function for the germination of Eudicotyledons. Shade avoidance is a mechanism used by the majority of Gymno- and Angiosperms to compete with plant neighbors (Gilbert et al. 2001). In addition to their own individual functions, phyB, -D, and -E, which are issued from a common ancestor, act redundantly to mediate *Arabidopsis* shade avoidance (Morelli and Ruberti 2002). The *Pinus* phy, a member of this branch, could concentrate all the functions in germination of phy B, -D, and -E in a single molecule.

In *Arabidopsis*, phots are involved in photosynthesis optimization by controlling chloroplast movements, phototropism, and stomata opening (Briggs and Christie 2002; Celaya and Liscum 2004; Kinoshita et al. 2001). In accordance with the phot functions, only the chloroplast-containing organisms contain phot encoding sequences. *Chlamydomonas* contains a unique chloroplast that probably does not undergo relocation. In this case, optimization of light capture is accomplished by the movement of the organisms themselves (phototaxis) (Schaller et al. 1997). This unicellular algae possesses only one ancestral phototropin known to be required for multiple aspects of sexual development (Huang and Beck 2003) and able to function in *Arabidopsis* (Onodera et al. 2005). *Chlamydomonas* lacks phy. It is possible that its single phot is responsible for *Chlamydomonas* movement in response to blue light intensity. It was not possible to establish the presence of phot sequences in Bryophytes but phots exist in Pterydophytes. Since phototropism can be mediated by both phots and phys (in higher plants and mosses, respectively), determining if phot sequences are actually absent from Bryophytes is an important issue.

Hybrid Phy–Phot in Ferns

Particular situations have been observed that diverged from the usual photoreceptor sequences and numbers. *Adiantum* was of particular interest because it contained at least five crys, two phys, one phot, and one phy–phot hybrid. The lack of sequences (genomic or EST) does not allow us to determine if all the other phy–phot hybrid-containing ferns also possess a large number of other photoreceptors. *Adiantum* possesses highly divergent crys. Assuming that these *Adiantum* cry sequences are not splice variant, this observation implies that crys can undergo rapid gene duplication/diversification in some plant species including *Adiantum*. Phy–phot encoding sequence has been found only in a particular group of ferns. This hybrid sequence probably appeared during a genomic rearrangement which occurs rapidly after the emergence of the Pteridophytes. The acquisition of this so-called phy3 sequence could be a critical step for the proliferation of ferns in a low-light environment (Kawai et al. 2003).

Acknowledgments. We thank Filippo Passardi for his critical reading and his constructive suggestions. We thank the University of Geneva and the Swiss National Science Foundation for research support: Grant 31-068003.02 to C.D. and the Marie Heim-Vögtlin program to P.L. and the Societe Académique de Genève.

References

Alba R, Cordonnier–Pratt MM, Pratt LH (2000a) Fruit-localized phytochromes regulate lycopene accumulation independently of ethylene production in tomato. Plant Physiol 123:363–370

Alba R, Kelmenson PM, Cordonnier–Pratt MM, Pratt LH (2000b) The phytochrome gene family in tomato and the rapid differential evolution of this family in angiosperms. Mol Biol Evol 17:362–373

Briggs WR, Beck CF, Cashmore AR, Christie JM, Hughes J, Jarillo JA, Kagawa T, Kanegae H, Liscum E, Nagatani A, Okada K, Salomon M, Rudiger W, Sakai T, Takano M, Wada M, Watson JC (2001) The phototropin family of photoreceptors. Plant Cell 13:993–997

Briggs WR, Christie JM (2002) Phototropins 1 and 2: versatile plant blue-light receptors. Trends Plant Sci 7:204–210

Brudler R, Hitomi K, Daiyasu H, Toh H, Kucho K, Ishiura M, Kanehisa M, Roberts VA, Todo T, Tainer JA, Getzoff ED (2003) Identification of a new cryptochrome class. Structure, function, and evolution. Mol Cell 11:59–67

Cashmore AR, Jarillo JA, Wu YJ, Liu D (1999) Cryptochromes: blue light receptors for plants and animals. Science 284:760–765

Celaya RB, Liscum E (2005) Phototropins and associated signaling: Providing the power of movement in higher plants. Photochem Photobiol 81: 73–80

Chen M, Chory J, Fankhauser C (2004) Light signal transduction in higher plants. Annu Rev Genet 38:87–117

Daiyasu H, Ishikawa T, Kuma K, Iwai S, Todo T, Toh H (2004) Identification of cryptochrome DASH from vertebrates. Genes Cells 9:479–495

