Adaptive Evolution in the Photosensory Domain of Phytochrome A in Early Angiosperms

Sarah Mathews, J. Gordon Burleigh,*[1] and Michael J. Donoghue†*

*Division of Biological Sciences, University of Missouri; and †Department of Ecology and Evolutionary Biology, Yale University

Flowering plant diversity now far exceeds the combined diversity of all other plant groups. Recently identified extant remnants of the earliest-diverging lines suggest that the first angiosperms may have lived in shady, disturbed, and moist understory habitats, and that the aquatic habit also arose early. This would have required the capacity to begin life in dimly lit environments. If so, evolution in light-sensing mechanisms may have been crucial to their success. The photoreceptor phytochrome A is unique among angiosperm phytochromes in its capacity to serve a transient role under conditions where an extremely high sensitivity is required. We present evidence of altered functional constraints between phytochrome A (*PHYA*) and its paralog, *PHYC*. Tests for selection suggest that an elevation in nonsynonymous rates resulted from an episode of selection along the branch leading to all angiosperm *PHYA* sequences. Most nucleotide sites (95%) are selectively constrained, and the ratio of nonsynonymous to synonymous substitutions on branches within the *PHYA* clade does not differ from the ratio on the branches in the *PHYC* clade. Thus, positive selection at a handful of sites, rather than relaxation of selective constraints, apparently has played a major role in the evolution of the photosensory domain of phytochrome A. The episode of selection occurred very early in the history of flowering plants, suggesting that innovation in phyA may have given the first angiosperms some adaptive advantage.

Introduction

The angiosperm crown group originated approximately 140 to 190 MYA (Sanderson and Doyle 2001) and fundamentally changed the structure and functioning of the terrestrial landscape. With the rise of the angiosperms, several lineages of plants that had been dominant since the Triassic declined in dominance or became extinct (Hickey and Doyle 1977; Doyle 1978; Knoll 1984; Friis, Chalconer, and Crane 1987; Lidgard and Crane 1988; Crane, Friis, and Pederson 1995; Lupia, Lidgard, and Crane 1999). Competing hypotheses have emphasized two divergent ecological settings and plant morphologies associated with the initial establishment of flowering plants (Doyle and Donoghue 1993). With respect to the light environment, the first angiosperms have been portrayed either as understory woody plants of wet tropical rain forests (Arber and Parkin 1907; Bessey 1915; Bews 1927; Axelrod 1952; Takhtajan 1969; Thorne 1974; Cronquist 1988) or as semi-woody to herbaceous colonizers of disturbed, open habitats (Stebbins 1974; Doyle and Hickey 1976; Hickey and Doyle 1977; Doyle and Donoghue 1986; Taylor and Hickey 1992, 1996; Wing and Boucher 1998).

Recent phylogenetic analyses have identified the extant remnants of the earliest-diverging angiosperm lineages, which include shrubs and woody vines that occur in shady habitats (Mathews and Donoghue 1999, 2000; Parkinson, Adams, and Palmer 1999; Qiu et al. 1999; Soltis, Soltis, and Chase 1999; Barkman et al. 2000; Doyle and Endress 2000; Graham et al. 2000; Savolainen et al. 2000; Soltis et al. 2000; Zanis et al. 2002); these are *Amborella* and a clade including the families Austrobaileyaceae, Illiciaceae, Schisandraceae, and Trimeniaceae. The water lilies also diverged early, but probably not before the lineage containing the woody understory species *Amborella trichopoda* of New Caledonia (Zanis et al. 2002). Implications for character evolution of the recently discovered fossil aquatic plant *Archaefructus liaoningensis* remain unclear in light of the low support for its placement in phylogenetic trees (Sun et al. 2002); however, similarities to *Cabomba* suggest that it may be a member of the water lily lineage (James Doyle, personal communication). Thus, the preponderance of phylogenetic evidence is consistent with the hypothesis of Feild et al. (2003) that the earliest angiosperms occupied shady, disturbed, and/or streamside habitats. In such an environment, innovations that enhanced their ability to establish in dimly lit environments may have been important to their early survival.

Photomorphogenesis, the control of plant form by ambient light conditions, is mediated by phytochrome photoreceptors, which absorb red light (R) and far-red light (FR), and by distinct blue- and ultraviolet-absorbing receptors (reviewed in Fankhauser and Chory 1997). Phytochrome evolution in land plants is marked by a series of gene duplications that have led to independently evolving and functionally distinct lines (Mathews and Sharrock 1997). A duplication preceding the origin of seed plants resulted in two distinct lines that persist in all extant seed plants. Phylogenetic analyses suggest that subsequent duplications occurred in each of these lines, leading to the four major forms found in angiosperms, phytochromes A, B, C, and E (phyA–C, E), encoded by *PHYA–C*, and *E* (Mathews, Lavin, and Sharrock 1995; Mathews and Sharrock 1997). *PHYA* and *PHYC* form one duplicate pair; *PHYB* and *PHYE* form a second duplicate pair. In cycads, *Ginkgo*, and conifers, a duplication in the *PHYA*/*C*-related line led to *PHYN* and *PHYO*, but the *PHYB*/*E*-related line, *PHYP*, did not diversify in other seed plants except for Pinaceae (Schneider-Poetsch et al. 1998; Clapham et al. 1999; Schmidt and Schneider-Poetsch 2002; Mathews and Donoghue, unpublished data). These phylogenetic analyses also suggest that the duplication leading to *PHYA* and *PHYC* occurred prior to the origin of angiosperms (Mathews, Lavin, and Sharrock 1995; Mathews and Sharrock 1997;

[1] Present address: Department of Computer Science, Iowa State University.

Key words: phytochrome A, positive selection, adaptive evolution, early angiosperms.

E-mail: mathewss@missouri.edu.

Mol. Biol. Evol. 20(7):1087–1097. 2003
DOI: 10.1093/molbev/msg123

Mathews and Donoghue 1999). *PHYA* and *PHYC* are about 50% identical (Sharrock and Quail 1989) and are functionally distinct (Casal, Sanchez, and Yanovsky 1997; Halliday, Thomas, and Whitelam 1997; Qin et al. 1997). The functional divergence of this gene pair is of interest in the context of early angiosperm history because phytochrome A (phyA) confers a competitive advantage for *Arabidopsis* seedlings establishing in dense shade (Yanovsky, Casal, and Whitelam 1995).

Phytochrome null mutants of *Arabidopsis*, tomato, and rice demonstrate that phyA induces seedling de-etiolation in response to continuous far-red light (FRc) via a far-red high irradiance response (FR-HIR; Nagatani, Reed, and Chory 1993; Parks and Quail 1993; Whitelam et al. 1993; van Tuinen et al. 1995; Takano et al. 2001). Etiolation is viewed as a specialized developmental pathway that is used to delay development in the dark or under low light conditions (e.g., McNellis and Deng 1995). Etiolated angiosperm seedlings require a light signal to de-etiolate, or to develop into a compact, green seedling with fully expanded seed leaves. Canopy shade is FR-enriched because chlorophyll strongly absorbs R; thus, in dense shade phyA induces de-etiolation. Conversely, open sunlight is R-enriched, and in these habitats, phyB induces de-etiolation. Additionally, phyA and phyB have contrasting antagonistic roles such that FRc suppresses phyB-mediated de-etiolation and Rc suppresses phyA-mediated de-etiolation (McCormac et al. 1992). The antagonistic role of phyA may best be understood in the context of the shade avoidance syndrome, controlled primarily by phyB, which induces stem elongation and accelerated development in response to neighbor detection (reviewed in Smith 2000). These responses may be too costly for seedlings, which have only limited resources (Smith, Xu, and Quail 1997). Subsequent to de-etiolation, shade avoidance responses come into play because phyA is lost rapidly in the light (Somers et al. 1991). It has been argued that seedlings emerging in the shade require a mechanism through which early shade avoidance reactions are suppressed, and furthermore, that the FR-HIR mediated by phyA provides such a mechanism (Smith, Xu, and Quail 1997). The importance of phyA in low light conditions is supported by the observation that *phyA* mutants of *Arabidopsis* die prematurely when grown in deep shade (Yanovsky, Casal, and Whitelam 1995). And while one *Arabidopsis* ecotype has a reduced sensitivity to FR (Maloof et al. 2001), the ability of this ecotype to survive in deep shade awaits further investigation.

