Supporting information

Puthiyaveetil *et al.* 10.1073/pnas.0803928105

SI Text

Genotyping of SALK T-DNA Lines. The phenotypes identified and described for chloroplast sensor kinase (CSK) are based on two completely characterized T-DNA lines, SALK_027360 and SALK_018074, which harbor insertions in the At1g67840 locus encoding the CSK protein. Seeds for SALK_027360 and SALK_018074 lines were obtained from the European Arabidopsis Stock Center (Nottingham), sown on soil and the F_1 plants were selfed. A PCR based approach was used to identify homozygous insertion lines among the F_1 plants. Seeds obtained from the genotyped, F_1 homozygous plants were used to characterize the phenotype. The PCR based genotyping approach involved a genomic PCR using genomic primers, a second PCR using genomic and T-DNA left border primers and a third reverse transcriptase PCR (RT-PCR) to confirm the absence of transcripts from the At1g67840 locus. The positions of the T-DNA insertion in SALK_027360 and SALK_018074 lines are indicated in supporting information (SI) Fig. S1. Fig. S1*B* shows results from a genomic PCR using 5′ TGGCCTCTTTTAGCTATGGGGA 3′ as the forward primer and 5′ TGCTCAAGACAAAGCCGTTGA 3′ as the reverse primer. The wild type *CSK* gene was amplified from the wild type sample and not from the SALK_027360 sample indicating T-DNA insertion. Fig. S1*C* shows results from a genomic PCR using 5′ TGGCCTCTTTTAGCTATGGGGA 3′ as forward primer and 5′ GCGTGGACCGCTTGCTGCAACT 3′ as reverse primer, which is also the T-DNA cassette left border primer. The length of the amplicon from this second PCR is indicative of the approximate insertion site in the SALK_027360. The actual insertion site was further determined by sequencing the amplificate from this PCR. Fig. S1*D* shows results from a reverse transcriptase PCR using 5′ GAGAGTTTCAGTCTCAGCCACA 3′ as forward primer and 5′ TTGCAATCAATTTTGTTCAAGTC 3′as reverse primer. The *CSK* mRNA was amplified only from the wild type and not from the SALK_027360 line, thus confirming that the At1g67840 locus was not transcribed in the SALK_027360 line. Fig. S1*E* shows results from a genomic PCR using 5′ GTAGAGTTCACACAGATGATTGAGAAA 3′ as the forward primer and 5′ GCTTCATTGGCTTCAGATACTGC 3′ as the reverse primer. The wild type *CSK* gene is amplified from the wild type sample and not from the SALK_018074 sample, indicating T-DNA insertion. Fig. S1*F* shows results from a genomic PCR using 5′ GCGTGGACCGCTTGCTGCAACT 3′ as the forward primer, which is also the T-DNA cassette left border primer, and 5′ GCTTCATTGGCTTCAGATACTGC 3′ as the reverse primer. The length of the amplificate from this PCR is indicative of the approximate insertion site in the SALK_018074. The actual insertion site was further determined by sequencing the amplificate from this PCR. Fig. S1*G* shows results from a reverse transcriptase PCR using 5′ ATGCTTCTTTCTGCAATCGC 3′ as the forward primer and 5′ CTATGCTTCATTGGCTTCAG 3′as the reverse primer. The *CSK* mRNA was amplified only from the wild type and not from the SALK_018074 line, thus confirming that the At1g67840 locus is not transcribed in the SALK_018074 line.

Conserved Sequence Features of the Chloroplast Sensor Kinase (CSK). Sequence prediction based on SMART, Pfam and InterPro database searches showed conserved functional domains in CSK (Fig. S3). These include an ATP-binding domain (HATPase_c) toward the C terminus, and a domain characteristic of a site of histidine autophosphorylation site and dimerisation (HisKA), consistent with a function for CSK as a sensor kinase. It should be mentioned that the HisKA domain in CSKs of *A. thaliana*, *P. patens* and *O. lucimarinus*, although recognizable in SMART, Pfam, and InterPro databases, is not predicted by the secondary database Prosite. N-terminal to the HisKA domain is a GAF domain, which presumably forms the redox sensor input domain of CSKs. GAF domain and the related PAS redox sensor input domain are known to sense redox signals via redox-active prosthetic groups such as heme (1) or flavin adenine dinucleotide (FAD) (2). The Prosite database predicts a nucleotide-binding motif in the GAF domain of the *Arabidopsis* CSK. This observation is consistent with the nucleotide binding properties of some GAF domains (3) and raises the interesting possibility that the GAF domain in CSK binds one or more redox-responsive cofactors such as uridine or flavin nucleotides for its redox-sensing function. Some redox sensor kinases are also known to employ redox-responsive cysteine residues, residing within the PAS/GAF domains (4, 5) or seen as part of other separate redox sensor input domains (6) for sensing redox signals. A conserved cysteine residue is seen within the GAF domain of plant and cyanobacterial CSKs (Fig. S2). While GAF domains of CSKs in *Ostreococcus*, *Phaeodactylum*, and a few *Synechococcus* species do not seem to have this cysteine conserved (Fig. S2), other positionally conserved cysteines are present in their GAF domains (Fig. S2).