Felsenstein J (1993) PHYLIP (Phylogeny Inference Package). University of Washington, Seattle

Gilbert IR, Jarvis PG, Smith H (2001) Proximity signal and shade avoidance differences between early and late successional trees. Nature 411:792–795

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696–704

Gyula P, Schafer E, Nagy F (2003) Light perception and signalling in higher plants. Curr Opin Plant Biol 6:446–452

Han L, Mason M, Risseeuw EP, Crosby WL, Somers DE (2004) Formation of an SCF(ZTL) complex is required for proper regulation of circadian timing. Plant J 40:291–301

Hiraga S, Sasaki K, Ito H, Ohashi Y, Matsui H (2001) A large family of class III plant peroxidases. Plant Cell Physiol 42:462–468

Huala E, Oeller PW, Liscum E, Han IS, Larsen E, Briggs WR (1997) Arabidopsis NPH1: a protein kinase with a putative redox-sensing domain. Science 278:2120–2123

Huang K, Beck CF (2003) Phototropin is the blue-light receptor that controls multiple steps in the sexual life cycle of the green alga Chlamydomonas reinhardtii. Proc Natl Acad Sci USA 100:6269–6274

Imaizumi T, Tran HG, Swartz TE, Briggs WR, Kay SA (2003) FKF1 is essential for photoperiodic-specific light signalling in Arabidopsis. Nature 426:302–306

Kawai H, Kanegae T, Christensen S, Kiyosue T, Sato Y, Imaizumi T, Kadota A, Wada M (2003) Responses of ferns to red light are mediated by an unconventional photoreceptor. Nature 421:287–290

Kinoshita T, Doi M, Suetsugu N, Kagawa T, Wada M, Shimazaki K (2001) Phot1 and phot2 mediate blue light regulation of stomatal opening. Nature 414:656–660

Lagarias DM, Crepeau MW, Maines MD, Lagarias JC (1997) Regulation of photomorphogenesis by expression of mammalian biliverdin reductase in transgenic arabidopsis plants. Plant Cell 9:675–688

Lamparter T (2004) Evolution of cyanobacterial and plant phytochromes. FEBS Lett 573:1–5

Lariguet P, Fankhauser C (2004a) The effect of light and gravity in hypocotyl growth orientation. In: Wada M (ed) 58th Yamada Conference. Springer, New York

Lariguet P, Fankhauser C (2004b) Hypocotyl growth orientation in blue light is determined by phytochrome A inhibition of gravitropism and phototropin promotion of phototropism. Plant J 40: 826–834

Lin C (2000) Photoreceptors and regulation of flowering time. Plant Physiol 123:39–50

Lin C, Shalitin D (2003) Cryptochrome structure and signal transduction. Annu Rev Plant Biol 54:469–496

Mas P, Kim WY, Somers DE, Kay SA (2003) Targeted degradation of TOC1 by ZTL modulates circadian function in *Arabidopsis thaliana*. Nature 426:567–570

Mathews S, Sharrock RA (1997) Phytochrome gene diversity. Plant Cell Environ 20:666–671

Mathews S, Burleigh JG, Donoghue MJ (2003) Adaptive evolution in the photosensory domain of phytochrome a in early angiosperms. Mol Biol Evol 20:1087–1097

Matsushita T, Mochizuki N, Nagatani A (2003) Dimers of the N-terminal domain of phytochrome B are functional in the nucleus. Nature 424:571–574

Mockler T, Yang H, Yu X, Parikh D, Cheng YC, Dolan S, Lin C (2003) Regulation of photoperiodic flowering by Arabidopsis photoreceptors. Proc Natl Acad Sci USA 100: 2140–2145

Moller SG, Ingles PJ, Whitelam GC (2002) The cell biology of phytochrome signalling. New Phytol 154:553–590

Montgomery BL, Lagarias JC (2002) Phytochrome ancestry: sensors of bilins and light. Trends Plant Sci 7:357–366

Morelli G, Ruberti I (2002) Light and shade in the photocontrol of Arabidopsis growth. Trends Plant Sci 7:399–404

Onodera A, Kong SG, Doi M, Shimazaki KI, Christie J, Mochizuki N, Nagatani A (2005) Phototropin from Chlamydomonas reinhardtii is functional in Arabidopsis thaliana. Plant Cell Physiol 46: 367–374

Panda S, Hogenesch JB, Kay SA (2002) Circadian rhythms from flies to human. Nature 417:329–335

Passardi F, Longet D, Penel C, Dunand C (2004) The plant peroxidase multigenic family in rice and its evolution in green plants. Phytochemistry 65:1879–1893