Ginkgo and some conifers have elements of the FR-HIR de-etiolation response (Burgin et al. 1999; Christensen et al. 2002), but the response may be less important in conifer seedlings, many of which are compact and green (not etiolated) in the dark (Burgerstein 1900; Bogorad 1950; Mukai et al. 1992). Moreover, none of the seed plant phytochromes are unambiguous homologs of *PHYA* (e.g., Schmidt and Schneider-Poetsch 2002; Mathews and Donoghue, unpublished data), suggesting that the FR-HIRs of angiosperms and other seed plants may have independent origins. Phytochrome A also allows dark-imbibed seeds to germinate in response to millisecond pulses of broad spectrum light via the very low fluence response (VLFR; Botto et al. 1996), a response that is unknown outside of angiosperms and that operates via a downstream pathway different from that of the FR-HIR (Yanovsky, Casal, and Luppi 1997; Cerdán et al. 1999; Casal 2000; Yanovsky et al. 2002). The VLF germination response may be important for germination in open environments, allowing germination in response to brief soil disturbances (Casal, Sanchez, and Yanovsky 1997), and/or it may be important in shady, disturbed sites where light intensities fluctuate. Evidence from analyses of spectral characteristics indicates that phyC is likely to differ from phyA in its photoresponse to shade (Eichenberg et al. 2000). *PHYC* transcripts occur in low abundance in *Arabidopsis* (Sharrock and Quail 1989) and rice (Basu et al. 2000), and they may be constitutively expressed (Basu et al. 2000). Although little is known about the function of phyC, overexpression of *Arabidopsis PHYC* in tobacco suggests that it is involved in expansion of cotyledons in response to red light (Halliday, Thomas, and Whitelam 1997; Qin et al. 1997) and thus may be more important in open habitats. Thus, if the first angiosperms were understory colonizers, and if innovations in phytochrome function were important to their success, it is the evolution of phyA function that might have been important.

Gene duplications are considered to be a significant force in genome evolution (e.g., Wagner 2001) and may also play a significant role in speciation (Lynch and Conery 2000). When a gene duplicates, one copy may be silenced or evolve a novel function, or the two copies may subdivide functions of the ancestral gene (e.g., Ohno 1970; Walsh 1995; Force et al. 1999; Hughes 1999; Lynch and Force 2000). In the case of *PHYA* and *PHYC*, both copies have been maintained and they encode different functions. Moreover, as noted above, at least some of the functions of phyA are unique to angiosperms, suggesting that functional diversification has involved processes other than, or in addition to, subdivision of ancestral functions. Thus, two alternative models of functional diversification following gene duplication might explain the divergence of phyA from phyC. First, the function of one copy might be selectively maintained while the other copy evolves without selective constraints, obtaining a new function (Ohno 1970; Dykhuizen and Hartl 1980). This model is supported by evidence that nucleotide substitution rates in phyA differ significantly from those in phyC (Mathews and Sharrock 1997; Alba et al. 2000). Second, one copy might evolve a new function via positive selection. This model would be supported by evidence that rates of nonsynonymous substitution exceed rates of synonymous substitution (Hill and Hastie 1987). Yang and Nielsen (2002) report evidence of positive selection along a branch separating phyA and phyC sequences from species that are nested well within the angiosperms. Thus, the divergence of phyA from phyC may have involved both the relaxation of selective constraints and positive selection.

If photoreceptor evolution played a role in the initial establishment of angiosperms, we would expect to find evidence of altered selective constraints in sequences from the earliest diverging angiosperms. Previously published analyses could not address this question because no early-

diverging species were sampled (Yang and Nielsen 2002). In this analysis we have explored a published data set that sampled the photosensory domain sequences of phyA and phyC from species representing all of the early-diverging lines of angiosperms (Mathews and Donoghue 2000) to determine if evolutionary constraints on phyA and phyC were altered early in the history of angiosperms, or if they were under positive selection. A further test of the hypothesis that the evolution of phyA function was important early in the history of angiosperms is not addressed in this study. It will require characterization of phyA responses in early-diverging angiosperms to determine whether they have the same responses as those described from derived eudicots (e.g., *Arabidopsis* and tomato) and a derived monocot (rice).

Materials and Methods

Taxonomic Sampling

We used a subset of the *PHYA* and *PHYC* sequences analyzed by Mathews and Donoghue (2000) to test for functional divergence and to perform tests for selection. As noted above, Yang and Nielsen (2002) have reported evidence of positive selection on the branch separating clades of full-length (~4 kb) *PHYA* from *PHYC*. Their use of full-length sequences restricted their sampling to clades well nested within the angiosperms, and they included no outgroup sequences. Thus, they could not precisely determine the branch on which selection had occurred. The data set of Mathews and Donoghue (2000) included photosensory domain sequences (~1.2 kb) from representatives of each of the early diverging lineages, as well as representatives of eudicot and monocot clades nested within angiosperms. We did not include the singletons (only *PHYA* or *PHYC* sequenced) from their data set and we included the published *PHYA*/*PHYC*-related sequence from *Picea PHYO* (U60264) in order to root the phylogeny. We conducted a Bayesian analysis of these data using MrBayes version 2.01 (Huelsenbeck and Ronquist 2001). These analyses used the general reversible model (e.g., Tavaré 1986) with rate variation among nucleotides following a discrete gamma distribution with four rate categories (Yang 1994), and they assumed a proportion of invariant sites estimated from the data. The default priors were used for all parameters. The Markov chain Monte Carlo (MCMC) search consisted of four chains, three of which were heated to a temperature of 0.2 (Huelsenbeck and Ronquist 2001). The Markov chain started with a random tree and was sampled once every 100 generations for 1 million generations. The posterior probability was then calculated from a consensus tree of all trees sampled after the Markov chain reached stationarity, which we estimated to be at 50,000 generations. The majority rule consensus of 9,501 trees inferred in Bayesian analysis is depicted in figure 1. Relationships that are strongly supported in this tree are the same as those that were strongly supported in parsimony analysis (Mathews and Donoghue 2000). Our tests for functional divergence (Gu 1999, 2001) were based on a tree that included sequences from 36 of the 45 species, excluding only those sequences that were incomplete at their 5′ or 3′ ends. Our tests for variable selection among lineages and sites (Yang and Nielsen 2002) were based on the tree of all 45 species.

Character Sampling

The apoprotein of eukaryotic phytochromes comprises approximately 1,100 to 1,200 amino acid residues. Light perception by phytochromes requires the covalent attachment of a plastid-derived linear-tetrapyrrole chromophore. Double-bond photoisomerization of the chromophore leads to photoreversible changes in the conformation and activity of the protein. Thus, a hallmark of phytochromes is their interconversion between the red light–absorbing form (Pr) and the far-red light–absorbing form (Pfr; Kendrick and Kronenberg 1994). A general model of phytochrome function postulates that absorption of R by Pfr induces biological activity and that absorption of FR by Pfr inhibits activity (e.g., Quail 1991). Shinomura, Uchida, and Furuya (2000) showed that the phyA-mediated VLFR fits this model because it depends on photoconversion from Pr to Pfr. However, they found that the phyA-mediated FR-HIR departs from this model because it depends on the conversion of Pfr to Pr (Shinomura, Uchida, and Furuya 2000).

The 3-dimensional structure of phytochromes remains undetermined, but functional domains have been identified through biochemical, genetic, and bioinformatics approaches (reviewed in Montgomery and Lagarias 2002). At a broad level, the molecule can be divided in half, with photosensory activity attributed to the n-terminal half and the regulatory activity attributed to the c-terminal half. Within each half, domains that confer specific functions have been identified. The photosensory domain comprises four subdomains (P1–P4) necessary for photosensory specificity. Three of these, P2–P4, are conserved among eukaryotic and prokaryotic phytochromes. The regulatory domain comprises two subdomains necessary for transducing light signals, a PAS-related subdomain and a histidine kinase–related subdomain (Montgomery and Lagarias 2002). Our sequences correspond to photosensory subdomains P3 and P4. Based on sequence homology, the 3-dimensional structure of these individual subdomains is inferred to be similar to that of the of the small yeast protein YLK9 (Montgomery and Lagarias 2002). The phytochrome chromophore is covalently bound to a conserved cysteine in the pocket formed by the P3 subdomain.

