Amino acid polymorphisms in *Arabidopsis* phytochrome B cause differential responses to light

Daniele L. Filiault[†], Carolyn A. Wessinger[†‡], Jose R. Dinneny[‡§], Jason Lutes[§¶], Justin O. Borevitz[§‖], Detlef Weigel[§††], Joanne Chory[§¶‡‡], and Julin N. Maloof[†‡‡]

[†]Section of Plant Biology, College of Biological Sciences, University of California, One Shields Avenue, Davis, CA 95616; [§]Plant Biology Laboratory, and [¶]Howard Hughes Medical Institute, Salk Institute for Biological Studies, La Jolla, CA 92037; [‖]Department of Ecology and Evolution, University of Chicago, Chicago, IL 60610; and [††]Department of Molecular Biology, Max Planck Institute for Developmental Biology, 72076 Tübingen, Germany

Contributed by Joanne Chory, December 28, 2007 (sent for review November 6, 2007)

Plants have a sophisticated system for sensing and responding to their light environment. The light responses of populations and species native to different habitats show adaptive variation; understanding the mechanisms underlying photomorphogenic variation is therefore of significant interest. In *Arabidopsis thaliana*, phytochrome B (PHYB) is the dominant photoreceptor for red light and plays a major role in white light. Because *PHYB* has been proposed as a candidate gene for several quantitative trait loci (QTLs) affecting light response, we have investigated sequence and functional variation in *Arabidopsis PHYB*. We examined *PHYB* sequences in 33 *A. thaliana* individuals and in the close relative *Arabidopsis lyrata*. From 14 nonsynonymous polymorphisms, we chose 5 for further study based on previous QTL studies. In a larger collection of *A. thaliana* accessions, one of these five polymorphisms, I143L, was associated with variation in red light response. We used transgenic analysis to test this association and confirmed experimentally that natural *PHYB* polymorphisms cause differential plant responses to light. Furthermore, our results show that allelic variation of *PHYB* activity is due to amino acid rather than regulatory changes. Together with earlier studies linking variation in light sensitivity to photoreceptor genes, our work suggests that photoreceptors may be a common target of natural selection.

hypocotyl | linkage disequilibrium | natural variation

Plants use three types of photoreceptors to survey their light environment: phytochromes for red and far-red light and cryptochromes and phototropins for blue light (1). Changes in the light environment sensed through these receptors affect many aspects of plant development. The phytochrome family in *Arabidopsis thaliana* consists of five genes, *PHYA-PHYE*, with partially redundant developmental functions (2–10). At the seedling stage, phytochromes regulate emergence from the soil. Seedlings that germinate underground or in the dark cannot photosynthesize and extend their hypocotyl upwards toward the soil surface. Light is a cue that the soil surface has been reached; light perception, therefore, causes inhibition of hypocotyl elongation and the beginning of photoautotrophic growth. The primary photoreceptors for this response are PHYA in far-red light, PHYB in red light, and cryptochromes in blue light. Later in development, plants use PHYB, PHYD, and PHYE to detect neighbor proximity by monitoring the ratio of red to far-red (R/FR) light (4, 5). Because chlorophyll absorbs red but not far-red light, low R/FR ratios indicate close neighbors or canopy shade. Neighbor perception can induce a variety of competitive responses, including stem and leaf petiole elongation and early reproduction, collectively called the shade-avoidance syndrome. Notably, the proper response to light depends on whether or not the plant is native to that environment. For example, the ability of low R/FR to induce shade-avoidance responses is reduced in species and populations normally growing under shady conditions (11), an example of adaptive variation in phytochrome-mediated responses (12, 13).

Phytochromes exist in two photoconvertible forms: Pr, an inactive red-light absorbing form, and Pfr, the active far-red light absorbing form. In sunny conditions, which are characterized by high R/FR ratios, most phytochrome is in the active Pfr form. Shade causes a decrease in Pfr and a concomitant induction of shade-avoidance responses. Phytochromes signal through a web of downstream factors, including a family of related bHLH transcription factors, the PIFs and PILs (14), GIGANTEA (GI) (15), and the bZIP transcription factors HY5 and HYH that integrate signals from multiple photoreceptor pathways (16). Light-regulated protein degradation, often mediated by the E3 ubiquitin ligase COP1 and associated proteins (17), also is important for phytochrome signaling.

An evolutionary question of considerable interest is how selection acts on the different components of metabolic and developmental pathways. For example, in the case of the anthocyanin biosynthesis pathway, which is responsible for producing flower color pigments, it has been found that genes acting later in the pathway evolve more quickly than those upstream (18, 19). Interestingly, for interpretation of light signals, a number of changes have been found at the top of the pathway in the photoreceptors themselves. For example, reduced FR sensitivity of the *A. thaliana* strain Lm-2 was traced to a single amino acid change in PHYA that reduced photoconversion and autophosphorylation (20). Cloning of an *A. thaliana* flowering time quantitative trait loci (QTL) revealed that it was caused by an amino acid substitution that stabilized the CRY2 protein, thereby increasing activation of the photoperiod pathway (21). Variation in *PHYC* is responsible for differences in both flowering time and hypocotyl elongation among *A. thaliana* accessions (22), and a naturally occurring deletion in *PHYD* increases stem elongation (3). More generally, phytochrome genes evolve more quickly than the average plant gene (23), and there is evidence for positive selection early in the diversification of *PHYA* (24).

