The Plant Cell, Vol. 17, 1894–1907, July 2005, www.plantcell.org

A Conserved Transcription Factor Mediates Nuclear Control of Organelle Biogenesis in Anciently Diverged Land Plants [W]

Yuki Yasumura,[1] Elizabeth C. Moylan,[2] and Jane A. Langdale[3]

Department of Plant Sciences, University of Oxford, Oxford OX1 3RB, United Kingdom

Land plant chloroplasts evolved from those found in the green algae. During land plant evolution, nuclear regulatory mechanisms have been modified to produce morphologically and functionally diverse chloroplasts in distinct developmental contexts. At least some of these mechanisms evolved independently in different plant lineages. In angiosperms, GOLDEN2-LIKE (GLK) transcription factors regulate the development of at least three chloroplast types. To determine whether GLK-mediated regulation of chloroplast development evolved within angiosperms or is a plesiomorphy within land plants, gene function was examined in the moss *Physcomitrella patens*. Gene expression patterns and loss-of-function mutant phenotypes suggested that *GLK* gene function is conserved between *P. patens* and *Arabidopsis thaliana*, species that diverged >400 million years ago. In support of this suggestion, moss genes partially complement Arabidopsis loss-of-function mutants. Therefore, GLK-mediated regulation of chloroplast development defines one of the most ancient conserved regulatory mechanisms identified in the plant kingdom.

INTRODUCTION

Oxygenic photosynthesis originated in aquatic single-celled cyanobacteria >2.5 billion years ago (Blankenship, 2001). All extant chloroplasts are derived from a primary endosymbiotic event that involved such cyanobacteria (Douglas, 1998; Moreira et al., 2000). The bacterium became integrated into the cell and was established as a membrane-bound organelle. Ancestral chloroplasts most likely contained chlorophylls *a*, *b*, and *c*, phycobilisomes, and unstacked thylakoid membranes (Moreira et al., 2000). As the three primary algal groups evolved, these features were differentially lost. Glaucophytes and rhodophytes (red algae) lost chlorophyll *b*, whereas chlorophytes (green algae) lost phycobilisomes and gained stacked thylakoids. Chlorophyll *c* was lost from all lineages except from the red algae that gave rise via secondary endosymbiosis to the chromists. Land plant chloroplasts evolved from those found in chlorophytes. In all cases, most genes were lost from the plastid genome and transferred to the nucleus (Martin et al., 2002). As a consequence, chloroplast biogenesis is regulated by the nucleus in all extant eukaryotic phototrophs.

The nucleus regulates chloroplast biogenesis by encoding structural components of the organelle, by regulating chloroplast division, and by regulating the developmental process per se. Although many nuclear genes that contribute functional components to the organelle have been described, including those of endosymbiotic origin that control division from within the organelle (Osteryoung et al., 1998), very few regulatory genes have been identified even in model organisms. One example, *Accumulation and Replication of Chloroplasts5*, encodes a cytoplasmically localized dynamin-like protein that regulates chloroplast division in *Arabidopsis thaliana* (Gao et al., 2003). A second example, the *Golden2-like* (*GLK*) genes, encode transcription factors that regulate chloroplast development in diverse species, namely in the monocot maize (*Zea mays*) and in the eudicot Arabidopsis (Langdale and Kidner, 1994; Hall et al., 1998; Rossini et al., 2001; Fitter et al., 2002). *GLK* genes are members of the plant-specific GARP family of transcription factors (Riechmann et al., 2000; Fitter et al., 2002). Within the GARP family, *GLK* genes are monophyletic, and gene duplications have occurred independently in the monocots and eudicots. In addition to the GARP DNA binding domain (DBD), *GLK* genes share a conserved C-terminal domain referred to as the GOLDEN2 C-terminal (GCT) box. In Arabidopsis, the two *GLK* genes act redundantly to regulate monomorphic chloroplast development (Fitter et al., 2002). By contrast, in maize, each gene acts alone in one of the two photosynthetic cell types that develop in the leaf (Rossini et al., 2001). These two cell types differentiate morphologically distinct chloroplasts. In combination, these observations suggest that *GLK* gene function is fundamentally required for chloroplast development in angiosperms and that at least in maize, specialization of chloroplast type is correlated with specialization of *GLK* gene function.

To determine whether GLK-mediated regulation of chloroplast development evolved within or before the land plants, we assessed gene activity in the bryophyte *Physcomitrella patens*. The bryophytes diverged from the rest of the land plants >400 million years ago and are the most distantly related group to the

[1] Current address: John Innes Centre, Colney Lane, Norwich NR4 7UH, UK.

[2] Current address: BioMed Central, Middlesex House, 32-42 Cleveland Street, London W1T 4LB, UK.

[3] To whom correspondence should be addressed. E-mail jane.langdale@plants.ox.ac.uk; fax 44-1865-275147.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Jane A. Langdale (jane.langdale@plants.ox.ac.uk).

[W] Online version contains Web-only data.

Article, publication date, and citation information can be found at www.plantcell.org/cgi/doi/10.1105/tpc.105.033191.

angiosperms (Gifford and Foster, 1989). In *P. patens*, the haploid gametophyte as opposed to the diploid sporophyte dominates the life cycle. As such, chloroplast biogenesis occurs in a different developmental context from that in angiosperms. The *P. patens* gametophyte consists of filamentous protonema (Figures 1A and 1B) and leafy shoots called gametophores (Figure 1D). There are two types of filaments: chloronema, which contain numerous chloroplasts (Figure 1A), and caulonema, which develop fewer and smaller chloroplasts (Figure 1B). The leaf-like structures of the gametophore are composed of a single layer of cells, the majority of which contain chloroplasts. The remaining cells constitute the conducting vessels that transport nutrients throughout the shoot. In addition to leaf-like structures, gametophores develop nonphotosynthetic rhizoid filaments that provide structural support for the shoot (Figures 1C and 1D). In total, therefore, three chloroplast-containing cell types differentiate during the *P. patens* life cycle.

We report here that there are two *Glk* genes in the *P. patens* genome. Mutations in both genes were generated by homologous recombination. Single mutants were phenotypically indistinguishable from wild-type plants, but double mutants were uniformly pale green. Therefore, the two genes function redundantly to regulate chloroplast biogenesis. Notably, similar aspects of chloroplast development are perturbed in double mutants of *P. patens* and Arabidopsis, suggesting that gene function is conserved between the two species. The partial complementation of Arabidopsis mutants with a moss *Glk* gene confirmed this suggestion. Thus, GLK-mediated regulation of chloroplast development is an ancient mechanism conserved across all land plants.

RESULTS

Two *Glk* Genes Are Present in the *P. patens* Genome

To determine whether *Glk* genes are present in the *P. patens* genome, genomic PCR was performed using degenerate primers designed from conserved DBD and GCT box sequences. Two *Glk* genes were isolated, both of which contain sequences encoding the conserved DBD and the GCT box (Figure 2A). Unlike the angiosperm *GLK* genes, the two moss genes share a high degree of amino acid similarity outside of these domains (75% identical), and neither gene contains introns. When *P. patens Glk* (Pp*Glk*) sequences were used as hybridization probes in DNA gel blot analysis of genomic DNA, no additional gene copies were identified (see Supplemental Figure 1 online). Phylogenetic analysis confirmed that Pp*Glk1* and Pp*Glk2* are most closely related to each other and to angiosperm *GLK* genes than to any other GARP genes (Figure 2B). To investigate relationships within the *GLK* gene family, a second phylogenetic analysis was conducted using cDNA sequence data from the DBD and the GCT box. This analysis placed the two Pp*Glk* genes in a sister group that is basal to the rest of the *GLK* genes (Figure 2C). This observation suggests that independent duplication events occurred in the bryophytes and angiosperms. The notable sequence identity between Pp*Glk1* and Pp*Glk2* further suggests either that this duplication was relatively recent in *P. patens* or that there is selective pressure to maintain two very similar genes in the genome.

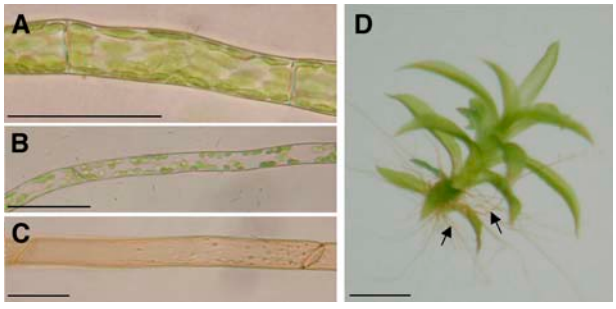

Figure 1. *P. patens* Gametophyte.

(A) Chloronemal cell.
(B) Caulonemal cell.
(C) Rhizoid cell.
(D) Gametophore with rhizoid filaments (arrows).
Bars = 50 μm (**[A]** to **[C]**) and 1 mm **(D)**.

Pp*Glk* Genes Are Expressed in Photosynthetic Tissue

To determine whether Pp*Glk* genes are expressed in tissues that differentiate chloroplasts, RT-PCR was performed using RNA isolated from protonema and gametophores. Figure 3A demonstrates that both genes are expressed in both tissues. Furthermore, in protonema (but not in gametophores), Pp*Glk* transcript levels are upregulated relative to tubulin transcripts in the light period of a diurnal light/dark cycle. This upregulation by light was confirmed by RNA gel blot analysis (see Supplemental Figure 2 online). Thus, Pp*Glk* gene expression patterns are consistent with a role in chloroplast development.