Rate Tests

We estimated the coefficient of functional divergence (θ) in order to test the hypothesis that functional constraints on *PHYA* and *PHYC* did not change after gene duplication (Gu 1999, 2001) by using DIVERGE version 1.04 (Gu 1999). DIVERGE uses the amino acid substitution rate at a site as an indication of positive selection on that site, and it examines whether rapidly evolving sites in one paralog are rapidly evolving in the other paralog. A shift in the distribution of rates among sites between the two paralogs may indicate functional diversification. In this test, theta (θ) measures the coefficient of rate correlation over sites between paralogs. Assuming a Poisson model

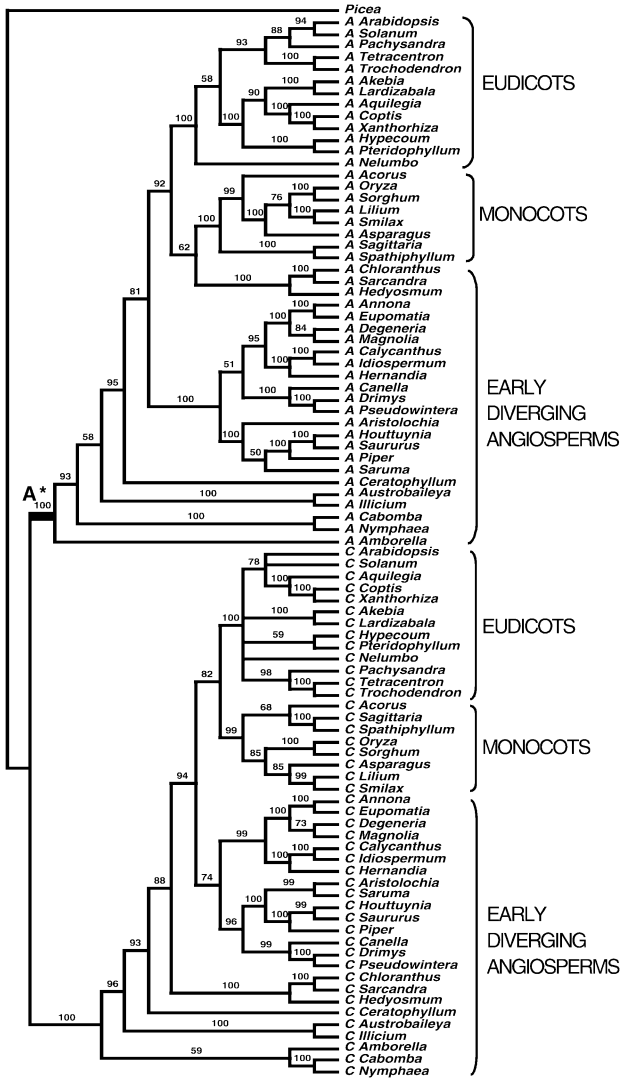

FIG. 1.—One of three trees with the best likelihood (lnL = –44453.46) from the Bayesian analyses of the phytochrome A and C data conducted with MrBayes version 2.01 (Huelsenbeck and Ronquist 2001). Posterior probabilities are above the branches and were calculated from a consensus tree of all trees sampled after the Markov chain reached stationarity, which we estimated to be at 50,000 generations. Branch A*, leading to all *PHYA*, is the foreground branch for detecting positive selection. GenBank accession numbers and voucher information for angiosperm *PHY* sequences are available in Mathews and Donoghue (2000). The *Picea* sequence used is GenBank accession U60264.

of amino acid substitution, the Gu test (1999, 2001) estimates subtree likelihoods for a given tree. In this case, the tree was similar to the tree in figure 1, pruned as described above (see T*axonomic Sampling*). After obtaining maximum likelihood (ML) estimates of branch lengths, the gamma shape parameter (α) of rate variation across sites, and the coefficient of functional divergence (θ_{AC}), a likelihood ratio test (LRT) can be constructed to test the null hypothesis that $\theta_{AC} = 0$ (with the alternative hypothesis being that $\theta_{AC} > 0$). Rejection of the null hypothesis provides statistical evidence that functional constraints shifted after gene duplication (Gu 1999, 2001).

Tests for Selection

We compared the rates of synonymous (silent; d_S) and nonsynonymous (amino-acid replacement; d_N) substitutions in *PHYA* and *PHYC* to distinguish whether or not positive selection had a role in functional divergence. The d_N/d_S ratio measures selective pressure on amino acids. If functional divergence occurs by neutral mutations that are fixed by random genetic drift (Ohno 1970), or by neutral mutations that preadapt it for a new function (Dykhuizen and Hartl 1980), d_N will not exceed d_S. If functional divergence occurs by positive selection, d_N may significantly exceed d_S. Thus, the ratio of d_N/d_S (ω) under the three conditions will vary. Under the first two conditions, $\omega < 1$; relaxing selective constraints should elevate the nonsynonymous substitution rate but the d_N/d_S ratio should never exceed the neutral expectation that $\omega = 1$. However, ω may exceed one if an amino acid change is selectively advantageous and it is fixed at a higher rate than a synonymous mutation. Thus, finding that ω is significantly greater than one provides convincing evidence of positive selection (Hill and Hastie 1987).

We used a tree-based likelihood ratio approach (e.g., Yang 1998), implemented in PAML (Yang 1997), to test for variable selective pressures among branches in the phylogeny of *PHYA* and *PHYC*. We compared the likelihoods of the one-ratio model, which assumes the same d_N/d_S ratio for each branch in the tree (table 1, model 1A), with two-ratio models, which assume two d_N/d_S ratios. One of the two-ratio models assumes that all *PHYA* branches have one ratio and that the *PHYC* and outgroup branches have another (table 1, model 2A), while the other assumes that the branch to all *PHYA* has a different ratio than all the other branches in the phylogeny (table 1, model 3A). We also compared the two-ratio models with a three-ratio model (table 1, model 4A), which assumes one d_N/d_S ratio for the branch to all *PHYA*, one for the internal *PHYA* branches, and one for all the rest. Likelihood estimation assumed the codon substitution model of Goldman and Yang (1994), which estimates separate nucleotide frequencies for each codon position and includes parameters for the transition/tranversion ratio and the d_N/d_S ratio.

The lineage-specific models just described assume that there is no variation in ω among sites. Thus, they will detect selection in a lineage only if the average d_N over all sites is greater than the average d_S. Conversely, site-specific models (Yang et al. 2000; Yang, Swanson, and Vacquier 2000) detect selection only if the average d_N over all lineages is greater than the average d_S. Thus, we also used likelihood ratio tests to compare site-specific models with models that allow the d_N/d_S ratio to vary both among sites and among lineages (Yang and Nielsen 2002). The branch-site models assume that each amino acid site falls into one of four classes of d_N/d_S ratio. Sites that are highly conserved across all branches comprise a class having a small d_N/d_S ratio (ω_0), whereas sites that are neutral or weakly constrained across all branches comprise a class having a ratio near or below 1 (ω_1). Sites that are highly conserved on background branches (ω_0) but that have a d_N/d_S ratio that may be greater than one (ω_2) on the branch being tested for selection—the foreground

Table 1
Parameter Estimates from Tests for Selection

Model	*p*	*l*	Parameter Estimates	Positively Selected Sites
Branch-Specific Models				
1A: One-ratio	1	−43610.20	$\omega = 0.067$	N/A
2A: Two-ratio[a]	2	−43610.20	$\omega_0 = 0.067$, $\omega_{AA}* = 0.067$	N/A
3A: Two-ratio[b]	2	−43607.39	$\omega_{0A} = 0.068$, $\omega_A* = 0.165$	N/A
4A: Three-ratio[c]	3	−43607.38	$\omega_0 = 0.067$, $\omega_A = 0.067$ $\omega_A* = 0.165$	N/A
Site-Specific Models				
M1: neutral	1	−49926.24	$p_0 = 0.167$, ($p_1 = 0.833$)	Not allowed
M3: discrete ($K = 2$)	3	−42521.82	$p_0 = 0.732$, ($p_1 = 0.268$) $\omega_0 = 0.026$, $\omega_1 = 0.197$	None
Branch-Site Models				
Model A	3	−49866.40	$p_0 = 0.027$, $p_1 = 0.134$ ($p_2 + p_3 = 0.840$), $\omega_2 = 0.000$	
Model B	3	−42516.21	$p_0 = 0.702$, $p_1 = 0.252$ ($p_2 + p_3 =$ **0.046**) $\omega_0 = 0.026$, $\omega_1 = 0.198$ $\omega_2 =$ **9.288**	***P* > 0.9:** Q118[d], P119D, T152V ***P* ≥ 0.7:** L31R, P106K, A135L ***P* > 0.5:** S88A, T252V, F309M, K319S, V362Y

NOTE.—*p* is the number of free parameters for estimation of ω; *l* is the log likelihood of the data given the model; *P* is the posterior probability of a site being under selection; *K* is the number of site classes. Parameters that indicate positive selection are in boldface; parameters in parentheses are not free parameters. Sites potentially under selection are identified using *Picea PHYO* (U60264) as the reference sequence. Substitutions at 10 of the 15 sites that are predicted to be under positive selection unambiguously occur on the foreground branch based on parsimony character mapping (see text), and their state changes on this branch are noted. Estimates of the transition/transversion ratios are 3.2 in the M1 and Model A analyses and 2.2 in all other analyses.

[a] Model 2A assumes one ω for all branches outside the PHYA clade (ω_0) and one ω for all PHYA branches ($\omega_{AA}*$).