Although phenotypic variation due to polymorphisms in *A. thaliana PHYA* (20), *PHYC* (22), and *PHYD* (3) has been described, *PHYB* plays by far the largest role for white-light responses in standard laboratory strains. In addition, *PHYB* is a candidate for a white and red light QTL in *A. thaliana* (25), for a flowering time QTL in tomato (26), and a bud set QTL in

Author contributions: D.L.F. and C.A.W. contributed equally to this work; D.L.F., J.O.B., D.W., J.C., and J.N.M. designed research; D.L.F., C.A.W., J.R.D., J.L., and J.N.M. performed research; D.L.F., C.A.W., and J.N.M. analyzed data; and D.L.F. and J.N.M. wrote the paper.

The authors declare no conflict of interest.

Data Deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. EU352775–EU352793).

[‡]Present Address: Department of Biology, Duke University, Box 90338, Durham, NC 27708.

[‡‡]To whom correspondence may be addressed. E-mail: chory@salk.edu or jnmaloof@ucdavis.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0712174105/DC1.

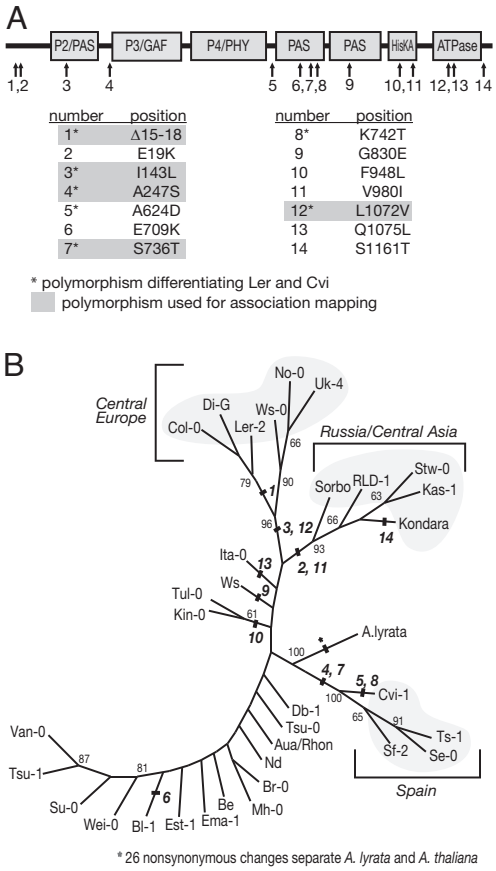

Fig. 1. Sequence diversity in *A. thaliana PHYB*. (*A*) Amino acid changes in *A. thaliana* PHYB. A schematic diagram of PHYB domains is shown, with arrows indicating each amino acid polymorphism. The table below indicates the specific amino acid change. The amino acid to the left of the number indicates the residue in the reference sequence from the Columbia (Col-0) accession. (*B*) Maximum likelihood phylogenetic tree, based on nucleotide sequence, showing *PHYB* divergence in *A. thaliana*. Bold italic numbers indicate where amino acid variants separate different branches. Plain text numbers in smaller font indicate bootstrap support.

poplar (27, 28). Here, we examine the effect of natural variation in *PHYB* on differential light response in *A. thaliana* and find that polymorphisms in PHYB proteins contribute to variation in *A. thaliana* light sensitivity.

Results and Discussion

To determine whether polymorphisms in *PHYB* could contribute to variation in light response, we began by examining *PHYB* sequence diversity in *Arabidopsis*. We sequenced the *PHYB* gene from L*er* and Cvi, two accessions known to segregate for *LIGHT2*, a QTL coincident with *PHYB* (25); 18 additional *A. thaliana* accessions; and the sister species *A. lyrata*. In our sequence analysis, we also included 14 *A. thaliana* accessions for which *PHYB* sequences were available in GenBank (29). In *A. thaliana*, we found 65 synonymous and 14 nonsynonymous polymorphisms. An additional 15 polymorphisms were identified in introns. Of the 14 nonsynonymous sites, 7 differentiated L*er* and Cvi (Fig. 1 *A* and *B*). A maximum likelihood phylogenetic tree revealed three well supported clades: one clade of six accessions contained the standard laboratory strains Col and L*er*, which originate from Central Europe; another one included five Russian and Central Asian accessions; and the third clade comprised of four Spanish accessions (Fig. 1*B*). The remaining 18 accessions grouped into a large, poorly resolved clade. In general, the pattern seen is consistent with previously reported genomic patterns of isolation by distance (30, 31). A similar tree was obtained by using neighbor-joining methods.