Conserved Functional Domains in CSK. Fig. S3 shows conserved functional domains in CSK and representative algal and cyanobacterial homologues (redrawn from SMART database predictions). Domain denotations: HATPase_c, ATP-binding domain; HisKA, site of histidine autophosphorylation and dimerisation; GAF, sensor domain. For the *Ostreococcus* and the *Physcomitrella* CSKs, the predicted GAF domain is below the curated threshold score of SMART database. For the *Arabidopsis* CSK, SMART database does not predict a GAF domain, even though the *Arabidopsis* CSK sensor domain shows significant sequence homology to the GAF domain of cyanobacterial Hik2. Domain boundary of GAF domains, which fall below the curated threshold score, or not predicted by SMART database, are shown in broken lines. The chloroplast-targeting signal is represented as a white rectangle at the N terminus.

Sequence Features of CSK Suggest Mechanisms of Redox Sensing. Further sequence prediction based on SMART database searches shows a conserved GAF domain in cyanobacterial CSK homologues and in algal and plant CSKs (Fig. S2 and Fig. S3). A GAF domain is a small ligand-binding domain first described for vertebrate c*G*MP specific phosphodiesterase, a cyanobacterial *A*denylate cyclase and the bacterial formate hydrogen lyase transcription activator *F*hlA. GAF domains are known to form sensor domains in a number of histidine kinases. GAF domains of redox histidine sensor kinases sense redox signals via a redox-active heme prosthetic group (7) or by a conserved redox-responsive cysteine residue (5). GAF domains also show homology and functional overlap with the well characterized redox sensor input domain, PAS (1, 8). A redox sensing and transcriptional regulatory role for CSK in chloroplasts is strongly suggested by its redox sensing sequence features (Fig. S2 and Fig. S3); by the observed signaling phenotypic effect of CSK (Fig. 4); and by the functional and phylogenetic relatedness of CSK to the known transcriptional regulatory two-component systems of non-green algae. Our proposed redox-sensing role for CSK is

further supported by the demonstration that the cyanobacterial homologue of CSK, Hik2, interacts with the known redox response regulator, RppA (9) and confers tolerance of photosystem II to environmental stress (10). The cognate response regulator partner of CSK in higher plant chloroplasts remains to be elucidated. However, we can conclude that, like CSK itself, any chloroplast response regulator (CRR) is now nuclear encoded and imported into chloroplasts as a protein precursor. A modified response regulator-like protein, TCP34 is imported as the product of nuclear gene and specifically binds promoters of reaction center genes in spinach chloroplasts (11).