To examine Pp*Glk* gene expression at the cellular level, a *uidA* reporter gene encoding β-glucuronidase (GUS) was inserted into the Pp*Glk1* and Pp*Glk2* gene loci via homologous recombination (see Supplemental Figure 3 online for complete details). Constructs were designed so that double crossovers in the Pp*Glk* 5′ and 3′ homologous regions resulted in transcripts consisting of the Pp*Glk* gene 5′ region, *uidA*, and the Pp*Glk* 3′ untranslated region (UTR). Thus, transcript abundance would be regulated by the endogenous Pp*Glk* promoter and the fused 3′ UTR. The translated protein would be a fusion between the N-terminal portion of PpGLK and GUS. Three independently transformed lines were obtained for both Pp*Glk1*:*uidA* and Pp*Glk2*:*uidA*. At each of the six loci, multiple repeats of vector and/or construct sequences were present. However, in each case, there were no additional copies of *uidA* and the repeats were downstream of the Pp*Glk*:*uidA*:Pp*Glk* 3′ UTR fusion. Therefore, the repeats were unlikely to interfere with the gene expression analysis. In all six Pp*Glk*:*uidA* lines, GUS activity was detected in both filaments and leafy shoots (Figures 3B to 3D). Therefore, Pp*Glk1* and Pp*Glk2* are expressed in all chloroplast-containing cells of the gametophyte. A further correlation between Pp*Glk* gene expression and chloroplast development was seen in rhizoids. GUS activity was not detected in achloroplastic rhizoid filaments, but where rhizoids transdifferentiated into caulonema, GUS activity was observed (Figures 3E and 3F). Thus, as in angiosperms, *Glk*

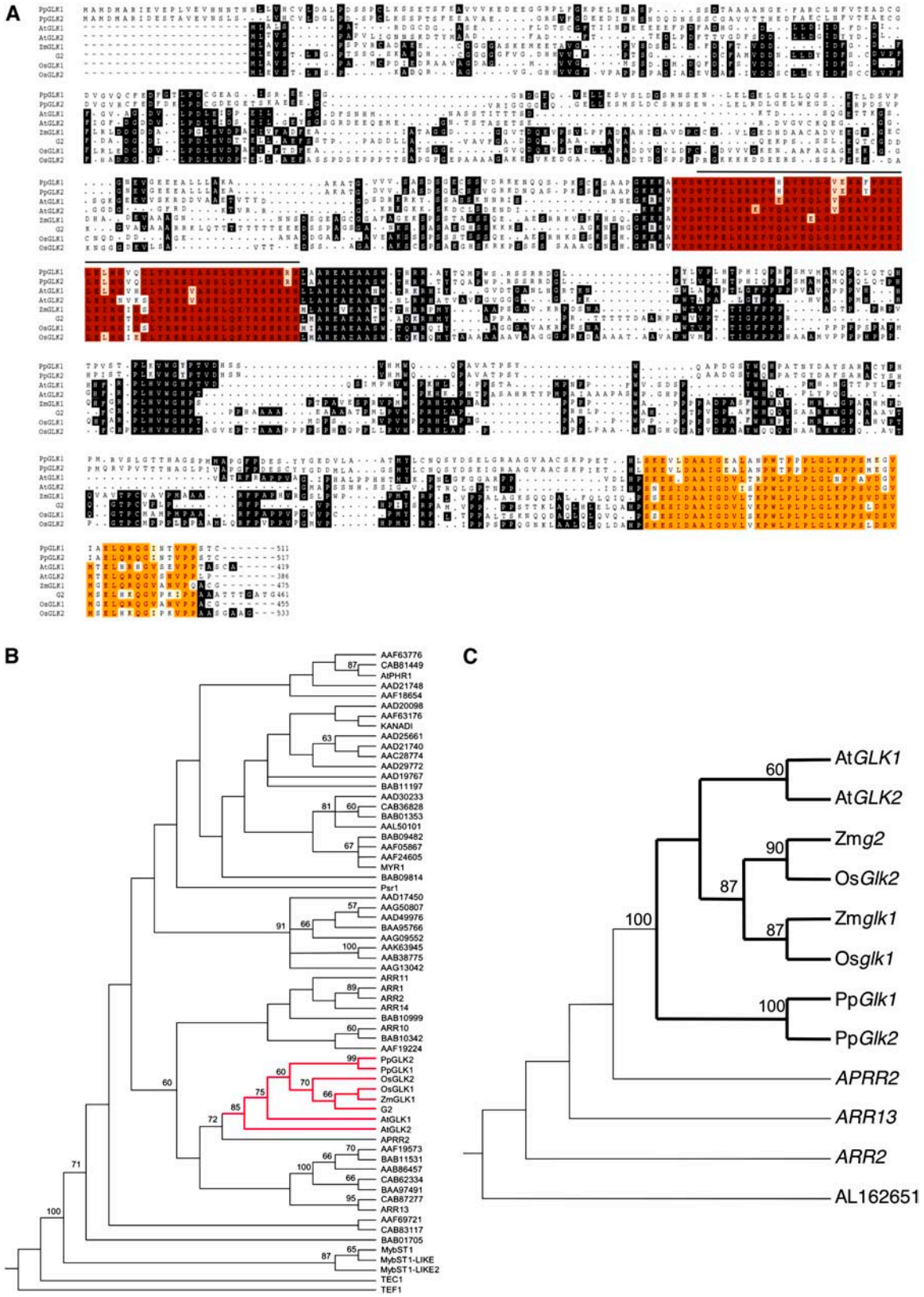

Figure 2. *Glk* Genes in *P. patens*.

gene expression in *P. patens* is induced by both light and developmental cues, and the presence of *Glk* transcripts is correlated with the presence of differentiated chloroplasts.

Generation of Pp*glk* Mutants

To determine the function of *Glk* genes in moss, loss-of-function alleles were generated by homologous recombination (see Supplemental Figure 4 online for complete details). Kanamycin resistance (35S:*nptII*) and hygromycin resistance (35S:*aphIII*) cassettes were used to disrupt Pp*Glk1* and Pp*Glk2*, respectively. Insertion constructs were designed so that double crossovers on either side of the antibiotic resistance cassette would lead to replacement of the endogenous DBD with the cassette. Wild-type moss was transformed with these constructs to generate the *glk* single mutants, Pp*glk1-1* to *-3* and Pp*glk2-1* to *-3*. Pp*glk1-1* was subsequently transformed with the Pp*Glk2* construct to generate the Pp*glk1-1 glk2-4* double mutant. Similarly, Pp*glk2-1* was transformed with the Pp*Glk1* construct to generate the Pp*glk1-5 glk2-1* double mutant. Therefore, three independent Pp*glk1* and Pp*glk2* single mutant lines and two independent double mutant lines were obtained. In each case, the Pp*Glk* locus was disrupted, but in some cases, vector and/or construct DNA was present in addition to the predicted antibiotic resistance gene. Regardless of the amount of DNA inserted at each locus, all recombined alleles were likely to be null.

In total, 311 transformed lines were generated. Of these, 106 contained inserts at the predicted site, but only 23 were free of additional inserts at nontarget sites. Therefore, 93% had undergone illegitimate recombination at one or more sites, whereas only 30% had undergone homologous recombination. These frequencies disagree with previously published data; however, the constructs used in the previous study did not contain *P. patens* genomic DNA (Schaefer, 2001). This raises the possibility that the additional integration events resulted from homologous recombination with reduced specificity rather than by illegitimate recombination. We discount this suggestion, however, because none of the Pp*Glk1* constructs was inserted into the Pp*Glk2* locus or vice versa, yet the two sequences are likely to be more similar to each other than to any other sequences in the genome. The presence of multiple construct–vector repeats at the targeted locus is rarely discussed in the literature but was a consistent feature in our experiments. The generation of *glk* double mutants hinted at a reason for the multiple inserts. In both Pp*glk1-1 glk2-4* and Pp*glk1-5 glk2-1*, there are multiple inserts in the first targeted locus but only a single insert at the second targeted locus. That is, Pp*glk1-1 glk2-4* has multiple copies of the *nptII* resistance cassette and one copy of the *aphIII* cassette, whereas Pp*glk1-5 glk2-1* has the opposite. Therefore, it is possible that there is selection for multiple copies of the resistance gene after primary transformation experiments.

Pp*glk* Double Mutants Are Pale Green

To examine the phenotype of Pp*glk* loss-of-function mutants, colonies of the wild type, the three lines of Pp*glk1* and Pp*glk2*, and the two lines of Pp*glk1 glk2* were compared. Mutant colonies appeared indistinguishable from the wild type with respect to size, shape, and tissue density. Pp*glk1* and Pp*glk2* single mutants also appeared indistinguishable from the wild type with respect to chlorophyll pigmentation levels (Figures 4A and 4B). However, both Pp*glk1 glk2* double mutant lines appeared pale green (Figures 4A and 4B). This pale green phenotype was observed with both gametophore-dominating (Figure 4A) and protonema-dominating (Figure 4B) colonies. These observations suggested that Pp*Glk1* and Pp*Glk2* function redundantly to promote photosynthetic development in *P. patens*.

To confirm that the Pp*Glk* loci were disrupted in each of the mutant lines, Pp*Glk* and Pp*Tubulin* gene products were amplified from protonemal tissue by RT-PCR (Figure 4C). Whereas Pp*Tubulin* transcripts accumulated in all nine lines, Pp*Glk1* transcripts did not accumulate in Pp*glk1* single mutants or in double mutants. Similarly, Pp*Glk2* transcripts did not accumulate in Pp*glk2* single mutants or in double mutants. Thus, in all six single mutant lines and in both double mutant lines, homologous recombination events produced null alleles.

To quantify the extent to which the loss of GLK function perturbs pigmentation levels, chlorophyll concentrations were measured in wild-type and mutant tissue (Figures 4D and 4E). In Pp*glk1* and Pp*glk2* single mutants, total chlorophyll levels were at least 70% of those seen in the wild type. In double mutants, however, levels were reduced to 40% of wild-type levels in both protonema (Figure 4D) and gametophores (Figure 4E). Therefore, Pp*Glk* genes act redundantly to promote chlorophyll accumulation in *P. patens*. To determine whether chlorophyll *a* and *b* levels are equally affected by the loss of GLK function, the concentration of each pigment was calculated (Figures 4D and 4E). Notably, levels of both chlorophyll *a* and *b* were reduced in lines in which the total chlorophyll level was lower than in the wild type. However, in double mutants, the two chlorophylls were reduced to different extents such that the ratio of *a* to *b* was higher than normal (Figures 4F and 4G). This observation suggests that the conversion of chlorophyll *a* to chlorophyll *b* is impaired in Pp*glk1 glk2* double mutants.