Because the 14 nonsynonymous polymorphisms all fell outside the functionally important GAF and PHY domains (Fig. 1*A*), we asked whether this pattern could be explained by selection. We modeled sequence evolution, using maximum likelihood methods to assess whether substitution rates differed between the PHY and GAF domains, which are important for chromophore binding and photoconversion (32), and the rest of the protein. For codon position 3, where polymorphisms are usually silent, there was no evidence for different rates ($P = 0.95$). However, the substitution rate for codon positions 1 and 2, where polymorphisms frequently cause amino acid changes, was significantly lower in the GAF and PHY domains ($P = 0.009$) compared with the rest of the protein. We only found evidence for depressed substitution rates at positions 1 and 2, suggesting that the reduced sequence diversity in the PHY and GAF domains is due to selection on protein sequence and function, rather than local variation in mutation rate.

The observed distribution of nonsynonymous sites might reflect that the GAF and PHY domains are more highly constrained than the rest of the protein. However, we found no evidence that mutations induced in the laboratory and known to compromise PHYB function unevenly affected the PHY and GAF domains: 12 of 29 missense mutations are in the GAF or PHY domains (32; $P = 0.4554$, Fisher's exact test, comparing mutation rate in the GAF and PHY domains with the rest of the protein). Thus, with respect to overall activity, the GAF and PHY domains are unlikely to be more highly constrained than the rest of the protein. An alternative explanation is that the observed pattern is a result of adaptive evolution, indicating that domains other than the PHY and GAF are better targets for selection. We used the McDonald–Kreitman test (33) to look for evidence of nonneutral selection. Although there is an excess of nonsynonymous polymorphisms that are fixed between species compared with ones that segregate within species, this excess is not significant [supporting information (SI) Table 1; $P > 0.26$], so the underlying cause of the observed substitution pattern remains unresolved.

To examine whether the differential distribution of polymorphic sites has persisted over a greater evolutionary distance, a phylogenetic shadowing (34) approach was taken. Using eShadow (35), both divergence threshold (DT) and hidden Markov model islands (HMMI) methods were used to identify conserved regions in four PHYB protein sequences from the Brassicales and 16 further eudicot sequences available in GenBank. Within the Brassicales, the HMMI algorithm predicted conservation in regions spanning most of the P2/PAS, GAF, and PHY domains (Fig. 2), with one interesting exception, an ≈50-aa region in the center of the GAF domain. This less conserved region corresponds to the portion of the GAF domain in bacterial phytochrome that contributes to the light-sensing knot structure (36). Compared with the HHMI method, the DT approach identified a slightly smaller conserved region from the C-terminal third of the P2/PAS domain to approximately the first 50 aa of the PHY domain, again excluding the light-sensing knot region in the GAF domain. Using eudicot sequences, similar patterns of conservation were seen with both the DT and HMMI models (SI Fig. 6). In these more divergent sequences, conservation is limited to the GAF domain excluding the light-sensing knot and the N-terminal end of the PHY domain. In summary, the patterns over longer time scales are predictive for the variation seen within *A. thaliana*, suggesting that the PHY and GAF domains are the most slowly-evolving domains during evolution. That the GAF contribution to the phytochrome knot

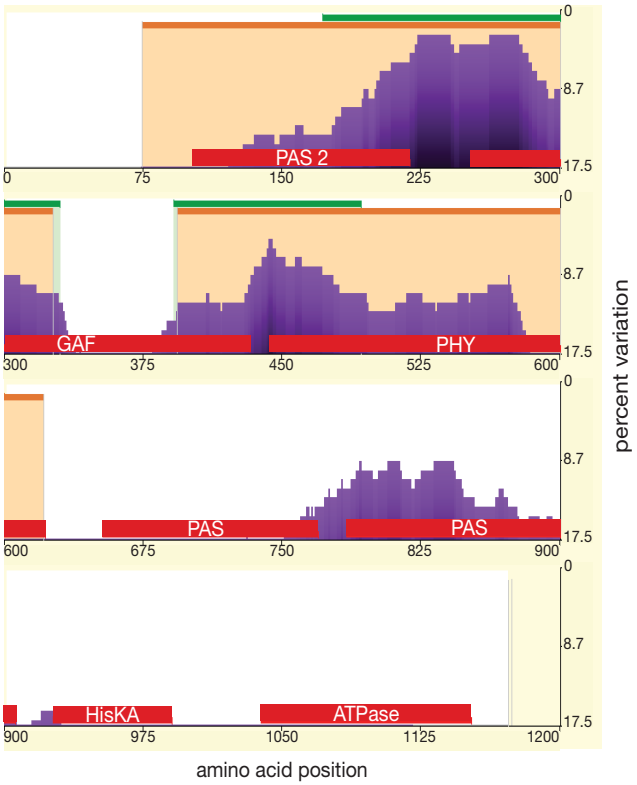

Fig. 2. Phylogenetic shadowing. Shown is the percentage variation by amino acid position for four Brassicales PHYB sequences. Regions determined to be slow-evolving are indicated by orange bars (HMMI) or green bars (DT). Functional domains are marked with red bars.