[b] Model 3A assumes one ω for all branches (ω_{0A}) except for the branch to all PHYA (ω_A*).

[c] Model 4A assumes one ω for all branches outside the PHYA clade (ω_0), one ω for the branch to all PHYA (ω_A*), and one ω for all the branches within the PHYA clade (ω_A).

[d] The amino acid at position 118 is variable within the PHYA clade.

branch—comprise a third class. Sites that are neutral or weakly constrained (ω_1) on the background branches but that may have a ratio greater than one (ω_2) on the foreground branch comprise a fourth class. Model A assumes that $\omega_0 = 0$ and that $\omega_1 = 1$, and model B allows ω_0 and ω_1 to vary. Both models estimate ω_2 from the data. LRTs compare models A and B with site-specific models M1 and M3, respectively. Models M1 and M3 are identical to models A and B, respectively, except that they assume that ω_0 and ω_1 are the same across all branches of the phylogeny. If $\omega_2 > 1$, and if either model A or B fits the data significantly better than model M1 or model M3, then these LRTs constitute a test of positive selection in a subset of the amino acids on the foreground branch. The specific sites that change along the foreground branch are identified using an empirical Bayes approach, which estimates the posterior probabilities that a site falls into a particular site class (Nielsen and Yang 1998).

Character Mapping

Amino acid substitutions that result in a new function may occur once, early in the history of a protein lineage, to character states that are conserved through subsequent speciation events. Tests for selection using branch-specific or site-specific models may not identify these sites, because they are most suitable for detecting recurrent diversifying selection (e.g., Anisimova, Bielawski, and Yang 2002). Similarly, the prediction of sites of functional specification using DIVERGE (Gu 1999) identifies just those sites that have changed often through time (Gu 2001), although sites that change infrequently are included in the estimation of θ (Gu 1999, 2001). To identify the full set of the amino acid substitutions that occurred on the branches to all PHYA and all PHYC, we translated the nucleotide alignment of Mathews and Donoghue (2000) and mapped all unambiguous changes onto the branches of their phylogeny by using MacClade (Maddison and Maddison 2000). The alignment is available at TreeBase (www.treebase.org) under accession number SN543. This data set includes singletons that we did not include in our pairwise comparisons and comprises a total of 100 *PHYA* and *PHYC* sequences to which we added the sequence from *Picea*. The species phylogeny inferred by Mathews and Donoghue (2000) in their duplicate gene rooting analysis unambiguously identifies the earliest diverging branches of the angiosperm phylogeny to be the same as those converged upon in many other analyses (Mathews and Donoghue 1999; Parkinson, Adams, and Palmer 1999; Qiu et al. 1999; Soltis, Soltis, and Chase 1999; Barkman et al. 2000; Doyle and Endress 2000; Graham et al. 2000; Savolainen et al. 2000; Soltis et al. 2000; Zanis et al. 2002). We used parsimony to infer ancestral states at the root node, because our focus was on sites that change just once or very rarely on the tree—that is, those that are conserved in the outgroups and in one gene copy but change to a conserved state on the branch to the other copy. When the focus is on rare changes, parsimony and maximum likelihood often yield the same results (e.g., Schluter et al. 1997; Cunningham, Omland, and Oakley 1998). To determine if this was the case in our study, we compared the ancestral states inferred using parsimony with those inferred using the maximum likelihood approach of Yang, Kumar, and Nei (1995) that is implement-

ed in PAML (Yang 1997). We further tested the results from the mapping exercise by adding some unpublished data (Mathews and Donoghue) from additional seed plant phytochromes to the alignment described above. In addition to the *PHYO* from *Picea*, a conifer, we included *PHYN* from *Pinus*, also a conifer, and we included both *PHYO* and *PHYN* from *Cycas*, a cycad, and from *Ginkgo*. As noted above, *PHYO* and *PHYN* of other seed plants are related to *PHYA* and *PHYC* of angiosperms, whereas *PHYP* from other seed plants is related to *PHYB* and *PHYE* of angiosperms.

Results

Rate Tests

Estimation of θ using DIVERGE (Gu 1999) strongly indicates that rates of amino acid substitution of *PHYA* and *PHYC* are not correlated. The null hypothesis that $\theta_{AC} = 0$ is strongly rejected ($2\Delta l = 78.82$, $p < 0.001$, $df = 1$). This suggests that there was a shift in the distribution of substitution rates among sites between the paralogs, and it is consistent with their functional diversification, either through selection or relaxation of selective constraints.

Tests for Selection

Tests for selection using branch-specific models indicate that the d_N/d_S ratio varies significantly between the branch to all *PHYA* and all other branches in the phylogeny of *PHYA* and *PHYC*. Just one of the two-ratio models fits the data significantly better than the one-ratio model. This is the model that assumes one ratio for the branch to all *PHYA* and one ratio for all other branches in the phylogeny (table 1, model 3A). The test statistic is $2\Delta l = 5.63$; $P = 0.02$. The two-ratio model that assumes one ratio for all *PHYA* branches (table 1, model 2A) does not fit the data better than the one-ratio model. Moreover, the three-ratio model (table 1, model 4A) does not fit the data significantly better than two-ratio model 3A ($2\Delta l = 0.02$; $P = 0.99$). Together these results indicate that all the variation in the d_N/d_S ratio can be explained by separating the branch to all *PHYA* from all other branches. They also suggest that the d_N/d_S ratio on the internal *PHYA* branches is not significantly different from the ratio on the *PHYC* branches. Thus, it does not appear that there was a general relaxation of constraints on *PHYA*.

Tests for selection using branch-site models identify sites under positive selection along the branch to all *PHYA*. Parameter estimates under model A suggest that 16% of amino acid sites are highly conserved across all lineages with $\omega_0 = 0$, and 79% are nearly neutral, with $\omega_1 = 1$, while a handful of sites, 5%, are under strong selection along the branch to all *PHYA*. Model A fits the data significantly better than model M1 ($2\Delta l = 119.68$; $P < 0.001$ with $df = 2$). Parameter estimates under model B (table 1) suggest that about 5% of the sites are under strong positive selection ($\omega_2 = 9.29$). In multiple analyses under the assumptions of model B, the ω_2 values ranged from 9.29 to 17.53, indicating the computational difficulty of the analysis (Z. Yang, personal communication). In the table we report the value corresponding to the tree with the highest likelihood. Model B fits the data significantly better than the site-specific model M3 ($2\Delta l = 11.22$; $P = 0.004$ with $df = 2$). Regardless of the ω_2 value, at $P > 0.5$, model B identifies eleven sites potentially under selection along the branch to all *PHYA* (table 1). This is approximately one third of the sites inferred by character mapping to have changed along this branch (see below).

Character Mapping

We found that 32 amino acid substitutions map unambiguously to the branch leading to all PHYA in the tree with just *Picea* PHYO as the outgroup, while just seven substitutions map unambiguously to the branch leading to all PHYC. The same ancestral states at these sites are inferred by parsimony using MacClade (Maddison and Maddison 2000) and by maximum likelihood using PAML (Yang 1997). Of the substitutions occurring along the branch to all PHYA, 12 involve changes in charge, and three of these are at sites where changes within the PHYA clade do not alter the charge. Of the substitutions occurring along the branch leading to all PHYC, three involve changes in charge, and one of these is at a site where changes within the PHYC clade do not alter the charge. None of these sites is inferred to have been under selection. When PHYN and additional PHYO sequences representing other seed plant groups are added to the base of the tree, the number of substitutions that map to the foreground branch decreases to 30 while the number and identity of the sites that change on the branch to all PHYC remains the same. The sites that no longer map to the foreground branch are not inferred to have been under selection.

Our alignment corresponds to subdomains P3 and P4 of the phytochrome photosensory domain. The structure of these domains is inferred to be similar to that of the yeast protein YLK9 (Montgomery and Lagarias 2002). The chromophore is bound to a conserved cysteine residue in P3. Substitutions that involve charged residues are of particular interest because they might influence interactions between the protein and chromophore and/or between protein domains (Wu and Lagarias 2000). Three of the amino acid substitutions that occur along the branch to all PHYA result in a change in charge that is not altered within the PHYA clade. At position 119, the change is from an uncharged proline to a potentially negatively charged aspartic acid or glutamic acid residue very near the point of chromophore attachment. This site is inferred to have been under positive selection ($P = 0.98$). The position lies within a predicted loop between the third and fourth beta strands of the P3 domain; this likely is an important region for modulating the entry of the chromophore, its dynamics, its protonation state and/or its spectrum. At position 188, the change is from an uncharged glutamine to a positively charged histidine residue within a predicted alpha helix of the P3 domain that might serve as a connector to other domains. At position 360, the change is from a negatively charged glutamic acid residue shared by PHYC and the outgroup sequences to a potentially uncharged lysine or a neutral uncharged residue in all PHYA. This site is variable in PHYA but is never a negatively charged residue. It lies within a predicted alpha helix of the P4

domain that may serve as a connector to other domains. The single amino acid substitution along the branch to all PHYC that results in a change in charge replaces an uncharged glutamine at position nine with a positively charged arginine or lysine residue that is not altered within the PHYC clade. The position might lie in the terminal alpha helix in the P2 domain that may serve as a connector to other domains. The substitutions at each of these four sites may have functional implications, although just one is inferred to have been at a positively selected site (P119, $P = 0.98$). Thus, character mapping may provide insight into sites of potential functional significance that is not gained from tests for selection and functional divergence.