1. Sardiwal S, *et al.* (2005) GAF domain in the hypoxia/NO-inducible *Mycobacterium tuberculosis* DosS protein binds haem. *J Mol Biol* 353:929–936.
2. Taylor BL (2007) Aer on the inside looking out: Paradigm for a PAS-HAMP role in sensing oxygen, redox and energy. *Mol Microbiol* 65:1415–1424.
3. Zoraghi R, Corbin JD, Francis SH (2004) Properties and functions of GAF domains in cyclic nucleotide phosphodiesterases and other proteins. *Mol Pharmacol* 65:267–278.
4. Malpica R, Franco B, Rodriguez C, Kwon O, Georgellis D (2004) Identification of a quinone-sensitive redox switch in the ArcB sensor kinase. *Proc Natl Acad Sci USA* 101:13318–13323.
5. Vuillet L, *et al.*(2007) Evolution of a bacteriophytochrome from light to redox sensor. *EMBO J* 26:3322–3331.
6. Swem LR, *et al.* (2003) Signal transduction by the global regulator RegB is mediated by a redox-active cysteine. *EMBO J* 22:4699–4708.
7. Kumar A, Toledo JC, Patel RP, Lancaster JR, Jr, Steyn AJ (2007) *Mycobacterium tuberculosis* DosS is a redox sensor and DosT is a hypoxia sensor. *Proc Natl Acad Sci USA* 104, 11568–11573.
8. Ho YS, Burden LM, Hurley JH (2000) Structure of the GAF domain, a ubiquitous signaling motif and a new class of cyclic GMP receptor. *EMBO J* 19:5288–5299.
9. Sato S, *et al.* (2007) A large-scale protein protein interaction analysis in *Synechocystis* sp. PCC6803. *DNA Res* 14:207–216.
10. Mikami K, *et al.* (2003) Histidine kinases, Hik2, Hik16 and Hik33, in *Synechocystis* sp. PCC 6803 are involved in the tolerance of photosystem II to environmental stress. *Plant Cell Physiol* 44:S82.
11. Weber P, *et al.* (2006) TCP34, a nuclear-encoded response regulator-like TPR protein of higher plant chloroplasts. *J Mol Biol* 357:535–549.

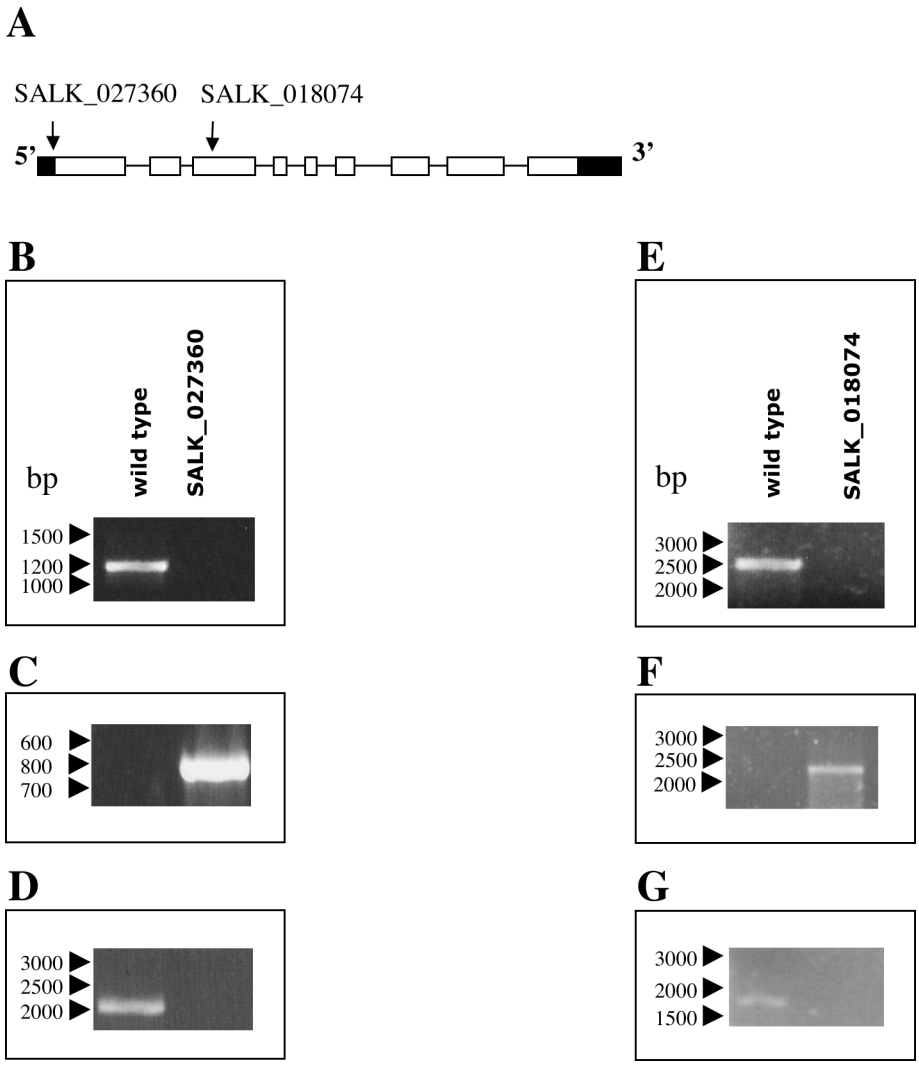

Fig. S1. (*A*) Schematic representation of the gene region of CSK indicating T-DNA insertion sites in two different SALK lines. Exons are represented as white rectangles and introns as the lines connecting them. The UTR regions at either end of the transcript are shown as filled rectangles. (*B–G*) Results from the genomic and the reverse transcriptase PCR methods used for genotyping SALK_027360 and SALK_018074 lines.