Thylakoid Stacking Is Perturbed in Pp*glk* Double Mutants

The reduced chlorophyll levels observed in Pp*glk1 glk2* double mutants suggested that chloroplast development in *P. patens* is

Figure 2. (continued).

(A) Aligned amino acid sequences of GLK proteins. Proteins are from moss (lines 1 and 2), Arabidopsis (lines 3 and 4), maize (lines 5 and 6), and rice (lines 7 and 8). The sequences are highly conserved over two domains: the DBD (red) and the GCT box (orange). The black line above the sequence indicates the region used for the parsimony analysis shown in **(B)**.

(B) Strict consensus tree of the GARP gene family with bootstrap support indicated above the nodes. The GLK proteins form a monophyletic group shown in red.

(C) Most parsimonious tree of the *GLK* gene family with bootstrap support indicated above the nodes.

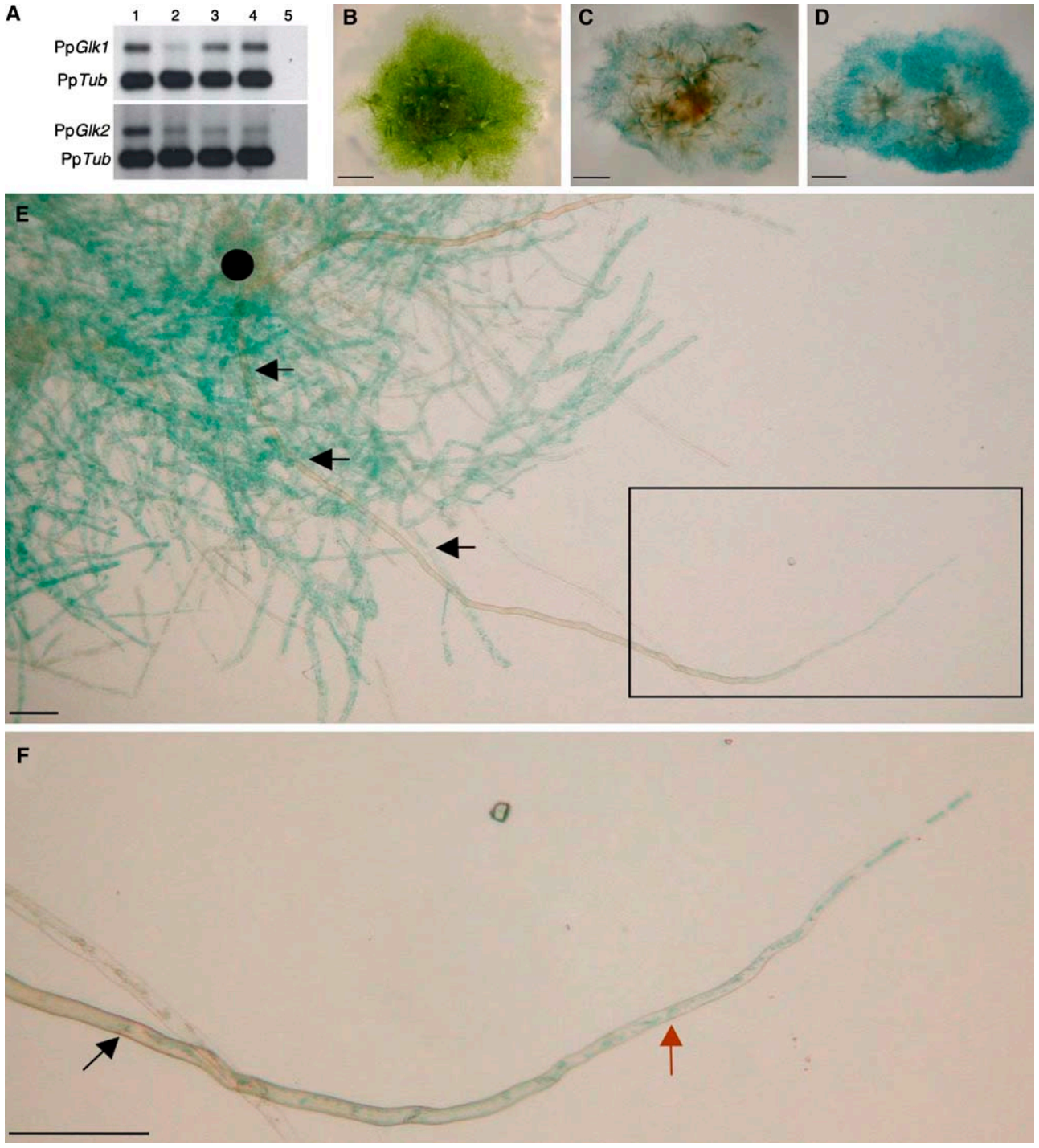

Figure 3. Pp*Glk* Gene Expression Patterns.

(A) RT-PCR analysis of Pp*Glk1* and Pp*Glk2* transcripts. Wild-type tissue was harvested from protonema (lanes 1 and 2) and gametophores (lanes 3 and 4), either in the light (lanes 1 and 3) or in the dark (lanes 2 and 4) period of the diurnal light/dark cycle. Lane 5, water control.
(B) to **(D)** Wild-type **(B)**, Pp*Glk1*:*uidA-3* **(C)**, and Pp*Glk2*:*uidA-3* **(D)** gametophytes stained for GUS activity. Bars = 2 mm.
(E) Pp*Glk2*:*uidA-3* gametophyte stained for GUS activity. An unstained rhizoid filament (arrows) extends from a young gametophore (dot). Bar = 0.1 mm.
(F) Box in **(E)** shown at higher magnification. The rhizoid filament develops into a caulonemal filament, in which GUS activity is detected. The black arrow indicates the cell wall between the rhizoid and cells in the transition zone. The red arrow indicates the cell wall between cells in the transition zone and the caulonemal cell. Bar = 0.1 mm.

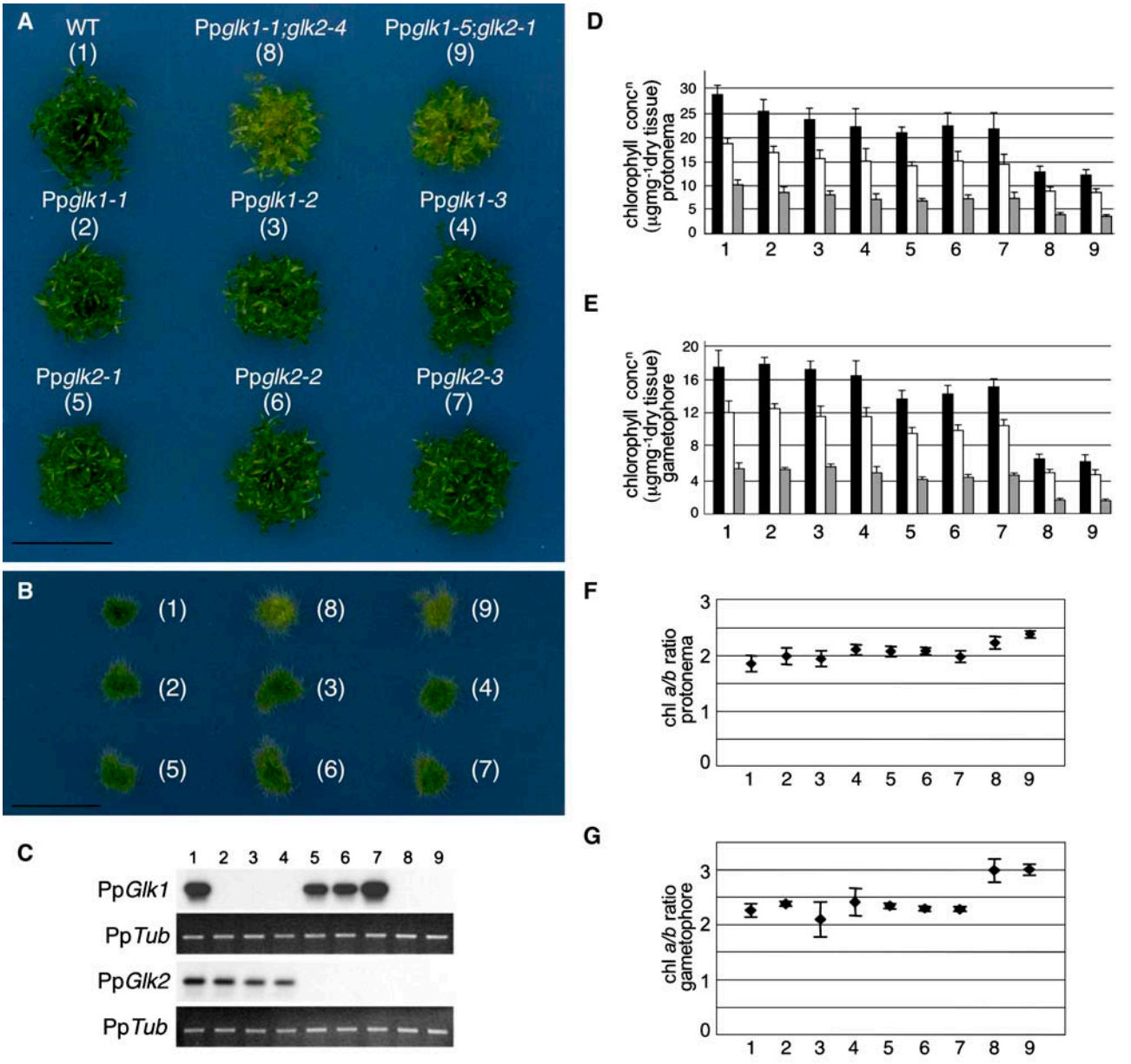

Figure 4. Pp*glk* Mutant Phenotype.