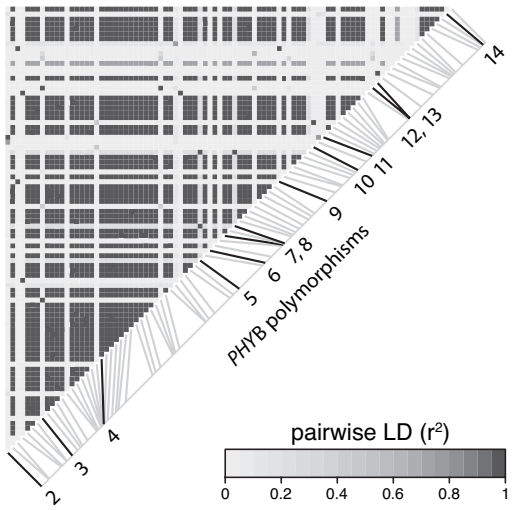

Fig. 3. *PHYB* linkage disequilibrium. Pairwise linkage disequilibrium (r^2) between single nucleotide polymorphisms in the *A. thaliana PHYB* gene. Polymorphic sites are arranged along the diagonal from lower left (5′) to upper right (3′). Numbered polymorphisms highlighted by darker lines correspond to the numbered amino acid polymorphisms in Fig. 1. To read the LD between any two sites, trace a column upwards from the 5′ site and a row leftward from the 3′ site; the square in which the column and row intersect indicates the LD.

is not conserved suggests that primary sequence may not be a particularly important determinant for this structural feature.

We next asked whether any of the observed sequence variation could impact *PHYB* activity. Association or linkage disequilibrium (LD) mapping uses historical recombination events in natural populations to associate polymorphisms with phenotypic variation (37). We used this technique to address whether any of the amino acid polymorphisms between *A. thaliana* accessions L*er* and Cvi are associated with natural variation in light response. A panel of 140 accessions that had been phenotyped for seedling light responses (20) was genotyped for the five nonsingleton polymorphisms that distinguish L*er* and Cvi. A whole-genome survey of LD in *A. thaliana* has shown that haplotype blocks are typically gene-sized (38). Consistent with this finding, there is strong LD across the *PHYB* gene (Fig. 3). Polymorphisms 1, 3, and 12 were specifically found to be associated with differences in hypocotyl elongation in red light ($P < 0.01$, $P < 0.005$, and $P < 0.05$, respectively; Fig. 4). Red light is precisely the condition where variation in *PHYB* would be most easily detected, suggesting that these associations are meaningful.

Because of the potentially confounding effects of population structure (39), the *PHYB* association can only be suggestive of being causal for differential light response across accessions. Therefore, we decided to compare the function of *PHYB*-L*er* and *PHYB*-Cvi experimentally. We first asked whether differences in *PHYB* activity (if any) would be more likely due to coding or regulatory changes. Three findings suggested that regulatory differences might not be important. First, protein blots did not show any obvious differences in PHYB levels among accessions (data not shown). Second, an analysis of published microarray data (40) revealed that *PHYB* mRNA levels are similar across accessions, with no significant difference between L*er* and Cvi ($P > 0.8$), and that there was no correlation between *PHYB* mRNA levels and hypocotyl elongation (Fig. 5 *A* and *B*). Third, an eQTL study that used microarrays to examine expression differences in L*er*, Cvi, and a derived mapping population found no evidence for differential *PHYB* expression (41). Therefore, we concluded that if *PHYB* was a source of variation in light sensitivity, then amino acid changes, rather than regulatory polymorphisms, between *PHYB*-L*er* and *PHYB*-Cvi were most likely the cause.

To examine possible differences in protein activity, we linked *PHYB* cDNAs from both L*er* and Cvi to the constitutive Cauliflower Mosaic Virus 35S promoter (35S), uncoupling *PHYB* function from any possible promoter differences. We transformed the null mutant *phyB-9*, which has greatly reduced light sensitivity and tall hypocotyls. Based on the allelic effects of the *LIGHT2* QTL (25) and our association mapping results, *35S*::*PHYB*-L*er* was expected to produce a more active PHYB than *35S*::*PHYB*-Cvi. Indeed, hypocotyls of *phyB-9 35S*::*PHYB*-L*er* plants were significantly shorter than those of *phyB-9 35S*::*PHYB*-Cvi ($P < 0.008$ for T1 generation, and $P < 0.0008$ for T2 generation plants; Fig. 5*C*), indicating that *PHYB-Cvi* confers less light responsiveness than *PHYB*-L*er* in a *phyb-9* (Col) background. It is possible that this difference is due to genetic interactions between the Col background and the different *PHYB* alleles. However, because *PHYB-Cvi* is associated with reduced light response across accessions and no loci were found to be epistatic with *LIGHT2* in L*er* X CviQTL analysis, it is more likely that the reduced response of *PHYB-Cvi* is inherent to the protein. In summary, *PHYB* polymorphisms are likely responsible for the significant associations across accessions and at least partially responsible for the *LIGHT2* QTL.