Discussion

Our results support the hypothesis that the photosensory domain of phytochrome A underwent adaptive evolution early in the history of flowering plants. The tests for functional divergence provide strong evidence that rates of substitution in *PHYA* and in its sister gene *PHYC* are not correlated. Tests for selection suggest that an elevation in nonsynonymous rates results from an episode of selection along the branch leading to all angiosperm *PHYA* sequences. The d_N/d_S ratio on the internal *PHYA* branches is not significantly different from the ratio on the *PHYC* branches. Moreover, most sites (95%) are highly conserved or evolving neutrally (table 1). Thus, positive selection at a handful of sites, rather than relaxation of selective constraints, apparently has played a major role in the evolution of the photosensory domain of phyA.

Strong evidence of positive selection has been reported mostly from analyses of recently diverged genes (Yang and Bielawski 2000), from analyses of genes that are expected to be under constant pressure to diversify, such as human immunodeficiency virus (HIV) and human influenza virus (Nielsen and Yang 1998; Crandall et al. 1999; Yang et al. 2000; Yang, Swanson, and Vacquier 2000), and from genes sampled from closely related species (Yang, Swanson, and Vacquier 2000). In plants, such evidence has been detected in disease-resistance genes (Bishop, Dean, and Mitchell-Olds 2000; Bergelson et al. 2001), where frequent diversification likely is advantageous. Evidence of positive selection may be much more difficult to obtain when adaptive change occurs early and is fixed, especially becausee adaptive change may result from evolution at a single site or at just a handful of sites (e.g., Perutz 1983; Hall 1984; Wilks et al. 1988; Chen, Greer, and Dean 1995, 1996; Newcomb et al. 1997). The nature of the challenge was emphasized by Golding and Dean (1998), who noted the exceeding difficulty of finding the few substitutions of large effect among the many other substitutions that have occurred in genes that diverged millions of years ago (Golding and Dean 1998:364). When adaptive evolution involves only a handful of amino acid sites, tests for selection that average rates across all sites of a locus may not detect evidence of positive selection (Endo, Ikeo, and Gojobori 1996; Crandall et al. 1999; Anisimova, Bielawski, and Yang 2001). Conversely, tests that allow the d_N/d_S ratio to vary among codons assume the same ratio on all branches of a phylogeny and thus may not detect episodic selection (Yang, Swanson, and Vacquier 2000).

Yang and Nielsen (2002) have addressed this in their branch-sites models, which show greater sensitivity to episodic selection. In the phytochrome data set that they analyzed they found no evidence of selection using either branch-specific or site-specific models alone, but they did find evidence of selection using branch-site models (Yang and Nielsen 2002), a finding similar to ours. However, their phylogeny of full-length *PHYA* and *PHYC* included only monocots (grasses) and eudiocots nested well within angiosperms. No outgroup was included and sites potentially under selection were identified with reference to the *PHYA* sequence from maize, a derived monocot (in figure 1 it would occur in the clade with the *PHYA* from other grasses, *Oryza* and *Sorghum*). Thus, although their analysis sampled more characters of the phytochrome molecule, they could not determine when in the history of the angiosperms adaptive evolution occurred; nor could they determine if selection occurred on the branch to *PHYA* or the branch to *PHYC*.

Episodic selection on *PHYA* very early in the history of flowering plants suggests that innovation in this photoreceptor may have given the first angiosperms some adaptive advantage. Among angiosperm phytochromes, phyA is unique in its capacity to serve a transient role under conditions where an extremely high sensitivity is required (Furuya and Schäfer 1996). This results from the fact that phyA is abundant in dark-grown seedlings but disappears rapidly in the light, due both to degradation and down-regulation (Somers et al. 1991), and from its role as mediator of very low fluence responses (VLFRs), which require exposure to only millisecond pulses of broad spectrum light (Botto et al. 1996). It also is unique in that the far-red high irradiance response (FR-HIR) depends on the conversion of Pfr to Pr induced by the absorption of continuous far-red light (FRc; Shinomura, Uchida, and Furuya 2000). Although angiosperm phyE is known to induce *Arabidopsis* seeds to germinate in response to FRc, seed germination in many species is inhibited by FRc (Hennig et al. 2002), and the primary activities of phyE and of the remaining angiosperm phytochromes are dependent primarily on the conversion of Pr to Pfr, induced by absorption of either continuous or pulsed red light.

Phylogenetic and ecophysiological evidence suggests that the earliest angiosperms may have occupied dimly lit, disturbed, and moist understory habitats, but that the aquatic habit, which requires the capacity to begin life underwater, also arose early (Mathews and Donoghue 1999, 2000; Parkinson, Adams, and Palmer 1999; Qiu et al. 1999; Soltis, Soltis, and Chase 1999; Barkman et al. 2000; Doyle and Endress 2000; Graham et al. 2000; Savolainen et al. 2000; Soltis et al. 2000; Zanis et al. 2002; Feild, Arens, and Dawson 2003). We speculate that having an extremely sensitive photoreceptor capable of inducing responses with minimal exposure to light might have conferred an adaptive advantage in these habitats, and that this is reflected in the signal of molecular adaptation in phyA along the branch leading to all *PHYA*.

It is important to use two approaches to test this hypothesis. First, phytochrome-mediated responses in

early-diverging flowering plants and additional outgroups need to be characterized. If our hypothesis were correct, we would expect extant remnants of the earliest-diverging lineages to show the full range of responses that are attributed to phyA in model species. Similarly, we would expect a full expression of these responses to be lacking in other seed plants. Both of these expectations need to be tested. We know that *phyA* null mutants of *Arabidopsis* fail to de-etiolate and then die prematurely when grown in deep shade (Yanovsky, Casal, and Whitelam 1995), possibly as a consequence of the inappropriate expression of shade avoidance responses (Smith, Xu, and Quail 1997). But phytochrome-mediated responses of early-diverging angiosperms are uncharacterized, and there are only limited data from other seed plants. Some members of the conifer family Pinaceae have a rudimentary FR-HIR (Burgin et al. 1999). However, the degree of etiolation, and thus the need to de-etiolate, varies among conifers; many are green in the dark (Burgerstein 1900; Bogorad 1950; Mukai et al. 1992). *Gingko* etiolates (Burgerstein 1900; Chinn and Silverthorne 1993) and also shows rudimentary elements of the FR-HIR de-etiolation response (Christensen et al. 2002). Some cycads (*Cycas* and *Zamia*) and some Gnetales (*Ephedra* and *Gnetum*) also etiolate (Bower 1882; Burgerstein 1900), but their ability to respond to FRc needs to be investigated. The VLFR also contributes to de-etiolation under low light conditions, although perhaps not enough to confer an advantage (Smith, Xu, and Quail 1997). No evidence has been presented that plants other than angiosperms have VLFRs, and we do not know if they occur in early-diverging angiosperms. It has been suggested that the very low fluence germination response is most important in open environments because seeds of many angiosperms germinate well in darkness, especially large seeds with adequate resources that are produced by many shade-tolerant species (Smith 1995). Small seeds, such as those that characterize early-diverging flowering plants (Feild, Arens, and Dawson 2003), however, are known to be viable in shady but disturbed habitats (Metcalfe and Grubb 1997; Grubb 1998), where the ability to respond to brief pulses of light may be important.