Pratt LH (1995) Phytochromes—Differential properties, expression patterns and molecular evolution. Photochem Photobiol 61:10–21

Sakai T, Kagawa T, Kasahara M, Swartz TE, Christie JM, Briggs WR, Wada M, Okada K (2001) Arabidopsis nph1 and npl1: blue light receptors that mediate both phototropism and chloroplast relocation. Proc Natl Acad Sci USA 98: 6969–6974

Sancar A (1994) Structure and function of DNA photolyase. Biochemistry 33:2–9

Sancar A (2003) Structure and function of DNA photolyase and cryptochrome blue-light photoreceptors. Chem Rev 103:2203–2237

Schaller K, David R, Uhl R (1997) How Chlamydomonas keeps track of the light once it has reached the right phototactic orientation. Biophys J 73:1562–1572

Schmidt M, Schneider-Poetsch HA (2002) The evolution of gymnosperms redrawn by phytochrome genes: the Gnetatae appear at the base of the gymnosperms. J Mol Evol 54:715–724

Schultz TF, Kiyosue T, Yanovsky M, Wada M, Kay SA (2001) A role for LKP2 in the circadian clock of *Arabidopsis*. Plant Cell 13:2659–2670

Somers DE, Schultz TF, Milnamow M, Kay SA (2000) ZEITLUPE encodes a novel clock–associated PAS protein from Arabidopsis. Cell 101:319–329

Suesslin C, Frohnmeyer H (2003) An Arabidopsis mutant defective in UV-B light-mediated responses. Plant J 33:591–601

Sullivan JA, Deng XW (2003) From seed to seed: the role of photoreceptors in Arabidopsis development. Dev Biol 260:289–297

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Wang H, Ma LG, Li JM, Zhao HY, Deng XW (2001) Direct interaction of Arabidopsis cryptochromes with COP1 in light control development. Science 294:154–158

Wikstrom N, Savolainen V, Chase MW (2001) Evolution of the angiosperms: calibrating the family tree. Proc R Soc Lond B Biol Sci 268:2211–2220

Yang HQ, Tang RH, Cashmore AR (2001) The signaling mechanism of Arabidopsis CRY1 involves direct interaction with COP1. Plant Cell 13:2573–2587

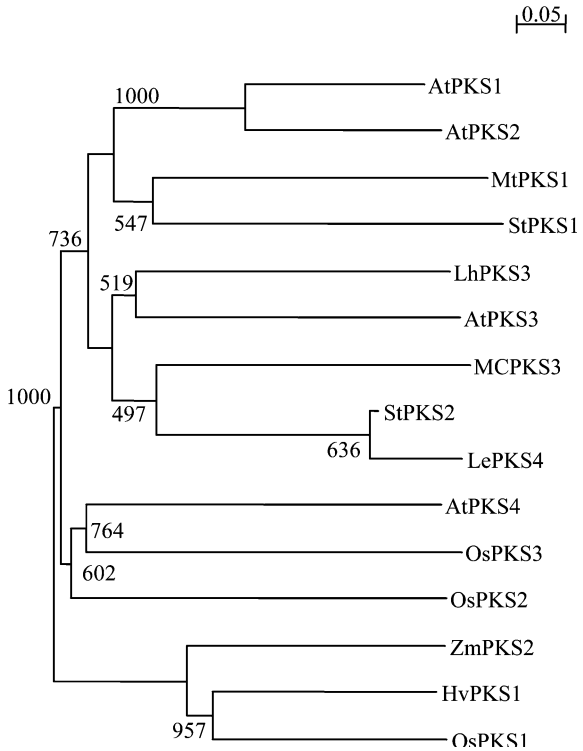
Supplemental Table. Name of the various photoreceptor sequences used in the phylogenetic analysis and their corresponding accession numbers. Few accession numbers are associated when the sequence used has been compiled. TC numbers correspond to the accessions obtained from TIGR (http://www.tigr.org/tdb/tgi/plant.shtml). SGN numbers correspond to the accessions obtained from Solanaceae Genomics Network (http://soldb.cit.cornell.edu/).