(A) and **(B)** Colony morphology of wild-type and single and double Pp*glk1 glk2* mutant plants in gametophore-dominating **(A)** and protonema-dominating **(B)** colonies. Bars = 1 cm.

(C) RT-PCR analysis of Pp*Glk* and tubulin (Pp*Tub*) transcripts in the lines shown in **(A)**. Lane numbers correspond to colony numbers in **(A)**.

(D) and **(E)** Chlorophyll concentration in protonema **(D)** and gametophores **(E)** of lines shown in **(A)**. Black bars, total chlorophyll; white bars, chlorophyll *a*; gray bars, chlorophyll *b*. Error bars represent SD.

(F) and **(G)** Chlorophyll *a*/*b* ratios in protonema **(F)** and gametophores **(G)** of lines shown in **(A)**. Error bars represent SD.

disrupted by the loss of GLK function. To determine whether chloroplast size and number are perturbed, protonema and gametophores were examined by light microscopy (Figures 5A to 5R). Although chloroplasts in double mutants were more translucent than in other lines, chloroplast size and number appeared similar in wild-type and mutant samples. By contrast, chloroplast ultrastructure was perturbed by the loss of GLK function (Figures 5S to 5V). Thylakoid membranes stack to form recognizable grana in chloroplasts of wild-type and single Pp*glk* mutant plants (Figure 5S, 5U, and 5V), but in Pp*glk1 glk2* double mutants, granal formation is disrupted (Figure 5T). Quantification of the number of thylakoid membrane layers per granal stack revealed an average

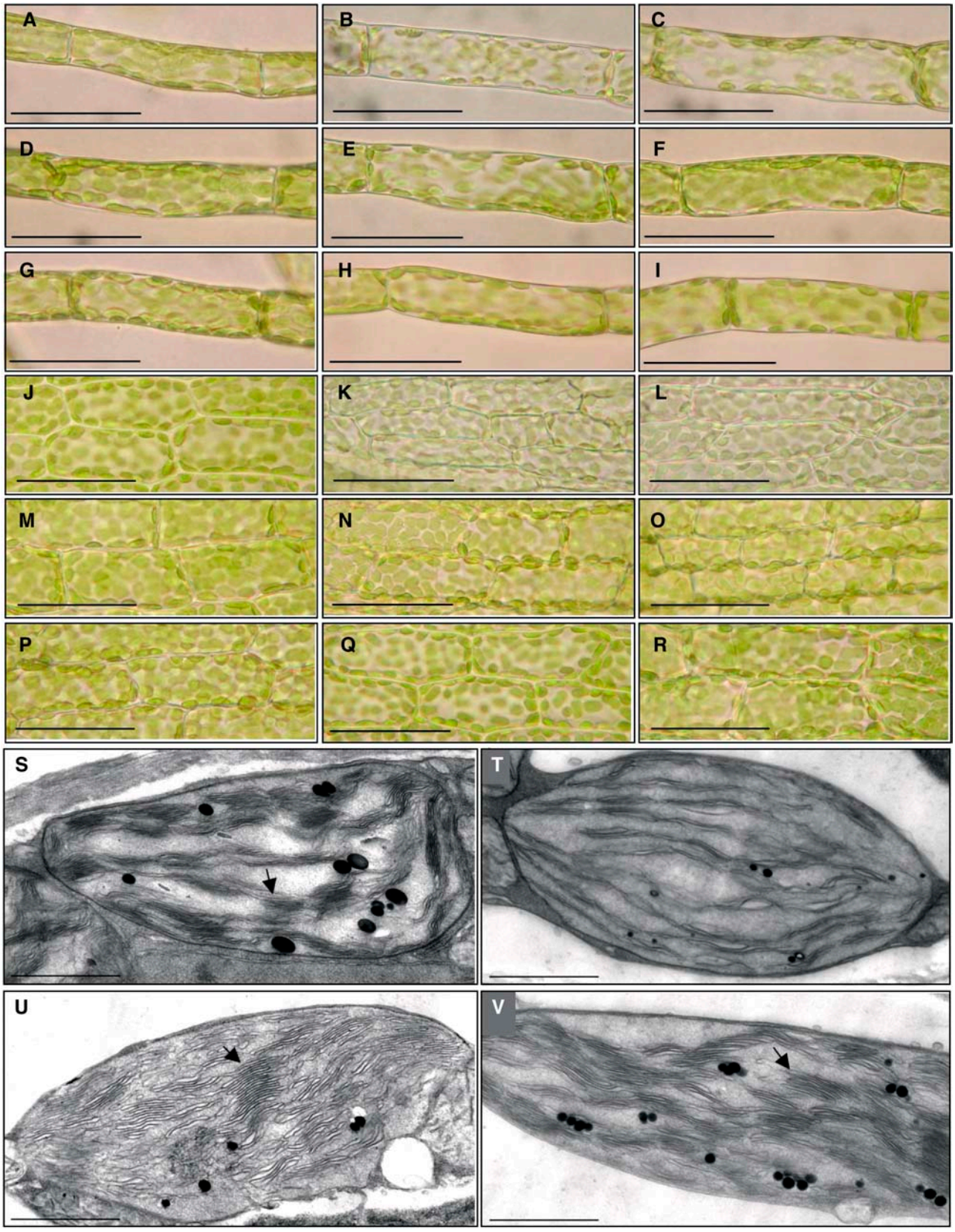

Figure 5. Chloroplast Morphology in Pp*glk* Mutants.

of seven in wild-type chloroplasts and three in double mutant chloroplasts (see Supplemental Figure 5 online). Notably, this phenotypic defect is very similar to that seen in chloroplasts of Arabidopsis At*glk1 glk2* double mutants (Fitter et al., 2002). Thus, *GLK* genes act to promote granal formation in both *P. patens* and Arabidopsis.

Mutant Chloroplasts Accumulate Reduced Levels of Light-Harvesting Chlorophyll *a*/*b* Binding Protein

To determine the extent of phenotypic similarity between *glk* double mutants of *P. patens* and Arabidopsis, markers of chloroplast development were examined. Because thylakoid stacking is reduced in double mutants, we first examined the integrity of the five thylakoid-bound photosynthetic complexes [photosystem I (PSI), PSII, NAD(P)H dehydrogenase (NDH), cytochrome f/b_6, and ATP synthase]. As the level of any individual subunit reflects the integrity of the complex as a whole, PsaD was used as a marker for PSI, D1 was used for PSII, NDH H subunit was used for NDH, cytochrome *f* was used for cytochrome f/b_6, and CF1α was used for ATP synthase. Figure 6A demonstrates that all five complexes are essentially intact in double mutants of both moss and Arabidopsis. Because all of the complexes are composed of both nucleus- and chloroplast-encoded subunits, this observation demonstrates that general chloroplast transcription, translation, and import pathways are functional in *glk* double mutants. Similarly, levels of the carbon fixation enzyme ribulose bisphosphate carboxylase and of the thylakoid assembly protein VESICLE-INDUCING PROTEIN IN PLASTIDS (VIPP1) are normal in both double mutants (Figure 6A). By contrast, levels of the membrane-bound light-harvesting chlorophyll *a*/*b* binding protein (LHCB) are reduced (Figure 6A). Previous work demonstrated that this decrease had a secondary effect on PSII integrity in Arabidopsis (Fitter et al., 2002). Here, only a slight reduction in the level of D1 protein is seen in Arabidopsis double mutants and no obvious reduction is seen in moss double mutants (Figure 6A). However, the reduction is not as dramatic as that seen previously. This observation adds further weight to the argument that PSII defects are an indirect consequence of reduced LHCB and chlorophyll levels as opposed to a direct effect of the loss of GLK function. The difference between the data presented here and those reported previously most likely reflects either the different growth conditions used (greenhouse versus growth chamber) and/or the time of day the tissue was harvested. Regardless, the most important observation in the context of this work is that the moss and Arabidopsis mutants exhibit nearly identical protein profiles. Thus, loss of GLK function leads to similar chloroplast defects in *P. patens* and Arabidopsis.

Reduced levels of LHCB in Arabidopsis *glk* double mutants are accompanied by reduced *LHCb1* and *LHCb6* transcript levels (Fitter et al., 2002). Similarly, reduced chlorophyll levels are accompanied by reduced At*HEMA1* (encoding glutamyl tRNA reductase) and At*CAO* (encoding chlorophyll *a* oxygenase) transcript levels (Fitter et al., 2002) (Figure 7C). To determine whether similar perturbations occur in moss double mutants, Pp*Cab* (homologous with At*LHCb*), Pp*HemA1*, and Pp*Cao* transcript levels were examined. As seen in Arabidopsis, levels of all three transcripts were reduced by at least 60% in *P. patens* double mutants compared with the wild type (Figures 6B and 6C). By contrast, both chloroplast- and nucleus-encoded ribulose bisphosphate carboxylase transcripts accumulated to at least 80% of wild-type levels (Figures 6B and 6C), also as seen in Arabidopsis. These consistent features of double mutant phenotypes in moss and Arabidopsis suggest that GLK function is conserved in bryophytes and angiosperms.

Pp*Glk1* Partially Rescues the Arabidopsis *glk* Double Mutant Phenotype

Although GLK function appears to be conserved in *P. patens* and Arabidopsis, GLK proteins have diverged to the extent that the DBD and the GCT box are the only common domains (Figure 2A). To determine whether the divergent regions are important for species-specific function, the entire Pp*Glk1* open reading frame was introduced into the At*glk1 glk2* double mutant under the control of the constitutive *Cauliflower mosaic virus* 35S promoter. Two independently transformed lines were obtained, both of which contained two independent insertions. In subsequent generations, these insertions segregated independently such that lines 8.2, 8.15, and 12.4 were homozygous for one of three of the insertion alleles isolated (see Supplemental Figure 6 online). For phenotypic characterization, lines 8.14 and 12.3 served as controls. These lines were derived from the two independent transformations followed by segregation away from the transgene in each case.