We used regression analysis to determine which of the three polymorphisms identified as significant in association mapping are most likely to be responsible for differences in PHYB activity. We regressed the polymorphisms both individually and in combination against the light responses. This process identified polymorphism 3 as the strongest candidate for the causative change. Specifically, models that included polymorphism 3 in

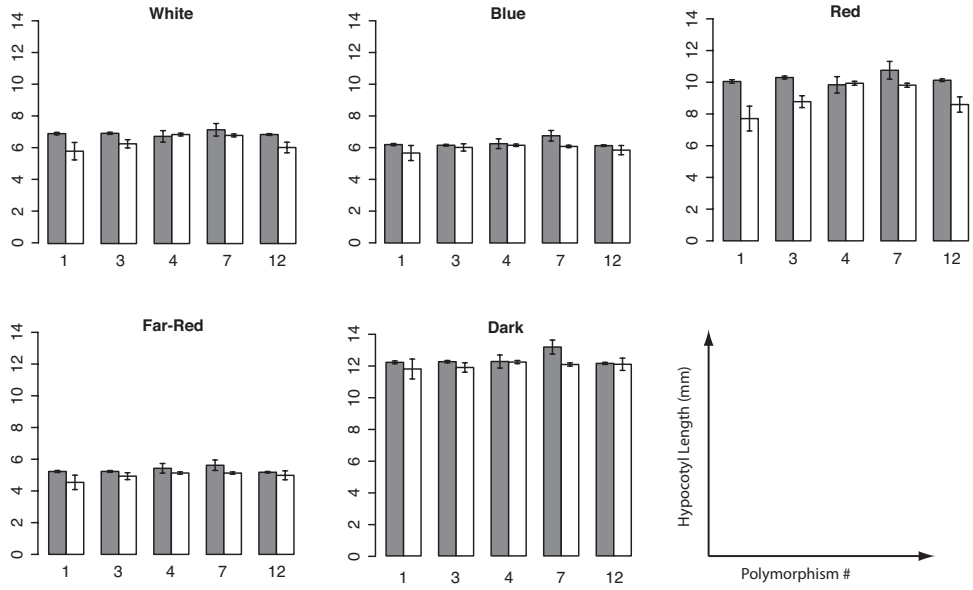

Fig. 4. Phenotypic association with *PHYB* polymorphisms. For each condition and polymorphism the bar height indicates average hypocotyl length of accessions with L*er* (gray) or Cvi (white) alleles. Hypocotyl data are from ref. 21 and represent an average of 14 hypocotyls per condition for each of 140 accessions. Error bars indicate the SEM.

combination with either 1 or 12 fit the data significantly better than models with polymorphism 1 or 12 alone. However, the combination of polymorphisms 3 plus 1 or 3 plus 12 did not provide any improvement over polymorphism 3 as the only explanatory variable (SI Table 2). Polymorphism 3 falls within a moderately conserved region of Brassicaceae PHYB. Comparison of the *A. thaliana* PHYB and PHYA sequences with the bacterial phytochrome *Dr*BphP sequence indicated that this polymorphism lies within the P2/PAS domain of the photosensory core (36). By superimposing secondary structural elements of bacterial phytochrome on *Arabidopsis* PHYB sequence, we found that polymorphism 3 was predicted to be on the surface in a span of three α-helices, away from the chromophore binding pocket. This surface is speculated to be involved in protein–protein interactions (36), suggesting a possible mode-of-action for polymorphism 3.

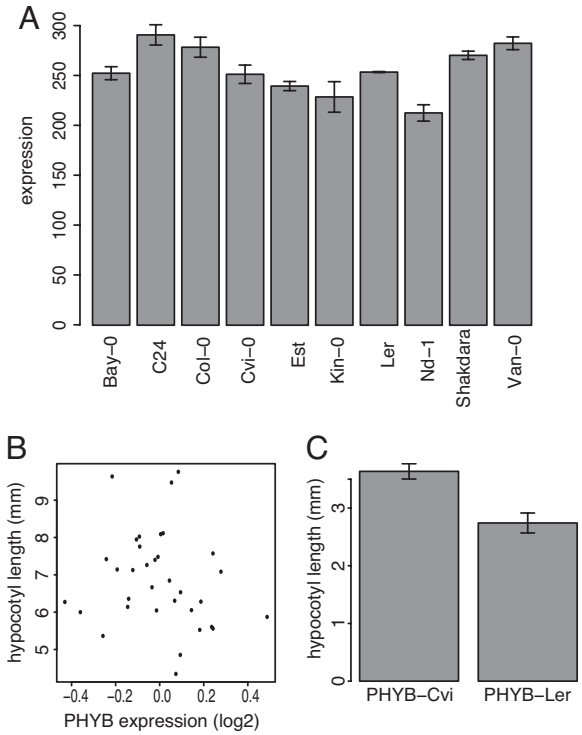

Fig. 5. PHYB protein variants confer different light sensitivities. (*A*) *PHYB* mRNA expression levels in 4-day-old seedlings, as determined on replicated microarrays (40). (*B*) Correlation between mean-centered *PHYB* mRNA expression levels and hypocotyl length. This set includes both replicated and unreplicated microarray measurements. (*C*) Hypocotyl length of *phyB-9* T2 plants transgenic with either phyB-L*er* or phyB-Cvi. An average of 56 plants from each of 40 (phyB-L*er*) or 44 (phyB-Cvi) independent single-insertion T1 lines was measured (4,711 plants total). Error bars indicate SEM.