	Cry	Phot	Phy	UV-B	ZTL	other
Adiantum capillus-veneris	AcCry1 AB012626 AcCry2 AB012627 AcCry3 AB012628 AcCry4 AB028928 AcCry5 AB028929	AcPhot AB037188 AcPhy3 AB012082	AcPhy3 AB012082 AcPhy1 AB016168 AcPhy2 AB016232			
Agrobacterium tumefaciens	AtCry AE008050		AtBphB AAL43154			
Anabaena/Nostoc	NosPholyase AP003590		Nosphy-like AB028873			
Arabidopsis thaliana	AtCry1 AY124863 AtCry2 NM_100320 AtCryDASH AB062926 At(6-4)PHR photolyase AB003687 AtPHR2 AF053366 AtPHR1 AF053365	AtPhot1 AY040062 AtPhot2 AF053941	AtPhyA X17341 AtPhyB X17342 AtPhyC X17343 AtPhyD X76609 AtPhyE X76610	AtULI3 At5G59920 AtUL3bis At4g01350	AtFKF1 AAF32298 AtZTL AF254413 AtLKP2 AB038797	AtHY5 AB005456 AtCop1 NM 128855 AtPKS1 AY063721 AtPKS2 AY088864 AtPKS3 AK118063 AtPKS4 BT015323
Avena sativa		AsPhot1 A/B AAC05083, AAC05084				
Ceratodon purpureus			CpPhy1 U56698 CpPhy2 U72993 CpPhy3 AY123149			
Ceratopteris richardii	CrNph1 BE642226	CrNPH1 BE641733			CrZTL BQ087648	
Chlamydomonas reinhardtii	CrCPH CREDNAPL	CrPhot CAC94941				
Dryopteris filix		DfPhy3 BAC55265				
Gloeobacter violaceus	Gvphotolyase BAC89690					
Hordeum vulgar						HvPKS1 BJ469192, BG368823
Hypolepis punctata		HpPhy3 AB089918				
Ipomoea nil			InPhyE U39787			
Lycopersicon esculentum and hirsutum	LeCrya AF130423 LeCryb AF348461 LeCryC AF130425	LePK AF143505 LePhot1 TC155942 LePhot2 TC163480	LePhyF AF178571 LePhyA AJ001913 LePhyB1 AJ002281 LePhyB2 AF122901 LePhyE AF178571		LeZTL1 AW934013 BF114438 LeZTL2 TC167826 LeZTL3 SGN-U220996	LePKS4 AI897288 LhPKS3 AW616220
Marcantia paleacea			MpPhy1 AB022917			
Medicago truncatula						MtPKS1 BE124133
Mesembryanthemum crystallinum					McZTL1 AY371290 McZTL2 AY371291	McPKS3 BG269652
Mougeotia Scalaris			MsPhy S52048			
Neurospora crassa		NcPhot XM325644	NcBphP1 XM325644			

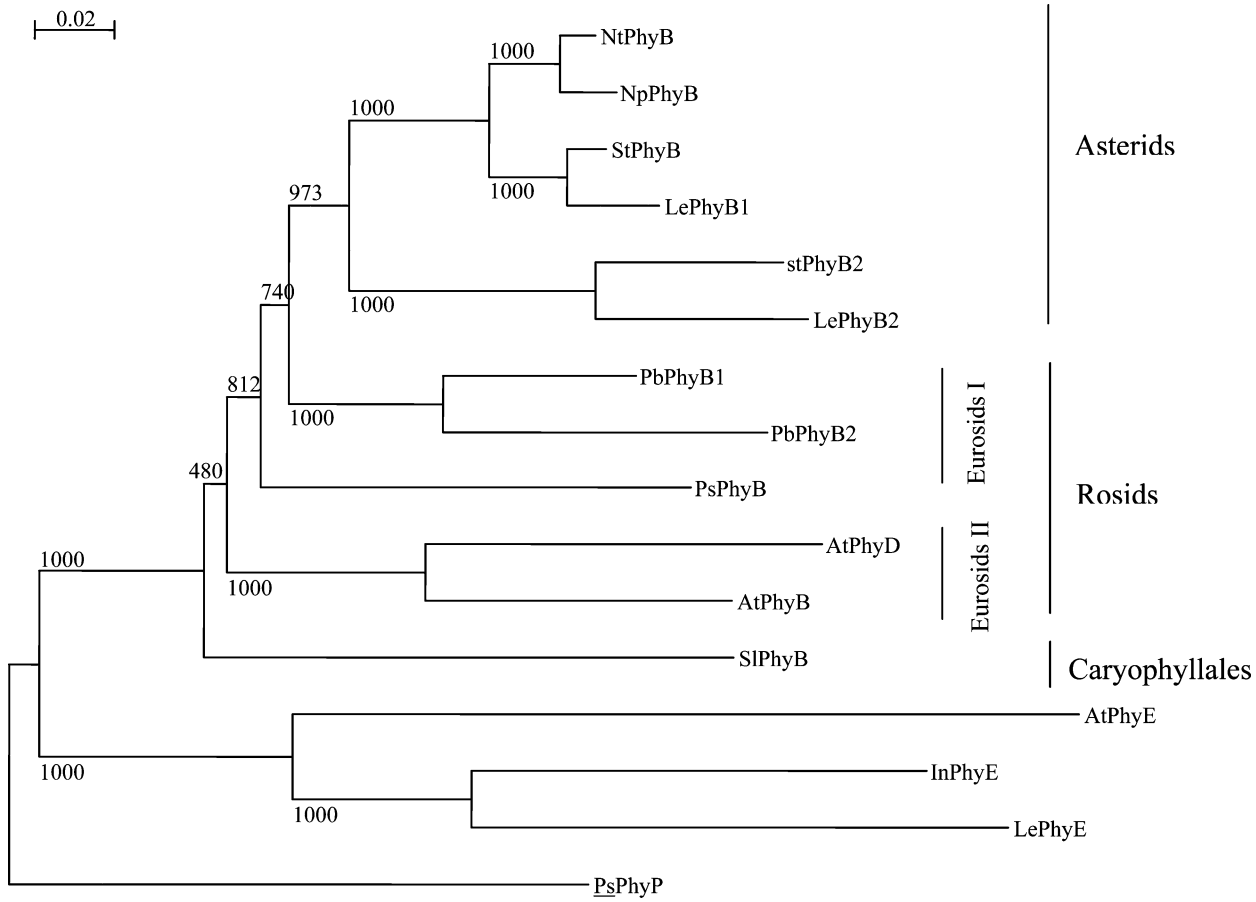
(Continued)