Second, in addition to characterizing phytochrome-mediated responses in early-diverging flowering plants and in other seed plants, mutagenesis experiments should be used to compare phytochrome-mediated responses of wild-type *Arabidopsis* with those of *Arabidopsis phyA* null mutants bearing sequence constructs with the ancestral states at the sites identified in our analyses as being of potential functional significance. If there is a connection between these sites and the known functions of phyA, altering them should also alter one or more of these functions. The fitness of *phyA* null mutants carrying these constructs can be determined to test our hypothesis that one or more of these functions could have conferred an adaptive advantage. The fitness of shade-avoidance phenotypes has been similarly tested using *phyB* mutants (Schmitt, McCormac, and Smith 1995; Pigliucci and Schmitt 1999). An important complementary test remains to be devised. Although it is possible to place a modified *PHYA* sequence in *Arabidopsis*, its ecology differs greatly from species like *Amborella trichopoda*. One is a weedy herb of open habitats; the other, a woody shade-tolerant shrub. It would be desirable to test the constructs in *Amborella*, which is not currently possible, or to find a way to test the fitness of phyA-mediated responses of woody shade-tolerant plants by somehow altering phyA expression. For example, it might be possible to test the far-red high irradiance de-etiolation response by using a brief light exposure to reduce the phyA pool before subjecting the emerging seedlings to continuous FR.

Many of the phytochrome mutants so far isolated have identified amino acid residues that are conserved, and possibly important, in all phytochromes (e.g., Quail et al. 1995; Maloof et al. 2001; Yanovsky et al. 2002). For example, one natural accession of *Arabidopsis thaliana* shows decreased sensitivity to FR, and its PHYA differs by a single amino acid from that of the lab strain Col-O. When this site was altered in PHYB, sensitivity to low-fluence R was also reduced, suggesting that the site is generally important for modulating light response in different phytochrome family members (Maloof et al. 2001). Similarly, an amino acid residue that is necessary for the phyA-mediated FR-HIR (Yanovsky et al. 2002) is also necessary for phyB-mediated response to continuous red light (Wagner and Quail 1995), suggesting that this site is important in different phytochromes. Our approach has identified amino acid residues that are uniquely altered in phyA. Manipulation of these sites has the potential to provide novel insight into the function of phyA.

Acknowledgments

We thank Taylor Feild for discussions on the ecophysiology of early-diverging angiosperms; Harry Smith for comments on our ideas about the importance of phyA; J. Clark Lagarias for sharing insight into the relevance of specific amino acid substitutions; Ziheng Yang for help with the branch-sites tests; and Bob Sharrock, Mannie Liscum, Keith Crandall, and three anonymous reviewers for helpful comments on drafts of this manuscript. Support for this work was provided by National Science Foundation grant DEB-9806397 to M.J.D. and S.M.

Literature Cited

Alba, R., P. M. Kelmenson, M. M. Cordonnier-Pratt, and L. H. Pratt. 2000. The phytochrome gene family in tomato and the rapid differential evolution of this family in angiosperms. Mol. Biol. Evol. **17**:362–373.

Anisimova, M., J. P. Bielawski, and Z. Yang. 2001. Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. Mol. Biol. Evol. **18**:1585–1592.

———. 2002. Accuracy and power of Bayes prediction of amino acid sites under positive selection. Mol. Biol. Evol. **19**:950–958.

Arber, E. A. N., and J. Parkin. 1907. On the origin of angiosperms. Bot. J. Linn. Soc. **38**:29–80.

Axelrod, D. I. 1952. A theory of angiosperm evolution. Evolution **6**:29–60.

Barkman, T. J., G. Chenery, J. R. Mcneal, J. Lyons-Weiler, W. J. Elisens, G. Moore, A. D. Wolfe, and C. W. DePamphilis. 2000. Independent and combined analyses of sequences from all three genomic compartments converge on the root of

flowering plant phylogeny. Proc. Natl. Acad. Sci. USA **97**: 13166–13171.

Basu, D., K., Dehesh, H.-J., Schneider-Poetsch, S. E., Harrington, S. R., McCouch, and P. H. Quail, 2000. Rice *PHYC* gene: structure, expression, map position and evolution. Plant Mol. Biol. **44**:27–42.

Bergelson, J., M. Kreitman, E. A. Stahl, and D. Tian. 2001. Evolutionary dynamics of plant *R*-genes. Science **292**:2281–2289.

Bessey, C. E. 1915. The phylogenetic taxonomy of flowering plants. Ann. Missouri Bot. Gard. **2**:109–164.

Bews, J. W. 1927. Studies in the ecological evolution of angiosperms. New Phytol. **26**:1–21.

Bishop, J. G., A. M. Dean, and T. Mitchell-Olds. 2000. Rapid evolution in plant chitinases: molecular targets of selection in plant-pathogen coevolution. Proc. Natl. Acad. Sci. USA **97**: 5322–5327.

Bogorad, L. 1950. Factors associated with the synthesis of chlorophyll in the dark in seedlings of *Pinus jeffreyi*. Bot. Gaz. **111**:221–241.

Botto, J. F, R. A. Sanchez, G. C. Whitelam, and J. J. Casal. 1996. Phytochrome A mediates the promotion of seed germination by very low fluences of light and canopy shade light in *Arabidopsis*. Plant Physiol. **110**:439–444.

Bower, F. O. 1882. The germination and embryogeny of *Gnetum gnemon*. Q. J. Micr. Sci. **22**:278–298.

Burgerstein, A. 1900. Ueber das Verhalten der Gymnospermen-Keimlinge im Licht und im Dunkeln. Ber. Dtsch. Bot. Ges. **18**:168–184.

Burgin, M. J., J. J. Casal, G. C. Whitelam, and R. A. Sanchez. 1999. A light-regulated pool of phytochrome and rudimentary high-irradiance responses under far-red light in *Pinus elliotti* and *Pseudotsuga menziesii*. J. Exp. Bot. **50**:831–836.

Casal, J. J. 2000. Phytochromes, cryptochromes, phototropin: photoreceptor interactions in plants. Photochem. Photobiol. **71**:1–11.

Casal, J. J., R. A. Sanchez, and M. J. Yanovsky. 1997. The function of phytochrome A. Plant Cell Environ. **20**:813–819.

Cerdán, P. D., M. J. Yanovsky, F. C. Reymunds, A. Nagatani, R. J. Staneloni, G. C. Whitelam, and J. J. Casal. 1999. Regulation of phytochrome B signaling by phytochrome A and FHY1 in *Arabidopsis thaliana*. Plant J. **18**:499–507.

Chen, R. A., A. Greer, and A. M. Dean. 1995. A highly active decarboxylating dehydrogenase with rationally inverted coenzyme specificity. Proc. Natl. Acad. Sci. USA **92**:11666–11670.

———. 1996. Redesigning secondary structure to invert coenzyme specificity in isopropylmalate dehydrogenase. Proc. Natl. Acad. Sci. USA **93**:12171–12176.

Chinn, E., and J. Silverthorne. 1993. Light-dependent chloroplast development and expression of a light-harvesting chlorophyll a/b-binding protein gene in the gymnosperm *Ginkgo biloba*. Plant Physiol. **103**:727–732.

Christensen, S., E. Laverne, G. Boyd, and J. Silverthorne. 2002. *Ginkgo biloba* retains functions of both type I and type II flowering plant phytochrome. Plant Cell Physiol. **43**:768–777.

Clapham, D. H., H. Ü. Kolukisaoglu, C.-T. Larsson, M. Qamaruddin, I. Ekberg, C. Wiegmann-Eirund, H. A. W. Schneider-Poetsch, and S. Von Arnold. 1999. Phytochrome types in *Picea* and *Pinus*. Expression patterns of *PHYA*-related types. Plant Mol. Biol. **40**:669–678.

Crandall, K. A., C. R. Kelsey, H. Imamichi, H. C. Lane, and N. P. Salzman. 1999. Parallel evolution of drug resistance in HIV: failure of nonsynonymous/synonymous substitution rate ratio to detect selection. Mol. Biol. Evol. **16**:372–382.

Crane, P. R., E. M., Friis, and K. R. Pederson, 1995. The origin and early diversification of angiosperms. Nature **374**:27–33.

Cronquist, A. 1988. The evolution and classification of flowering plants. The New York Botanical Garden, New York.

Cunningham, C. W., K. E. Omland, and T. H. Oakley. 1998. Reconstructing ancestral character states: a critical reappraisal. Trends Ecol. Evol. **13**:361–366.

Doyle, J. A. 1978. Origin of angiosperms. Annu. Rev. Ecol. Syst. **9**:365–392.

Doyle, J. A., and M. J. Donoghue. 1986. Seed plant phylogeny and the origin of angiosperms: an experimental cladistic approach. Bot. Rev. **52**:321–431.

———. 1993. Phylogenies and angiosperm diversification. Paleobiology **19**:141–167.

Doyle, J. A., and P. K. Endress. 2000. Morphological phylogenetic analysis of basal angiosperms: comparison and combination with molecular data. Int. J. Plant Sci. **161**:S121–S153.

Doyle, J. A., and L. J. Hickey. 1976. Pollen and leaves from the mid-Cretaceous Potomac group and their bearing on early angiosperm evolution. Pp. 139–206 *in* C. B. Beck, ed. Origin and early evolution of angiosperms. Columbia University Press, New York.