On a macroscopic level, all three 35S:Pp*Glk1* lines appeared greener than the double mutant but less green than 35S:At*GLK1* lines (Figure 7A). Measurement of chlorophyll concentrations in each line provided further resolution of pigmentation levels and showed that 35S:At*GLK1* lines contained more chlorophyll than the wild type (137%) and that 35S:Pp*Glk1* lines contained between 71 and 88% of wild-type levels (Figure 7B). Thus, the three independent 35S:Pp*Glk1* transgenes at least partially rescue the chlorophyll deficiency in Arabidopsis double mutants.

To determine whether the restored chlorophyll levels in 35S:Pp*Glk1* transformants are correlated with expression of the transgene and with restored levels of At*LHCb*, At*HEMA1*, and

Figure 5. (continued).

(A) to **(R)** Light micrographs of chloronemal (**[A]** to **[I]**) and gametophore (**[J]** to **[R]**) cells showing chloroplast density in the wild type (**[A]** and **[J]**), *glk1-1* (**[D]** and **[M]**), *glk1-2* (**[E]** and **[N]**), *glk1-3* (**[F]** and **[O]**), *glk2-1* (**[G]** and **[P]**), *glk2-2* (**[H]** and **[Q]**), *glk2-3* (**[I]** and **[R]**), *glk1-1 glk2-4* (**[B]** and **[K]**), and *glk1-5 glk2-1* (**[C]** and **[L]**). Bars = 50 μm.

(S) to **(V)** Representative images of chloroplast ultrastructure in the wild type **(S)**, Pp*glk1-5 glk2-1* **(T)**, Pp*glk1-1* **(U)**, and Pp*glk2-2* **(V)**. At least 50 chloroplasts were examined for each line. Arrows point to grana. Bars = 1 μm.

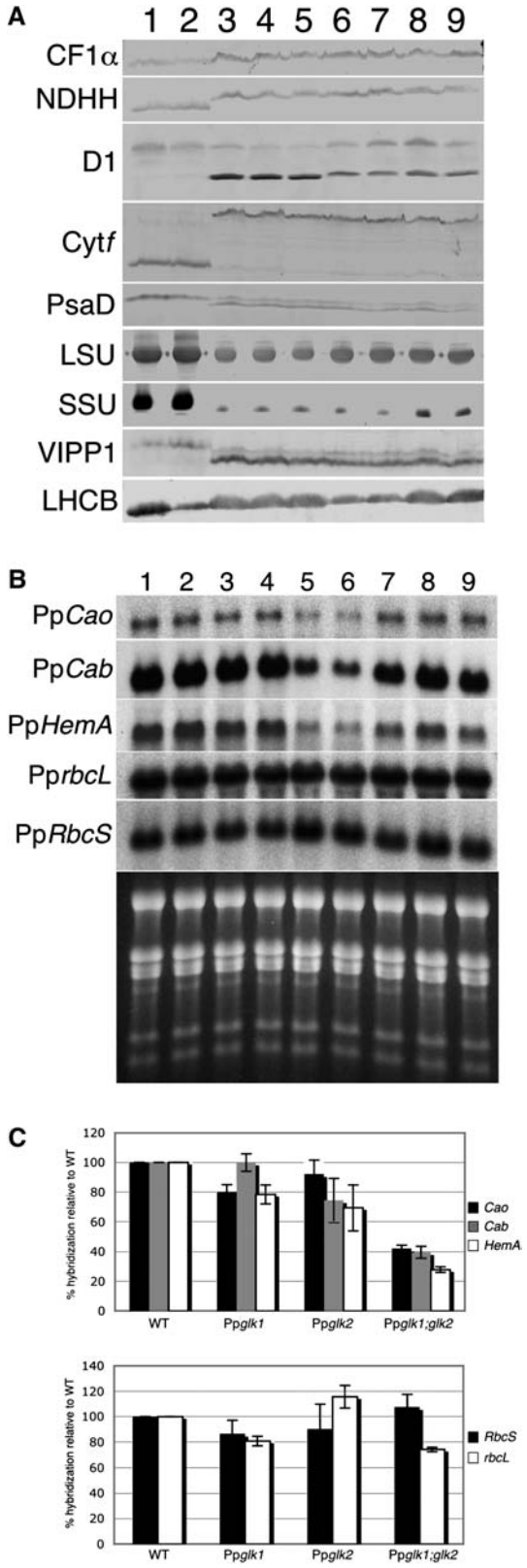

Figure 6. Photosynthetic Defects in Pp*glk* Mutants.

(A) Protein gel blot analysis of Arabidopsis (lanes 1 and 2) and *P. patens* (lanes 3 to 9) proteins. Wild-type Arabidopsis (lane 1) is compared with At*glk1 glk2* double mutant (lane 2). Wild-type moss (lane 3) is compared with Pp*glk1-1* (lane 4), Pp*glk1-2* (lane 5), Pp*glk1-1 glk2-4* (lane 6), Pp*glk1-5 glk2-1* (lane 7), Pp*glk2-1* (lane 8), and Pp*glk2-2* (lane 9). Equal volumes of protein were loaded after extraction at a ratio of 1 mL of buffer per gram of tissue. Some of the proteins migrated at different sizes in Arabidopsis and moss. Where this was the case, the region of the gel encompassing both proteins is shown.

(B) RNA gel blot analysis of photosynthetic gene expression patterns in moss single and double mutant lines. Lanes correspond to wild-type (lane 1), *glk1-1* (lane 2), *glk1-2* (lane 3), *glk1-3* (lane 4), *glk1-1 glk2-4* (lane 5), *glk1-5 glk2-1* (lane 6), *glk2-1* (lane 7), *glk2-2* (lane 8), and *glk2-3* (lane 9). Pp*HemA* encodes glutamyl tRNA reductase, Pp*Cao* encodes chlorophyll *a* oxygenase, Pp*Cab* encodes chlorophyll *a*/*b* binding protein, Pp*rbcL* encodes the large subunit of ribulose bisphosphate carboxylase, and Pp*RbcS* encodes the small subunit. Probes were hybridized to one of four replicate blots. Ethidium bromide fluorescence of rRNA on one of these blots is shown to assess relative loading levels. Hybridization signals were quantified using a phosphor imager and adjusted according to ethidium fluorescence levels of the 18S rRNA on the relevant blot, measured using a Kodak EDAS 290 camera.

(C) Graphic representation of relative hybridization levels for each gene shown in **(B)**. Hybridization relative to the wild type was calculated for each individual moss line. Values were then averaged for each single mutant and for the double mutant. Error bars represent SD.

At*CAO* transcripts, RNA gel blot analysis was performed. Figure 7C demonstrates that transgene expression levels are correlated with the degree of phenotypic recovery, as judged by transcript levels of the marker genes At*LHCb*, At*HEMA1*, and At*CAO* (Figure 7C). However, line 8.2 is anomalous in that the transgene is expressed at very low levels, the marker transcript levels are barely restored, and yet the plants exhibit the highest chlorophyll levels (Figure 7B). We assume that this anomaly is a consequence of the transgene insertion site. The phenotypes of lines 8.15 and 12.4 demonstrate that the 35S:Pp*Glk1* construct complements the Arabidopsis mutant less effectively than the 35S:*AtGLK1* construct. Given the divergence time between the two species, this observation is perhaps not surprising. Presumably, GLK target sequences have coevolved with the GLK protein; thus, the moss protein might bind the Arabidopsis targets with reduced efficacy. However, it is possible that the failure to fully complement the mutant phenotype may be a consequence of reduced translation of the moss protein or of different codon use. Regardless of which explanation is correct, the overall phenotype of the rescued lines demonstrates that the moss gene can partially substitute for the Arabidopsis gene.

DISCUSSION

The bryophyte and vascular plant lineages diverged >400 million years ago (Gifford and Foster, 1989). Since that time, developmental mechanisms have diversified such that distinct lineages exhibit diverse morphologies. Comparative studies have suggested that the *P. patens* and Arabidopsis genomes contain largely similar sets of genes (Nishiyama et al., 2003). However, it is unknown how many of these common genes are regulated similarly in the different developmental contexts of the two

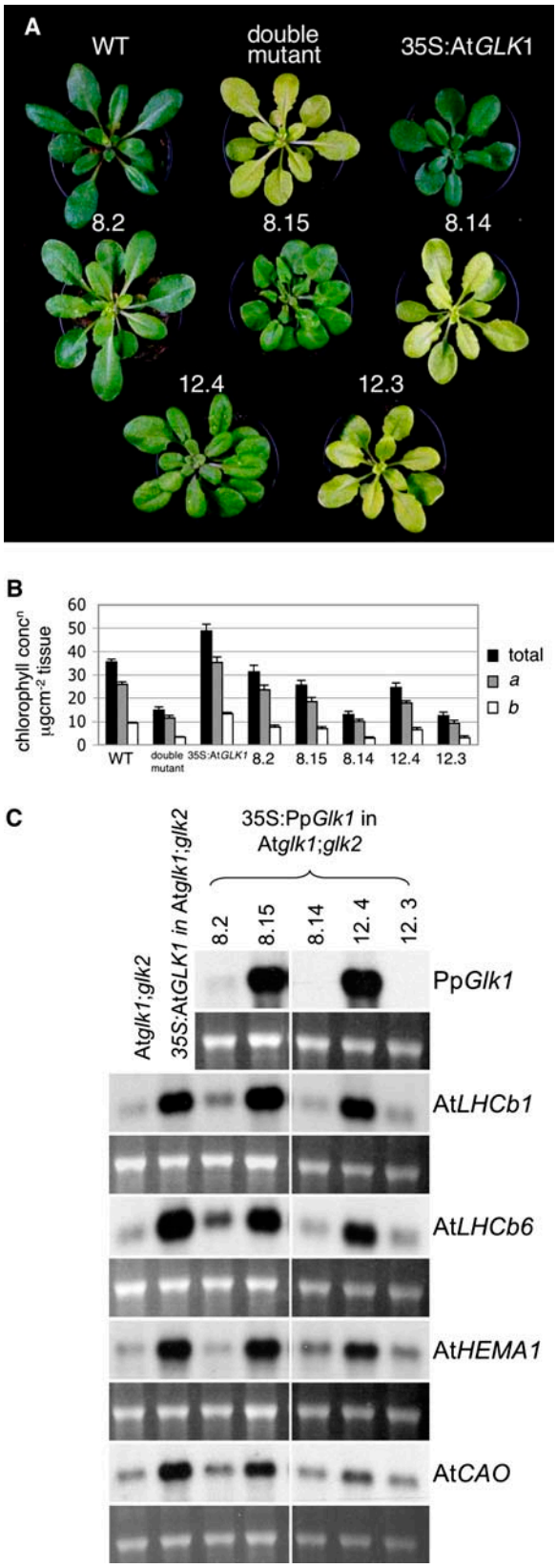

Figure 7. Complementation of At*glk* Mutant Phenotypes.