Determining the genes responsible for natural variation and identifying the underlying QTL remains an important challenge (42). We have shown that amino acid polymorphisms in the major photoreceptor for red light contribute to variation in photomorphogenesis across *A. thaliana* accessions. There is a long-standing debate whether developmental variation is primarily due to changes in gene expression or in protein activity. Our findings with PHYB, together with those for PHYA, PHYC, and CRY2 (20–22), weigh in on the side of protein changes. Second, it has been proposed that in some pathways variants are more likely to be found downstream, rather than upstream, because of relatively relaxed constraints on downstream genes (18, 19) or selective sweeps acting on upstream genes (43). Here, we report that there is functional variation at the top of the PHYB pathway, complementing previous studies of other light-response pathways that also pointed to the photoreceptors themselves being responsible for variation in light sensing (20–22,44). Similarly, another study has shown evidence for selection upstream in a floral development pathway (43).

Changes in *PHYA* and *CRY2* are limited to individual accessions (20, 21), making it difficult to know whether they confer adaptive advantages or are simply only mildly deleterious mutations that have not yet been purged from nature. In contrast, alternative variants that cause differential light sensitivity are more common in the case of *PHYC* (22) and *PHYB* (this work), suggesting more strongly that they may be important in adaptation. If any of these changes are indeed adaptive, it could

indicate that overall light sensitivity rather than a particular downstream process is more prone to selection. It will therefore be interesting to study the molecular evolution of the different light-sensing pathways as comprehensive whole-genome variation data become available.

Materials and Methods

DNA Sequencing and Assembly. Detailed sequencing methods are described in *SI Text*. All polymorphisms were confirmed in multiple sequencing reads. *PHYB* sequence for *Arabidopsis lyrata* subspecies lyrata was assembled from whole genome shotgun reads, using the Staden package (45). Sequences have been deposited in GenBank (29) as accession nos. EU352775–EU352793.

Phylogenetic Analysis. *PHYB* sequences were obtained either from the above sequencing or from GenBank (SI Table 3) and were aligned with ClustalX (46, 47). PHYLIP (48) was used to bootstrap the dataset 100 times, determine maximum likelihood and neighbor joining trees, and find majority rule consensus trees. PAUP* software, Version 4.0b10 (49), was used to determine whether there were different rates of nucleotide substitution in the P2/PAS, GAF, and PHY domains relative to the rest of the protein within *A. thaliana* by comparing different substitution models, as described in *SI Text*.

Phylogenetic Shadowing. Sequences obtained from GenBank (SI Table 3) were used as input for eShadow (http://eshadow.dcode.org). HMMI analysis of Brassicales used the following probabilities: $eS = 0.85$, $eF = 0.80$, and $T = 0.2$. Brassicales DT analysis used a maximum percent variation of 5% and a minimum length of 80 aa. For eudicot sequences, HMMI analysis used probabilities of $eS = 0.75$, $eF = 0.60$, and $T = 0.1$, whereas DT analysis used a maximum percent variation of 20% and a minimal length of 80 aa.

Association Mapping. Simple sequence length polymorphism, cleavable amplified polymorphic sequence (CAPS) and derived CAPS assays were designed and used to genotype >100 *A. thaliana* accessions for the five nonsingleton nonsynonymous polymorphisms between Ler and Cvi (SI Tables 4 and 5). A permutation-based approach was used to determine association with phenotypic differences. First, a *t* statistic for correlation between genotype and phenotype was calculated for each polymorphism. To establish the significance of these associations, 10,000 permuted datasets were analyzed, and, for each permutation, the highest *t* statistic across all polymorphisms was recorded. An association was deemed significant if its *t* statistic was larger than the appropriate quantile *t* statistic from the permuted dataset.

LD Analysis. Ninety-two SNPs were identified from a MAFFT multiple sequence alignment of 33 genomic PHYB sequences (SI Tables 3 and 6). One was removed from the analysis because of incomplete data. The LDheatmap package (50) in the R statistical environment (51) was used to determine and plot pairwise linkage disequilibrium, using allelic correlation (r^2).

Transgenic Analysis. The null *phyB-9* allele was transformed with *PHYB*-Cvi and *PHYB*-Ler as described in *SI Text*. Two independent transformations were performed with a total of 483 *PHYB*-Cvi and 662 *PHYB*-Ler T1 transformants assayed in nine independent experiments. Significance was determined by using a linear mixed-effects model with *PHYB* construct as a fixed effect and transformation plus experiment (within transformation) as random effects. To confirm these results, kanamycin resistant T2 progeny from 40 *PHYB-Cvi* and 44 *PHYB-Ler* T1 plants were analyzed in two independent experiments. Experiment was not a significant factor, so the mixed-effects model contained *PHYB* construct (fixed effect), plate (random effect), and T1 parent (random effect).