Supplemental Table. Continued.

Nicotiana plumbaginifolia			NpPhyB Y14676		
Nicotiana tabaccum			NtPhyB L10114		
Onoclea sensibilis		OsPhy3 BAC55267			
Oryza sativa	OsCry2 AB073547 OsCrya + b AB073547 OsCryDASH AK072287 OsPHR2 AK105611 OsCPD-photolyase AB096003 Os(6-4)PHR AP005915	OsPhot1 BAA84780 OsPhot2 BAA84779	OsPhyA X14172 OsPhyB X57563 OsPhyC AB018442	OsZTL2 AK100677 OsZTL1 AK111850 OsZTL3 AP005412 (Genomic)	OsPKS1 NM 196426 OsPKS2 TC271291 OsPKS3 TC272048
Physcomitrella patens	PpCrya + b AB027528, AB060693		PpPhy1 AY123146 PpPhy2 AY123147 PpPhy3 AY123148 PpPhy4 AY123145		
Pinus sylvestris	PsCry TC41388		PsPhyN AJ271627 PsPhyP X96738 PsPhyO AJ286245		
Pisum sativum		PsPhot1A, AAM15725 PsPhot1B AY295348 PsPK3 U11553	PsPhyB AF069305		
Populus balsamifera			PbPhyB1 AF309806 PbPhyB2 AF309807		
Rhyzobium leguminosarum			RlBphb AJ416905		
Solanum tuberosum			StPhyB Y14572 StPhyB2 NP005890, SGN-U256068, SGN-U256952, TC130366, SGN-U265279		StPKS1 BQ517558 StPKS2 SGN-U193421, BQ517559
Sorghum bicolor	SbCry1 TC93286	Sbphot2 TC98322 Sbphot1 BM325071 SbNPH TC108391 SbPK4 AA738536	SbPhyA U56729 SbPhyB AF369906 SbPhyC U56731	SbZTL1bis CD431693, SbZTL1 BE360254 BM326460	
Spinacia oleracea		SoPk X73298			
Stellaria longipes			S1PhyB AF544028		
Triticum aestivum	TaCry1 TC240073 TaCry2 TC249699				
Vicia faba		VfPhot1 AB095909 VfPhot2 AB095910			
Xanthomonas axonopodis			XaPhy-like AE011887		
Xenopus laevis	XlCryDASH AY049033				
Zea mays					ZmPKS1 CC630621(genomic)ZmPKS2, CC709502, CC668319



Supplemental Figure. Phylogenetic relationships of the PKS proteins from Angiosperms based on the protein sequence alignments. All branches are drawn to scale and the scale bar represent 0.05 substitution per nucleotide.



Supplemental Figure. Phylogenetic relationships of the phytochromes B/D/E from Eudicotyledon based on the protein sequence alignments. All branches are drawn to scale and the scale bar represent 0.02 substitution per nucleotide.