organisms. The work described here presents an example of a regulatory pathway that is conserved between these two divergent species. In both *P. patens* and Arabidopsis, GLK function is required for correct thylakoid stacking within the chloroplast and for the accumulation of both chlorophyll and LHCB. As such, it is likely that *GLK* gene function was required for chloroplast development in the last common ancestor of bryophytes and flowering plants. Therefore, GLK-mediated regulation of chloroplast development defines one of the most ancient conserved regulatory mechanisms identified in the plant kingdom, other notable examples being microRNA-mediated control of gene expression (Floyd and Bowman, 2004), regulation of state transitions in the chloroplast (Bellafiore et al., 2005), and protoporphyrin signaling from the chloroplast to the nucleus (Falciatore et al., 2005).

Phylogenetic analysis supports independent duplication events for *Glk* genes in *P. patens*, in Arabidopsis, and in monocots (Fitter et al., 2002). Therefore, although redundant *GLK* gene function is a common feature between Arabidopsis and moss, it must have arisen independently. Fully redundant genes are rare, as selection acts to maintain only essential functions. In Arabidopsis, the At*GLK* genes are only partially redundant because the expression of the two genes is differentially regulated. At*GLK1* expression is induced by light, whereas At*GLK2* expression is regulated by both light and the circadian clock (Fitter et al., 2002). Furthermore, At*GLK2* but not At*GLK1* is expressed in siliques. These differences in expression profiles must be sufficient to maintain functional copies of both genes. In *P. patens*, there are no obvious differences in expression profiles between Pp*Glk1* and Pp*Glk2*; however, Pp*Glk* genes are expressed during the haploid stage of the life cycle. For genes that function in haploid cells, there may be additional selective pressure to maintain duplicate copies, particularly if the genes are essential for cell function. An alternative explanation for the presence of two genes invokes a recent duplication event in *P. patens* that has yet to be selected upon. The latter suggestion is supported by the observation that the two *Glk* genes in *P. patens* are much more similar to each other than the two Arabidopsis genes are to each other, even outside of the functional DBD and GCT domains (Figure 2A).

Loss of GLK function in both moss and Arabidopsis perturbs chloroplast development to the extent that granal formation is

(A) Four-week-old seedling phenotypes of the wild type, At*glk1 glk2* double mutant, 35S:At*GLK1* in the double mutant background, and 35S:Pp*Glk1*in the double mutant background (lines 8 and 12). Lines 8.2, 8.15, and 12.4 contain a transgene, whereas lines 8.14 and 12.3 do not.
(B) Chlorophyll concentration in leaves of lines shown in **(A)**.
(C) RNA gel blot analysis of lines shown in **(A)**. At*LHCb1* and At*LHCb6* encode light-harvesting complex proteins, At*HEMA1* encodes glutamyl tRNA reductase, and At*CAO* encodes chlorophyll *a* oxygenase. Ethidium bromide fluorescence of 26S rRNA is shown to assess relative loading levels. Hybridization signals were quantified using a phosphor imager and adjusted according to ethidium fluorescence levels measured using a Kodak EDAS 290 camera. Relative hybridization levels in lanes 1 to 7 were as follows: Pp*Glk1*, (–), (–), 6, 67, 0, 100, 0; At*LhCb1*, 9, 100, 17, 133, 12, 135, 20; At*LHCb6*, 11, 100, 19, 64, 14, 57, 19; At*HEMA1*, 13, 100, 13, 93, 37, 91, 38; At*CAO*, 30, 100, 36, 70, 36, 65, 51.

reduced, LHCBs accumulate to lower levels than normal, and chlorophyll biosynthesis is reduced. Although reduced LHCB and chlorophyll levels have been correlated with reduced thylakoid stacking, reduced LHCB accumulation does not in itself prevent granal formation (Andersson et al., 2003). Therefore, it is possible that perturbations to LHCB and chlorophyll accumulation are a secondary consequence of *glk* mutations and that GLK function is primarily required to facilitate thylakoid stacking. How thylakoids stack to form grana, however, is a debatable topic with little resolution. There are three main proposals. The first states that grana form as an inevitable consequence of surface charges on the thylakoid membrane (Barber, 1980). The second suggests that there has to be heterogeneity of PSI and PSII within the thylakoids (Martienssen et al., 1989). The third results from studies of barley (*Hordeum vulgare*) mutants and invokes the existence of a stabilizing factor that promotes thylakoid stacking (Simpson et al., 1989). Because *GLK* genes encode transcription factors, it is difficult to speculate how gene function could contribute to charges on the membrane. Similarly, because *glk* mutations do not primarily affect the accumulation of protein complexes on the thylakoids, any effect on heterogeneity of PSI and PSII would be indirect. Thus, the most plausible explanation is that GLK proteins regulate the transcription of a thylakoid-stabilizing factor(s). If this were the case, GLK function would facilitate granal formation.

The ability to stack thylakoids into grana arose within the chlorophyte lineage (Song and Gibbs, 1995). Thylakoids exist as single lamellae in prokaryotic cyanobacteria and eukaryotic glaucocystophytes in which the light-harvesting antennae are not composed of LHC proteins but of phycobilisomes. In rhodophytes, LHC protein is present for PSI, but PSII is associated with extrinsic phycobilisomes and thylakoids do not stack (Dodge, 1973). True grana that show segregation of PSI and PSII occur only in land plants and in the green algae that are most closely related to plants (i.e., the charophytes). More distant green algae such as *Chlamydomonas reinhardtii* have appressed thylakoids that are not organized into grana and lack segregation of the two photosystems (Song and Gibbs, 1995; Bertos and Gibbs, 1998). Notably, tBLASTX searches of the *C. reinhardtii* genome database (http://genome.jgi-psf.org) identified five sequences with >50% amino acid identity with the GARP DBD but no sequences similar to the GCT box. The apparent absence of *GLK* genes in *C. reinhardtii* is consistent with the idea that *GLK* gene sequences evolved concurrently with grana. Therefore, we speculate that *GLK* gene function evolved to facilitate thylakoid stacking and the consequent photosystem segregation found in all land plant chloroplasts.

METHODS

Plant Strains and Growth Conditions

Physcomitrella patens subsp *patens* (Engle, 1968) was kindly provided by Celia Knight (University of Leeds, UK). Moss cultures were maintained under sterile conditions on BCD medium (Grimsley et al., 1977) and grown at 25°C with a 16-h-light/8-h-dark cycle at a light intensity of 50 to 60 $\mu mol \cdot m^{-2} \cdot s^{-1}$. *Arabidopsis thaliana* ecotype Columbia was grown at 22°C in a greenhouse with a 16-h-light/8-h-dark cycle.

Gene Isolation

Pp*Glk* gene fragments were amplified from genomic DNA by degenerate PCR using primers YY3-5′ (5′-GTNGAYTGGACNCCNGA-3′), YY4-3′ (5′-ACRTCNCCDATNGCNGCRTC-3′), and YY11 (5′-GCNGTNGARSAR-YTNGG-3′). The remaining coding sequences were obtained by 3′ rapid amplification of cDNA ends using gene-specific primers YY5 (5′-TGG-TCTCGAAGTTCACGGC-3′) for Pp*Glk1* and YY21 (5′-GCCGCATCAT-CCGATATCC-3′) for Pp*Glk2* or by ligation-anchored PCR using gene-specific primers YY14R (5′-TCCAAGCTGTTCCACCGC-3′) and YY18R (5′-CAAGCATCTAGCGAATTCACC-3′) for Pp*Glk1* and YY23R (5′-GGA-GCCTGAGTGTATGTGC-3′) for Pp*Glk2*. Sequences flanking Pp*Glk1* were obtained by thermal asymmetric interlaced PCR using primers YY15R (5′-CTTTCGCCAACAGCAGCG-3′), YY16R (5′-GCTACATT-CACCACTATCGG-3′), and YY18R for 5′ sequences and YY38 (5′-TCGCTGCGTGCAGTAAACCACC-3′), YY8-5′ (5′-GGAGACGCATTTA-TCGAAGG-3′), and YY9-5′ (5′-CCTGTTAGCTTTAGTTGTTTGCTG-3′) for 3′ sequences. Sequences flanking Pp*Glk2* were amplified by inverse PCR from self-ligated *Hin*dIII-digested genomic DNA using primers YY29 (5′-CCCTGTTTTGCAATAAGGCTTCC-3′) and YY35R (5′-AATCTTC-GTGGGTGGTGACC-3′).

Sequence Alignment

Sequence alignments were generated using the PILEUP function of the Genetics Computer Group (Madison, WI) software package, with default settings. Conserved amino acids were highlighted using the PRETTYBOX function. Coloring and sequence alignment were manually modified using CANVAS software (Deneba Systems, Miami, FL).