Dykhuizen, D., and D. L. Hartl. 1980. Selective neutrality of 6PGD allozymes in *E. coli* and the effects of genetic background. Genetics **96**:801–817.

Eichenberg, K., I. Baurle, N. Paulo, R. A. Sharrock, W. Rudiger, and E. Schafer. 2000. *Arabidopsis* phytochromes C and E have different spectral characteristics from those of phytochromes A and B. FEBS Lett. **470**:107–112.

Endo, T., K. Ikeo, and T. Gojobori. 1996. Large-scale search for genes on which positive selection may operate. Mol. Biol. Evol. **13**:685–690.

Fankhauser, C., and J. Chory. 1997. Light control of development. Annu. Rev. Cell Dev. Biol. **13**:203–229.

Feild, T. S., N. C. Arens, and T. E. Dawson. 2003. Emerging perspectives on the ancestral ecophysiology of angiosperms. Intl. J. Plant Sci. In press.

Force, A., M. Lynch, F. B. Pickett, A. Amores, Y. Yan, and J. Postlethwait. 1999. Preservation of duplicate genes by complementary, degenerative mutations. Genetics **151**:1531–1545.

Friis, E. M., W. G., Chalconer, and Crane, P. R. 1987. The origins of angiosperms and their biological consequences. Cambridge University Press, Cambridge.

Furuya, M, and E. Schäfer. 1996. Photoperception and signaling of induction reactions by different phytochromes. Trends Plant Sci. **1**:301–306.

Golding, G. B., and A. M. Dean. 1998. The structural basis of molecular adaptation. Mol. Biol. Evol. **15**:355–369.

Goldman, N., and Z. Yang. 1994. A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol. Biol. Evol. **11**:725–736.

Graham, S. W., P. A. Reeves, A. C. E. Burns, and R. G. Olmstead. 2000. Microstructural characters in noncoding chloroplast DNA: interpretation, evolution, and utility of indels and inversions in basal angiosperm phylogenetic inference. Int. J. Plant Sci. **161**(Suppl.):S83–S96.

Grubb, P. J. 1998. Seeds and fruits of tropical rainforest plants: interpretation of the range of seed size, degree of defense, and flesh/seed quotients. Pp. 1–24 *in* D. M. Newberry, H. H. T. Prins, and N. D. Brown, eds. Dynamics of tropical communities. Blackwell Scientific, Oxford.

Gu, X. 1999. Statistical method for testing functional divergence after gene duplication. Mol. Biol. Evol. **16**:1664–1674.

———. 2001. Maximum-likelihood approach for gene family evolution under functional divergence. Mol. Biol. Evol. **18**: 453–464.

Hall, B. G. 1984. The evolved β-galactosidase system of *Escherichia coli*. Pp. 165–185 *in* R. P. Mortlock, ed.

Microorganisms as model systems for studying evolution. Plenum Press, New York.
Halliday, K. J., B. Thomas, and G. C. Whitelam. 1997. Expression of heterologous phytochromes A, B, or C in transgenic tobacco plants alters vegetative development and flowering time. Plant J. **12**:1079–1090.
Hennig, L., W. M. Stoddart, M. Dieterle, G. C. Whitelam, and E. Schäfer. 2002. Phytochrome E controls light-induced germination in *Arabidopsis*. Plant Physiol. **128**:194–200.
Hickey, L. J., and J. A. Doyle. 1977. Early Cretaceous fossil evidence for angiosperm evolution. Bot. Rev. **43**:1–104.
Hill, R. E., and N. D. Hastie. 1987. Accelerated evolution in the reactive centre regions of serine protease inhibitors. Nature **326**:96–99.
Huelsenbeck, J. P., and F. R. Ronquist. 2001. MrBayes: Bayesian inference of phylogeny. Bioinformatics **17**:754–755.
Hughes, A. L. 1999. Adaptive evolution of genes and genomes. Oxford University Press, New York.
Kendrick, R. E., and G. H. M. Kronenberg. 1994. Photomorphogenesis in plants. Kluwer Academic, Dordrecht, The Netherlands.
Knoll, A. H. 1984. Patterns of extinction in the fossil record of vascular plants. Pp. 21–68 *in* M. H. Nitecki, ed. Extinctions. University of Chicago Press, Chicago.
Lidgard, S., and P. R. Crane. 1988. Quantitative analyses of the early angiosperm radiation. Nature **331**:344–346.
Lupia, R., S. Lidgard, and P. R. Crane. 1999. Comparing palynological abundance and diversity: implications for biotic replacement during the Cretaceous angiosperm radiation. Paleobiology **25**:305–340.
Lynch, M., and J. S. Conery. 2000. The evolutionary fate and consequences of duplicate genes. Science **290**:1151–1155.
Lynch, M., and A. Force. 2000. The probability of duplicate gene preservation by subfunctionalization. Genetics **154**: 459–473.
Maddison, D. R., and W. P. Maddison. 2000. MacClade 4. Analysis of phylogeny and character evolution. Sinauer Associates, Sunderland, Mass.
Maloof, J., J. O. Borevitz, T. Dabi et al. (11 co-authors). 2001. Natural variation in light sensitivity of *Arabidopsis*. Nat. Genet. **29**:441–446.
Mathews, S., and M. J. Donoghue. 1999. The root of angiosperm phylogeny inferred from duplicate phytochrome genes. Science **286**:947–950.
———. 2000. Basal angiosperm phylogeny inferred from duplicate phytochromes A and C. Int. J. Plant Sci. **161**(Suppl. 6):S41–S55.
Mathews, S., M. Lavin, and 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.
Mathews, S., and R. A. Sharrock. 1997. Phytochrome gene diversity. Plant Cell Environ. **20**:666–671.
McCormac, A. C., G. C. Whitelam, M. T. Boylan, P. H. Quail, and H. Smith. 1992. Contrasting responses of etiolated and light-adapted seedlings to red:far-red ratio: a comparison of wild type, mutant and transgenic plants has revealed differential functions of members of the phytochrome family. J. Plant Physiol. **140**:707–714.
McNellis, T. W., and X.-W. Deng. 1995. Light control of seedling morphogenetic pattern. Plant Cell **7**:1749–1761.
Metcalfe, D. J., and P. J. Grubb. 1997. The responses to shade of seedlings of very small-seeded tree and shrub species from tropical rain forest in Singapore. Funct. Ecol. **11**:215–221.
Montgomery, B. L., and J. C. Lagarias. 2002. Phytochrome ancestry: sensors of bilins and light. Trends Plant Sci. **7**:357–366.
Mukai, Y., K. Tazaki, T. Fujii, and N. Yamamoto. 1992. Light-independent expression of three photosynthetic genes, *cab*, *rbcS*, and *rbcL*, in coniferous plants. Plant Cell Physiol. **33**:859–866.
Nagatani, A., J. W. Reed, and J. Chory. 1993. Isolation and initial characterization of *Arabidopsis* mutants that are deficient in phytochrome A. Plant Physiol. **102**:269–277.
Newcomb, R. D., P. M. Campbell, D. L. Ollis, E. Cheah, R. J. Russell, and J. G. Oakshott. 1997. A single amino acid substitution converts a carboxylesterase to an organophosphorous hydrolase and confers insecticide resistance on a blowfly. Proc. Natl. Acad. Sci. USA **94**:7464–7468.
Nielsen, R., and Z. Yang. 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics **148**:929–936.
Ohno, S. 1970. Evolution by gene duplication. Springer-Verlag, Berlin.
Parkinson, C. L., K. L. Adams, and J. D. Palmer. 1999. Multigene analyses identify the three earliest lineages of extant flowering plants. Curr. Biol. **9**:1485–1488.
Parks, B. M., and P. H. Quail. 1993. hy8, a new class of *Arabiodopsis* long hypocotyl mutants deficient in functional phytochrome A. Plant Cell **5**:39–48.
Perutz, M. F. 1983. Species adaptation in a protein molecule. Mol. Biol. Evol. **1**:1–28.
Pigliucci, M., and J. Schmitt. 1999. Genes affecting phenotypic plasticity in *Arabidopsis*: pleiotropic effects and reproductive fitness of photomorphogenic mutants. J. Evol. Biol. **12**: 551–562.
Qin, M., R. Kuhn, S. Moran, and P. H. Quail. 1997. Overexpressed phytochrome C has similar photosensory specificity to phytochrome B but a distinctive capacity to enhance primary leaf expansion. Plant J. **12**:1163–1172.
Qiu, Y.-L., J. Lee, F. Bernasconi-Quadroni, D. E. Soltis, P. S. Soltis, M. Zanis, E. A. Zimmer, Z. Chen, V. Savolainen, and M. W. Chase. 1999. The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes. Nature **402:**404–407.
Quail, P. H. 1991. Phytochrome: a light-activated molecular switch that regulates plant gene expression. Annu. Rev. Genet. **25**:389–409.
Quail, P. H., M. T. Boylan, B. M. Parks, T. W. Short, and D. Wagner. 1995. Phytochromes: photosensory perception and signal transduction. Science **268**:675–680.
Sanderson, M. J., and J. A. Doyle. 2001. Sources of error and confidence intervals in estimating the age of angiosperms from *rbc*L and 18S rDNA data. Am. J. Bot. **88**: 1499–1516.
Savolainen, V., M. W. Chase, S. B. Hoot, C. M. Morton, D. E. Soltis, and C. Bayer. 2000. Phylogenetics of flowering plants based on combined analysis of plastid *atp*B and *rbc*L gene sequences. Syst. Biol. **49**:306–362.
Schmidt, M., and H. A. W. Schneider-Poetsch. 2002. The evolution of gymnosperms redrawn by phytochrome genes: The Gnetatae appear at the base of the gymnosperms. J. Mol. Evol. **54**:715–724.
Schmitt, A., A. C. McCormac, and H. Smith. 1995. A test of the adaptive plasticity hypothesis using transgenic and mutant plants disabled in phytochrome-mediated elongation responses to neighbors. Am. Nat. **146**:937–953.
Schneider-Poetsch, H. A. W., Ü., Kolukisaoglu, D. H., Clapham, J., Hughes, and Lamparter, T. 1998. Non-angiosperm phytochromes and evolution of vascular plants. Physiol. Plant. **102**:612–622.
Schluter, D, T. Price, A. Ø. Mooers, and D. Ludwig. 1997. Likelihood of ancestor states in adaptive radiation. Evolution **51**:1699–1711.