Microarray Analysis. Microarray data (40) were imported into R (51) and Bioconductor (52) and normalized by RMA (53). Because we were only querying a single gene, a simple *t* test was used to determine whether there was evidence for differential *PHYB* expression between Ler and Cvi ($n = 2$ and 3 replicates, respectively).

ACKNOWLEDGMENTS. We thank Andreah Wallace and Gabriel Trainer for technical assistance. This work was funded by National Science Foundation Grant DBI-0227103 (to J.N.M and Cynthia Weinig), National Insitutes of Health Grant GM62932 (to J.C. and D.W.), The Howard Hughes Medical Institute, and National Insitutes of Health Training Grant T32 GM070377 (to D.L.F.).

1. Chen M, Chory J, Fankhauser C (2004) Light signal transduction in higher plants. *Annu Rev Genet* 38:87–117.
2. Reed JW, Nagatani A, Elich TD, Fagan M, Chory J (1994) Phytochrome A, phytochrome B have overlapping but distinct functions in *Arabidopsis* development. *Plant Physiol* 104:1139–1149.
3. Aukerman MJ, *et al.* (1997) A deletion in the PHYD gene of the *Arabidopsis Wassilewskija* ecotype defines a role for phytochrome D in red/far-red light sensing. *Plant Cell* 9:1317–1326.
4. Devlin PF, Patel SR, Whitelam GC (1998) Phytochrome E influences internode elongation and flowering time in *Arabidopsis*. *Plant Cell* 10:1479–1487.
5. Devlin PF, *et al.* (1999) Phytochrome D acts in the shade-avoidance syndrome in *Arabidopsis* by controlling elongation growth and flowering time. *Plant Physiol* 119:909–915.
6. Monte E, *et al.* (2003) Isolation and characterization of phyC mutants in *Arabidopsis* reveals complex crosstalk between phytochrome signaling pathways. *Plant Cell* 15:1962–1980.
7. Halliday KJ, Whitelam GC (2003) Changes in photoperiod or temperature alter the functional relationships between phytochromes and reveal roles for phyD and phyE. *Plant Physiol* 131:1913–1920.
8. Clack T, Mathews S, Sharrock RA (1994) The phytochrome apoprotein family in *Arabidopsis* is encoded by five genes: the sequences and expression of PHYD and PHYE. *Plant Mol Biol* 25:413–427.
9. Franklin KA, *et al.* (2003) Phytochromes B, D, and E act redundantly to control multiple physiological responses in *Arabidopsis*. *Plant Physiol* 131:1340–1346.
10. Franklin KA, Davis SJ, Stoddart WM, Vierstra RD, Whitelam GC (2003) Mutant analyses define multiple roles for phytochrome C in *Arabidopsis* photomorphogenesis. *Plant Cell* 15:1981–1989.
11. Maloof JN, Borevitz JO, Weigel D, Chory J (2000) Natural variation in phytochrome signaling. *Semin Cell Dev Biol* 11:523–530.
12. Dorn LA, Pyle EH, Schmitt J (2000) Plasticity to light cues and resources in *Arabidopsis thaliana*: Testing for adaptive value and costs. *Evolution* 54:1982–1994.
13. Donohue K, Messiqua D, Pyle EH, Heschel MS, Schmitt J (2000) Evidence of adaptive divergence in plasticity: Density- and site-dependent selection on shade-avoidance responses in Impatiens capensis. *Evolution* 54:1956–1968.
14. Duek PD, Fankhauser C (2005) bHLH class transcription factors take centre stage in phytochrome signalling. *Trends Plants Sci* 10:51–54.
15. Huq E, Tepperman JM, Quail PH (2000) GIGANTEA is a nuclear protein involved in phytochrome signaling in *Arabidopsis*. *Proc Natl Acad Sci USA* 97:9789–9794.
16. Holm M, Ma LG, Qu LJ, Deng XW (2002) Two interacting bZIP proteins are direct targets of COP1-mediated control of light-dependent gene expression in *Arabidopsis*. *Genes Dev* 16:1247–1259.
17. Serino G, Deng XW (2003) The COP9 signalosome: regulating plant development through the control of proteolysis. *Annu Rev Plant Biol* 54:165–182.
18. Lu Y, Rausher MD (2003) Evolutionary rate variation in anthocyanin pathway genes. *Mol Biol Evol* 20:1844–1853.
19. Rausher MD, Miller RE, Tiffin P (1999) Patterns of evolutionary rate variation among genes of the anthocyanin biosynthetic pathway. *Mol Biol Evol* 16:266–274.
20. Maloof JN, *et al.* (2001) Natural variation in light sensitivity of *Arabidopsis*. *Nat Genet* 29:441–446.
21. El-Din El-Assal S, Alonso-Blanco C, Peeters AJ, Raz V, Koornneef M (2001) A QTL for flowering time in *Arabidopsis* reveals a novel allele of CRY2. *Nat Genet* 29:435–440.
22. Balasubramanian S, *et al.* (2006) The PHYTOCHROME C photoreceptor gene mediates natural variation in flowering and growth responses of *Arabidopsis thaliana*. *Nat Genet* 38:711–715.
23. Alba R, Kelmenson PM, Cordonnier-Pratt MM, Pratt LH (2000) The phytochrome gene family in tomato and the rapid differential evolution of this family in angiosperms. *Mol Biol Evol* 17:362–373.
24. 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.
25. Borevitz JO, *et al.* (2002) Quantitative trait loci controlling light and hormone response in two accessions of *Arabidopsis thaliana*. *Genetics* 160:683–696.
26. Jiménez-Gómez JM, *et al.* (2007) Quantitative genetic analysis of flowering time in tomato. *Genome* 50:303–315.
27. Ingvarsson PK, García MV, Hall D, Luquez V, Jansson S (2006) Clinal variation in phyB2, a candidate gene for day-length-induced growth cessation and bud set, across a latitudinal gradient in European aspen (*Populus tremula*). *Genetics* 172:1845–1853.
28. Frewen BE, *et al.* (2000) Quantitative trait loci and candidate gene mapping of bud set and bud flush in populus. *Genetics* 154:837–845.
29. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL (2007) *GenBank Nucleic Acids Res* 35:D21–D25.
30. Nordborg M, *et al.* (2005) The pattern of polymorphism in *Arabidopsis thaliana*. *PLoS Biol* 3:e196.
31. Sharbel TF, Haubold B, Mitchell-Olds T (2000) Genetic isolation by distance in *Arabidopsis thaliana*: biogeography and postglacial colonization of Europe. *Mol Ecol* 9:2109–2118.
32. Rockwell NC, Su YS, Lagarias JC (2006) Phytochrome structure and signaling mechanisms. *Annu Rev Plant Biol* 57:837–858.
33. McDonald JH, Kreitman M (1991) Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature* 351:652–654.
34. Boffelli D, *et al.* (2003) Phylogenetic shadowing of primate sequences to find functional regions of the human genome. *Science* 299:1391–1394.
35. Ovcharenko I, Boffelli D, Loots GG (2004) eShadow: a tool for comparing closely related sequences. *Genome Res* 14:1191–1198.