Phylogenetic Analysis

Fifty-one additional GARP protein sequences, two TEA family proteins (TEC and TEF), potato (*Solanum tuberosum*) MybST1, and Arabidopsis MybST1-like sequences were obtained from GenBank (http://www.ncbi.nlm.nih.gov). Sequences were aligned manually using Se-Al software (version 1.0) (Rambaut et al., 1996), excluding regions of ambiguous alignment (sequence alignments are included in the supplemental data online). Phylogenetic analyses were performed using PAUP (for Phylogenetic Analysis Using Parsimony) software (version 4.0b4; Sinaur Associates, Sunderland, MA). Parsimony analyses were conducted as in Moylan et al. (2004). Multiple most parsimonious trees were obtained, from which a strict consensus tree was computed. The robustness of clades in the strict consensus trees was evaluated by nonparametric bootstrap analysis (Felsenstein, 1985). Bootstrap values were obtained from 100 pseudoreplicates.

For the GARP tree (Figure 2B), parsimony analysis was conducted on aligned DBD amino acid sequences of eight GLK proteins and 51 additional GARP proteins. Fifty-four amino acid characters were included (region indicated in Figure 2A), of which 2 were constant and 52 were parsimony-informative. Myb-related proteins that are not classified as GARP proteins were also included as outgroups, two of which were TEA family members and three of which were MybST1 proteins. The tree was rooted on TEF1. Parsimony analysis resulted in 32 equally most parsimonious trees with lengths of 621 steps. Accession numbers are either indicated on the tree or as follows: AtPHR1, NP_194590; KANADI, AAK59989; MYR1, AAK01148; Psr1, AAD55945 (*Chlamydomonas reinhardtii*); ARR11, CAA06431; ARR1, BAA74528; ARR2, BAA74527; ARR14, AAD12696; ARR10, CAA16597; PpGLK2, AAV54521; PpGLK1, AAV54520; OsGLK2, AAK50393 (*Oryza sativa*); OsGLK2, AAK50394 (*O. sativa*); ZmGLK1, AAK50392 (*Zea mays*); G2, AAG32325 (*Z. mays*); AtGLK1, AAK20120; AtGLK2, AAK20121; APRR2, CAA17145; ARR13, AAC77865; MybST1, S51839 (*S. tuberosum*); MybST1-LIKE, NP_177158;

MybST1-LIKE2, CAB78068; TEC1, CAA85028 (*Saccharomyces cerevisiae*); TEF1, AAB00791 (*Homo sapiens*). Unless indicated otherwise, all proteins are from Arabidopsis.

For the GLK tree (Figure 2C), parsimony analysis was conducted on cDNA sequences aligned across the DBD and the C-terminal region of eight *GLK* genes and four other GARP genes. A total of 396 cDNA characters were included in the analysis, of which 89 were constant, 76 were variable, and 231 were parsimony-informative. The tree was rooted on AL162651. Parsimony analysis resulted in a single most parsimonious tree with a length of 937 steps. With respect to relationships within the GLK family, the data in Figure 2C are more informative than those in Figure 2B, because more characters were used in the analysis shown in Figure 2C.

Moss Transformation

Moss was transformed after polyethylene glycol–mediated DNA uptake into protoplasts, essentially according to Schaefer et al. (1991). The constructs used are illustrated in Supplemental Figures 3A and 4A online. Protoplasts were regenerated for 5 d on BCD medium and then transferred to fresh media containing antibiotics for selection (G418 disulfate at 50 μg/mL or hygromycin B at 20 μg/mL). After 2 weeks, colonies were transferred back to nonselective BCD medium for another 2 weeks. Stable transformants were subsequently selected by plating once again on selective media. Transformants were screened by PCR to check for construct integration at the target site. Selected colonies were further tested by DNA gel blot analysis to determine whether they contained multiple copies of the construct and, if so, whether the extra construct sequences were integrated at nontarget sites.

GUS Staining

A spot inoculum of moss protonema was grown for 18 d on BCD medium, subtended by a cellophane disc. After that time, the disc was removed from the plate and the tissue was fixed and stained for GUS activity essentially according to Imaizumi et al. (2002).

Transmission Electron Microscopy

Seven-day-old protonema were harvested 8 h into the light period, prefixed in 4% paraformaldehyde and 3% glutaraldehyde for 5 h, and then fixed in 1.33% OsO_4 for 2 h. After dehydration through an acetone series, samples were embedded in TAAB resin and sectioned using a glass knife on a Sorvall MT-2 ultramicrotome (Sorvall, Newtown, CT). Sections were stained with 0.2% lead citrate and examined with a Zeiss (LEO) Omega 912 electron microscope (Zeiss; LEO Electron Microscope, Oberkochen, Germany) equipped with a Proscan cooled slow-scan charge-coupled device camera (2048 × 2048 pixels). All digital images were captured using the integrated SIS image-analysis package (Soft Imaging Software, Münster, Germany).

DNA and RNA Analysis

DNA was extracted from moss colonies using cetyl-trimethyl-ammonium bromide as described at http://biology4.wustl.edu/moss/methods.php and from Arabidopsis using urea-based buffer as in Chen and Dellaporta (1994). RNA was isolated, and DNA and RNA gel blots were prepared and hybridized as described by Langdale et al. (1988) using gene-specific probes as follows: At*GLK1*, 225-bp fragment corresponding to positions 217 to 441 of accession number AY026772; Pp*Glk1* 5′ region, 245-bp fragment corresponding to positions 1027 to 1271 of accession number AY741684; Pp*Glk1* 3′ region, 292-bp fragment corresponding to positions 1927 to 2218 of accession number AY741684; Pp*Cao*, 360-bp fragment corresponding to positions 57 to 516 of accession number BJ172430; Pp*Cab*, 248-bp fragment corresponding to positions 2147 to 2394 of accession number M23532; Pp*HemA*, 420-bp fragment amplified using primers PpHemA-F (5′-GCGTGTACAACGTAGACGAC-3′) and PpHemA-R (5′-TGGCTTTTGCCTCCACAAGC-3′), which were deduced from sequence data available in physcobase (http://moss.nibb.ac.jp; clones pph27g23 and pph33m12); Pp*RbcS*, 160-bp fragment corresponding to positions 57 to 216 of accession number X76634; Pp*rbcL*, 346-bp fragment corresponding to positions 255 to 600 of accession number X74156; At*LHCb1*, EST AP2L15e09R; At*LHCb6*, EST 23C1T7; At*CAO*, EST 103D24T7; At*HEMA1*, a gift from M. Terry (McCormac and Terry, 2002). Hybridization signals were quantified using an imaging plate (BAS-MS 2025; Fujifilm, Tokyo, Japan) in a Bio-Rad FX molecular imager and Quantity One software (Bio-Rad, Hercules, CA). Signals were standardized to the level of ethidium bromide fluorescence of the 26S or 18S rRNA (measured using a EDAS-290 camera from Kodak [Rochester, NY]) and are presented as relative arbitrary units.

Protein Analysis

Two hundred to 300 mg of moss protonemal tissue and 300 mg of Arabidopsis leaf tissue was harvested 6 h into the light period, flash-frozen, and ground in liquid N_2. Homogenization buffer (18% sucrose, 10 mM $MgCl_2$, 100 mM Tris-HCl, pH 8, and 40 mM β-mercaptoethanol) was added at a ratio of 1 mL/g tissue and the mix was filtered through cheesecloth. After centrifugation at 13,000*g* for 10 min, stop dye (1% SDS, 0.1% bromphenol blue, 10 mM EDTA, and 20% Ficoll) was added to the supernatant at a ratio of 0.33:1. The pellet was resuspended in pellet buffer (2% SDS, 6% sucrose, and 40 mM β-mercaptoethanol) at a ratio of 550 μL/mL decanted supernatant. Finally, stop dye was added to the suspension at a ratio of 0.16:1. Gel blot analysis was performed as described by Fitter et al. (2002). LHCB antibody was raised against maize protein and was a gift from W. Taylor (Commonwealth Scientific and Industrial Research Organization, Canberra, Australia). Ribulose-1,5-bis-phosphate carboxylase/oxygenase antibody was raised against wheat protein and was a gift from J. Gray (Cambridge University, UK). PsaD, CF1α, and cytochrome *f* antibodies were raised against maize proteins and were a gift from A. Barkan (University of Oregon, Eugene, OR). VIPP1 antibody was a gift from J. Soll (University of Kiel, Germany), and NDH H antibody was a gift from K. Steinmuller (University of Dusseldorf, Germany). D1 antibody was a gift from P. Nixon (Imperial College London, UK).

RT-PCR

cDNA was generated using Superscript II reverse transcriptase–treated (Invitrogen) and DNaseI-treated (amplification grade; Invitrogen) RNA. The data shown in Figure 3 were obtained by multiplex PCR amplification of cDNA using 5 μM digoxigenin-labeled dUTP in addition to 0.2 mM deoxynucleotide triphosphate (17 cycles for Pp*Glk1* and 18 cycles for Pp*Glk2*). Each reaction contained two pairs of primers, one for Pp*Glk* (Pp*Glk1*, 5′-GATAAGCAGGGAAGAGGGTG-3′ and 5′-TAGTGCCTAACGAAACTCGC-3′; Pp*Glk2*, 5′-AGGACAACTCGTCCTCGTG-3′ and 5′-AGTGGATATCGGATGATGCG-3′) and the other for a tubulin fragment (5′-TGCTGCTGGATAATGAAGCG-3′ and 5′-CGTGCTGTTCGAAATCATGC-3′). After gel electrophoresis, fragments were transferred to Nytran filters (Schleicher & Schuell, Keene, NH) and visualized using the detection method according to the manufacturer's instructions (Roche, Indianapolis, IN). The RT-PCR data in Figure 4 were also obtained by multiplex PCR amplification (20 cycles) with the primers described above. In this case, however, tubulin fragments were visualized by ethidium bromide staining of the gel and Pp*Glk* fragments were detected by gel blot analysis using the 140-bp fragment corresponding to positions 1349 to 1488 of accession number AY741684 (which hybridizes with both Pp*Glk1* and Pp*Glk2* PCR products).