Sharrock, R. A., and 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.

Shinomura, T, K. Uchida, and M. Furuya. 2000. Elementary processes of photoreception by phytochrome A for high-irradiance response of hypocotyl elongation in *Arabidopsis*. Plant Physiol. **122**:147–156.

Smith, H. 1995. Physiological and ecological function within the phytochrome family. Annu. Rev. Plant Physiol. Plant Mol. Biol. **46**:289–315.

———. 2000. Phytochromes and light signal perception by plants—an emerging synthesis. Nature **407**:585–591.

Smith, H., Y. Xu, and P. H. Quail. 1997. Antagonistic but complementary actions of phytochromes A and B allow optimum seedling de-etiolation. Plant Physiol. **114**:637–641.

Soltis, P. S., D. E. Soltis, and M. W. Chase. 1999. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. Nature **402**:402–404.

Soltis, P. S., D. E. Soltis, M. W. Chase et al. (16 co-authors). 2000. Angiosperm phylogeny inferred from a combined data set of 18s rDNA, *rbc*L, and *atp*B sequences. Bot. J. Linn. Soc. **133**:381–461.

Somers, D. E., R. A. Sharrock, J. M. Tepperman, and P. H. Quail. 1991. The *hy3* long hypocotyls mutant of *Arabidopsis* is deficient in phytochrome B. Plant Cell **3**:1263–1274.

Stebbins, G. L. 1974. Flowering plants, evolution above the species level. Harvard University Press, Cambridge.

Sun, G., Q. Ji, D. L. Dilcher, S. Zheng, K. C. Nixon, and X. Wang. 2002. Archaefructaceae, a new basal angiosperm family. Science **296**:899–904.

Takano, M., H. Kanegae, T. Shinomura, A. Miyao, H. Hirochika, and M. Furuya. 2001. Isolation and characterization of rice phytochrome A mutants. Plant Cell **13**:521–534.

Takhtajan, A. L. 1969. Flowering plants: origin and dispersal. Smithsonian Institution, Washington, D.C.

Tavaré, S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. Pp. 57–86 *in* Lectures on mathematics in the life sciences. Vol. 17. American Mathematical Society, Providence, R.I.

Taylor, D. W., and L. J. Hickey. 1992. Phylogenetic evidence for the herbaceous origin of angiosperms. Plant Syst. Evol. **180**:137–156.

———. 1996. Evidence for and implications of an herbaceous origin for angiosperms. Pp. 232–266 *in* D. Taylor and L. Hickey, eds. Flowering plant origin, evolution, and phylogeny. Chapman and Hall, New York.

Thorne, R. F. 1974. A phylogenetic classification of the Annoniflorae. Aliso **8**:147–209.

Van Tuinen, A., L. H. J. Kerckhoffs, A. Nagatani, R. E. Kendrick, and M. Koornneef. 1995. Far-red light insensitive, phytochrome A deficient mutants of tomato. Mol. Gen. Genet. **246**:133–141.

Wagner, A. 2001. Birth and death of duplicated genes in completely sequenced eukaryotes. Trends Genet. **17**:237–239.

Wagner, D., and P. H. Quail. 1995. Mutational analysis of phytochrome B identifies a small COOH-terminal-domain region critical for regulatory activity. Proc. Natl. Acad. Sci. USA **92**:8596–8600.

Walsh, J. B. 1995. How often do duplicated genes evolve new functions? Genetics **110**:345–364.

Whitelam, G. C., E. Johnson, J. Peng, P. Carol, M. L. Anderson, J. S. Cowl, and N. P. Harberd. 1993. Phytochrome A null mutants of *Arabidopsis* display a wild-type phenotype in white light. Plant Cell **5**:757–768.

Wilks, H. M, K. W. Hart, R. Feeney, C. R. Dunn, H. Muirhead, W. N. Chia, D. A. Barstow, T. Atkinson, A. R. Clarke, and J. J. Holbrook. 1988. A specific, highly active malate dehydrogenase by redesign of a lactate dehydrogenase framework. Science **242**:1541–1544.

Wing, S. L., and L. D. Boucher. 1998. Ecological aspects of the Cretaceous flowering plant radiation. Annu. Rev. Earth Planet. Sci. **26**:379–421.

Wu, S.-H., and J. C. Lagarias. 2000. Defining the bilin lyase domain: lessons from the extended phytochrome superfamily. Biochemistry **39**:13487–13495.

Yang, Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J. Mol. Evol. **39**:306–314.

———. 1997. PAML: a program for phylogenetic analysis by maximum likelihood version. CABIOS **13**:555–556(http://abacus.gene.ucl.ac.uk/software/paml.html).

———. 1998. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. Mol. Biol. Evol. **15**:568–573.

Yang, Z., and J. P. Bielawski. 2000. Statistical methods for detecting molecular adaptation. Trends Ecol. Evol. **15**: 496–503.

Yang, Z., S. Kumar, and M. Nei. 1995. A new method of inference of ancestral nucleotide and amino acid sequences. Genetics **141**:1641–1650.

Yang, Z., and R. Nielsen. 2002. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. Mol. Biol. Evol. **19**:908–917.

Yang, Z., R. Nielsen, N. Goldman, and A.-M. K. Pedersen. 2000. Codon-substitution models for heterogenous selection pressure at amino acid sites. Genetics **155**:431–449.

Yang, Z., W. J. Swanson, and V. D. Vacquier. 2000. Maximum-likelihood analysis of molecular adaptation in abalone sperm lysin reveals variable selective pressures among lineages and sites. Mol. Biol. Evol. **17**:1446–1455.

Yanovsky, M. J., J. J. Casal, and J. P. Luppi. 1997. The *VLF* loci, polymorphic between ecotypes of Landsberg *erecta* and Columbia dissect two branches of phytochrome A signaling pathways that correspond to the very-low fluence and high-irradiance responses of phytochrome. Plant J. **12**:659–667.

Yanovsky, M. J., J. J. Casal, and G. C. Whitelam. 1995. Phytochrome A, phytochrome B and HY4 are involved in hypocotyl growth responses to natural radiation in *Arabidopsis*: weak de-etiolation of the *phyA* mutant under dense canopies. Plant Cell Environ. **18**:788–794.

Yanovsky, M. J., J. P. Luppi, D. Kirchbauer, O. B. Ogorodnikova, V. A. Sineshchekov, E. Adam, S. Kircher, R. J. Staneloni, E. Schäfer, F. Nagy, and J. J. Casal. 2002. Missense mutation in the PAS2 domain of phytochrome A impairs subnuclear localization and a subset of responses. Plant Cell **14**:1591–1603.

Zanis, M. J., D. E. Soltis, P. S. Soltis, S. Mathews, and M. J. Donoghue. 2002. The root of the angiosperms revisited. Proc. Natl. Acad. Sci. USA **99**:6848–6853.

Keith Crandall, Associate Editor

Accepted March 11, 2003