36. Wagner JR, Brunzelle JS, Forest KT, Vierstra RD (2005) A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature* 438:325–331.
37. Flint-Garcia SA, Thornsberry JM, Buckler ES (2003) Structure of linkage disequilibrium in plants. *Annu Rev Plant Biol* 54:357–374.
38. Kim S, *et al.* (2007) Recombination and linkage disequilibrium in *Arabidopsis thaliana*. *Nat Genet* 39:1151–1155.
39. Zhao K, *et al.* (2007) An *Arabidopsis* example of association mapping in structured samples. *PLoS Genet* 3:e4.
40. Lempe J, *et al.* (2005) Diversity of flowering responses in wild *Arabidopsis thaliana* strains. *PLoS Genet* 1:109–118.
41. Keurentjes JJ, *et al.* (2007) Regulatory network construction in *Arabidopsis* by using genome-wide gene expression quantitative trait loci. *Proc Natl Acad Sci USA* 104:1708–1713.
42. Weigel D, Nordborg M (2005) Natural variation in *Arabidopsis*. How do we find the causal genes? *Plant Physiol* 138:567–568.
43. Olsen KM, Womack A, Garrett AR, Suddith JI, Purugganan MD (2002) Contrasting evolutionary forces in the *Arabidopsis thaliana* floral developmental pathway. *Genetics* 160:1641–1650.
44. Olsen KM, Halldorsdottir SS, Stinchcombe JR, Weinig C, Schmitt J, Purugganan MD (2004) Linkage disequilibrium mapping of *Arabidopsis* CRY2 flowering time alleles. *Genetics* 167:1361–1369.
45. Staden R, Beal KF, Bonfield JK (2000) The Staden package, 1998. *Methods Mol Biol* 132:115–130.
46. 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.
47. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25:4876–4882.
48. Felsenstein J (2005) *PHYLIP (Phylogeny Inference Package), Version 3.6.* Available at http://evolution.genetics.washington.edu/phylip.html.
49. Swofford DL (2003) *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods)* (Sinaur Associates, Sunderland, MA).
50. Shin J-H, Blay S, McNeney B, Graham J (2006) LDheatmap: An R Function for Graphical Display of Pairwise Linkage Disequilibria Between Single Nucleotide Polymorphisms. *J Stat Soft* 16:Code Snippet 3.
51. R Development Core Team (2006) *R: A Language and Environment for Statistical Computing.* Available at www.r-project.org.
52. Gentleman RC, *et al.* (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 5:R80.
53. Irizarry RA, *et al.* (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4:249–264.