Chlorophyll Assays

Protonema and gametophores were harvested 2 h into the light period and freeze-dried for 4 h using a Freeze Dryer Modulyo (Edwards, Crawley, UK). A total of 0.4 mg of protonemal tissue or 0.5 mg of gametophore tissue was ground in liquid nitrogen and added to 1 mL of 80% acetone. After a brief spin, the absorbance of the supernatant was measured at 645 and 663 nm. The amount of chlorophyll in each sample was calculated using the following formulae: total chlorophyll concentration (μg/mL) = $OD_{645} \times 20.2 + OD_{663} \times 8.02$; chlorophyll *a* concentration (μg/mL) = $OD_{645} \times 12.7 + OD_{663} \times 2.69$; chlorophyll *b* concentration (μg/mL) = $OD_{645} \times 22.9 + OD_{663} \times 4.68$ (Arnon, 1949). The values obtained were corrected to obtain chlorophyll concentrations in micrograms per milligram of dry tissue. For Arabidopsis, a 0.5-cm^2 leaf disc was ground and the chlorophyll concentration was corrected to micrograms per square centimeter of leaf disc.

Arabidopsis Transformation

Arabidopsis plants were transformed with a pBin+ (van Engelen et al., 1995) construct containing the entire Pp*Glk1* coding sequence driven by the *Cauliflower mosaic virus* 35S promoter (pKIKI 56). Plants were transformed as described previously (Clough and Bent, 1998). Transformed plants were selected on the basis of kanamycin resistance and were self-pollinated to generate T2, T3, and T4 populations that segregated the transgenes.

Sequence data from this article have been deposited with the GenBank data library under accession numbers AY741684 (Pp*Glk1*) and AY741685 (Pp*Glk2*).

ACKNOWLEDGMENTS

We thank Celia Knight and the University of Leeds transformation service for helping us to establish moss work in our laboratory, Daphne Stork for excellent technical assistance, and John Baker for photography. The advice on transmission electron microscopy from Hugh Dickinson and Michael Shaw was invaluable. We are grateful for the support and encouragement of all members of the laboratory throughout the course of this work. Additional thanks to Miltos Tsiantis, Mark Waters, and Mark Fricker for constructive comments on the manuscript. This project was funded by a University of Oxford Kobe Scholarship to Y.Y. and by Biotechnology and Biological Science Research Council and Gatsby Charitable Foundation grants to J.A.L.

Received April 6, 2005; revised May 9, 2005; accepted May 10, 2005; published May 27, 2005.

REFERENCES

Andersson, J., Wentworth, M., Walters, R.G., Howard, C.A., Ruban, A.V., Horton, P., and Jansson, S. (2003). Absence of the Lhcb1 and Lhcb2 proteins of the light-harvesting complex of photosystem II: Effects on photosynthesis, grana stacking and fitness. Plant J. **35,** 350–361.

Arnon, D.I. (1949). Copper enzymes in isolated chloroplast polyphenoloxidase in *Beta vulgaris*. Plant Physiol. **24,** 1–15.

Barber, J. (1980). Membrane surface charges and potentials in relation to photosynthesis. Biochim. Biophys. Acta **594,** 253–308.

Bellafiore, S., Barneche, F., Peltier, G., and Rochaix, J.D. (2005). State transitions and light adaptation require chloroplast thylakoid protein kinase STN7. Nature **433,** 892–895.

Bertos, N.R., and Gibbs, S.P. (1998). Evidence for a lack of photosystem segregation in *Chlamydomonas reinhardtii* (Chlorophyceae). J. Phycol. **34,** 1009–1016.

Blankenship, R.E. (2001). Molecular evidence for the evolution of photosynthesis. Trends Plant Sci. **6,** 4–6.

Chen, J., and Dellaporta, S.L. (1994). Urea-based plant DNA miniprep. In The Maize Handbook, M. Freeling and V. Walbot, eds (New York: Springer Verlag), pp. 526–528.

Clough, S.J., and Bent, A.F. (1998). Floral dip: A simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. Plant J. **16,** 735–743.

Dodge, J.D. (1973). The Fine Structure of Algal Cells. (London: Academic Press).

Douglas, S.E. (1998). Plastid evolution: Origins, diversity, trends. Curr. Opin. Genet. Dev. **8,** 655–661.

Engle, P.P. (1968). The induction of biochemical and morphological mutants in the moss *Physcomitrella patens*. Am. J. Bot. **55,** 438–446.

Falciatore, A., Merendino, L., Barneche, F., Ceol, M., Meskauskiene, R., Apel, K., and Rochaix, J.D. (2005). The FLP proteins act as regulators of chlorophyll synthesis in response to light and plastid signals in Chlamydomonas. Genes Dev. **19,** 176–187.

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. Evolution **39,** 783–791.

Fitter, D.W., Martin, D.J., Copley, M.J., Scotland, R.W., and Langdale, J.A. (2002). *GLK* gene pairs regulate chloroplast development in diverse plant species. Plant J. **31,** 713–727.

Floyd, S.K., and Bowman, J.L. (2004). Gene regulation: Ancient microRNA target sequences in plants. Nature **428,** 485–486.

Gao, H., Kadirjan-Kalbach, D., Froehlich, J.E., and Osteryoung, K.W. (2003). ARC5, a cytosolic dynamin-like protein from plants, is part of the chloroplast division machinery. Proc. Natl. Acad. Sci. USA **100,** 4328–4333.

Gifford, E.M., and Foster, A.S. (1989). Morphology and Evolution of Vascular Plants. (New York: W.H. Freeman).

Grimsley, N.H., Ashton, N.W., and Cove, D.J. (1977). The production of somatic hybrids by protoplast fusion in the moss, *Physcomitrella patens*. Mol. Gen. Genet. **154,** 97–100.

Hall, L.N., Rossini, L., Cribb, L., and Langdale, J.A. (1998). GOLDEN2: A novel transcriptional regulator of cellular differentiation in the maize leaf. Plant Cell **10,** 925–936.

Imaizumi, T., Kadota, A., Hasebe, M., and Wada, M. (2002). Cryptochrome light signals control development to suppress auxin sensitivity in the moss *Physcomitrella patens*. Plant Cell **14,** 373–386.

Langdale, J.A., and Kidner, C.A. (1994). *bundle sheath defective*, a mutation that disrupts cellular differentiation in maize leaves. Development **120,** 673–681.

Langdale, J.A., Rothermel, B.A., and Nelson, T. (1988). Cellular patterns of photosynthetic gene expression in developing maize leaves. Genes Dev. **2,** 106–115.

Martienssen, R.A., Barkan, A., Freeling, M., and Taylor, W.C. (1989). Molecular cloning of a maize gene involved in photosynthetic membrane organization that is regulated by *Robertson's Mutator*. EMBO J. **8,** 1633–1640.

Martin, W., Rujan, T., Richly, E., Hansen, A., Cornelsen, S., Lins, T., Leister, D., Stoebe, B., Hasegawa, M., and Penny, D. (2002). Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc. Natl. Acad. Sci. USA **99,** 12246–12251.

McCormac, A.C., and Terry, M.J. (2002). Light-signalling pathways leading to the co-ordinated expression of *HEMA1* and *Lhcb*

during chloroplast development in *Arabidopsis thaliana*. Plant J. **32,** 549–559.

Moreira, D., Le Guyader, H., and Philippe, H. (2000). The origin of red algae and the evolution of chloroplasts. Nature **405,** 69–72.

Moylan, E.C., Bennett, J.R., Carine, M.A., Olmstead, R.G., and Scotland, R.W. (2004). Phylogenetic relationships among *Strobilanthes* s.l. (Acanthaceae): Evidence from ITS nrDNA, *trnL-F* cpDNA, and morphology. Am. J. Bot. **91,** 724–735.

Nishiyama, T., Fujita, T., Shin, I.T., Seki, M., Nishide, H., Uchiyama, I., Kamiya, A., Carninci, P., Hayashizaki, Y., Shinozaki, K., Kohara, Y., and Hasebe, M. (2003). Comparative genomics of *Physcomitrella patens* gametophytic transcriptome and *Arabidopsis thaliana*: Implication for land plant evolution. Proc. Natl. Acad. Sci. USA **100,** 8007–8012.

Osteryoung, K.W., Stokes, K.D., Rutherford, S.M., Percival, A.L., and Lee, W.Y. (1998). Chloroplast division in higher plants requires members of two functionally divergent gene families with homology to bacterial *ftsZ*. Plant Cell **10,** 1991–2004.

Rambaut, A., Grassly, N.C., Nee, S., and Harvey, P.H. (1996). Bi-De: An application for simulating phylogenetic processes. Comput. Appl. Biosci. **12,** 469–471.

Riechmann, J.L., et al. (2000). Arabidopsis transcription factors: Genome-wide comparative analysis among eukaryotes. Science **290,** 2105–2110.

Rossini, L., Cribb, L., Martin, D.J., and Langdale, J.A. (2001). The maize *golden2* gene defines a novel class of transcriptional regulators in plants. Plant Cell **13,** 1231–1244.

Schaefer, D., Zryd, J.P., Knight, C.D., and Cove, D.J. (1991). Stable transformation of the moss *Physcomitrella patens*. Mol. Gen. Genet. **226,** 418–424.

Schaefer, D.G. (2001). Gene targeting in *Physcomitrella patens*. Curr. Opin. Plant Biol. **4,** 143–150.

Simpson, D.J., Vallon, O., and von Wettstein, D. (1989). Freeze-fracture studies on barley plastid membranes. VIII. In *viridis*^{113}, a mutant completely lacking photosystem II, oxygen evolution enhancer 1 (OEE1) and the alpha-subunit of cytochrome b-559 accumulate in appressed thylakoids. Biochim. Biophys. Acta **975,** 164–174.

Song, X.-Z., and Gibbs, S.P. (1995). Photosystem I is not segregated from photosystem II in the green alga *Tetraselmis subcordiformis*: An immunogold and cytochemical study. Protoplasma **189,** 267–280.

van Engelen, F.A., Molthoff, J.W., Connor, A.J., Nap, J.P., Periera, A., and Stiekema, W.J. (1995). pBINPLUS: An improved plant transformation vector based on pBIN19. Transgenic Res. **4,** 288–290.