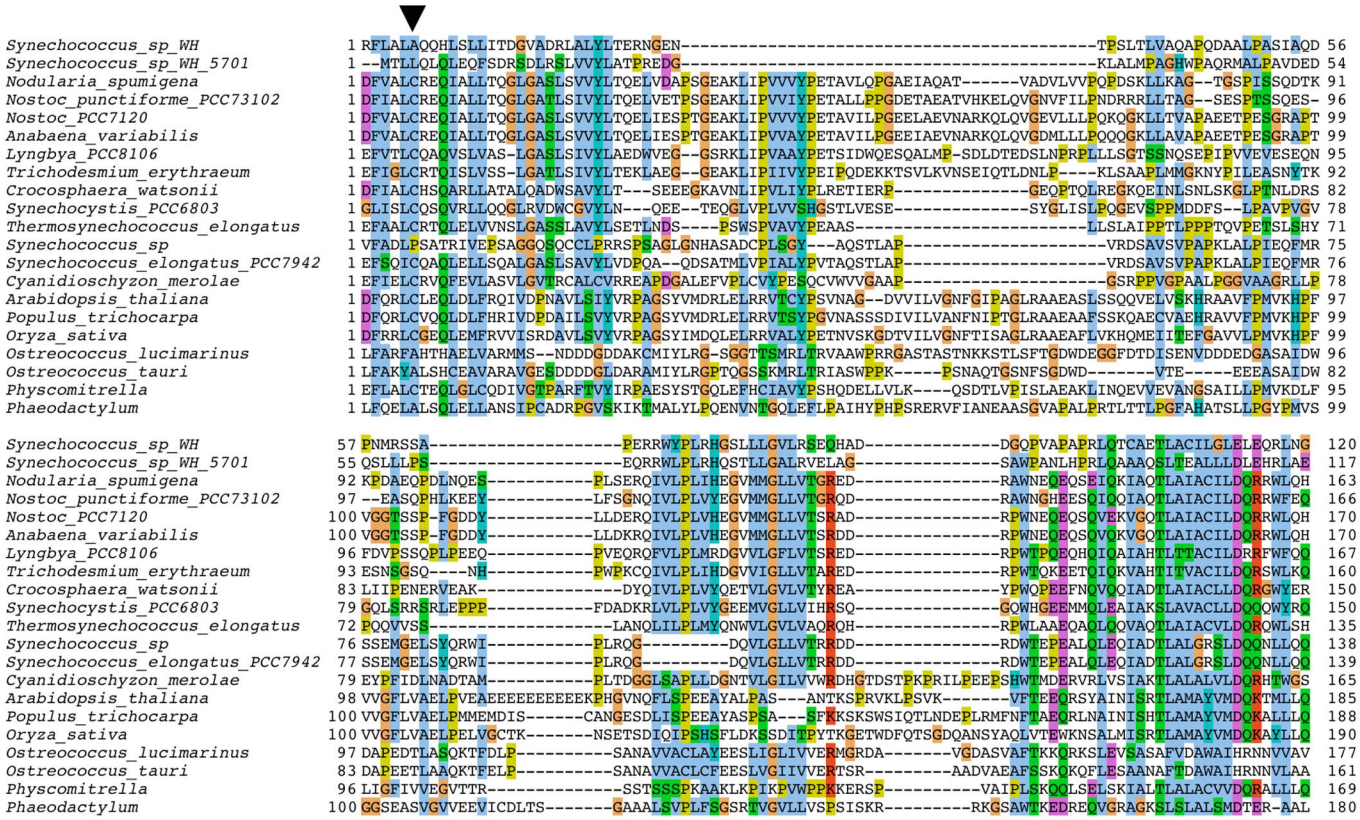

Fig. S2. Sequence alignment of the predicted GAF domain of the *Arabidopsis thaliana* CSK and its plant, algal, and cyanobacterial homologues (as delineated and named by the SMART database). The conserved cysteine residue in the cyanobacterial and plant CSKs is indicated by the arrowhead.

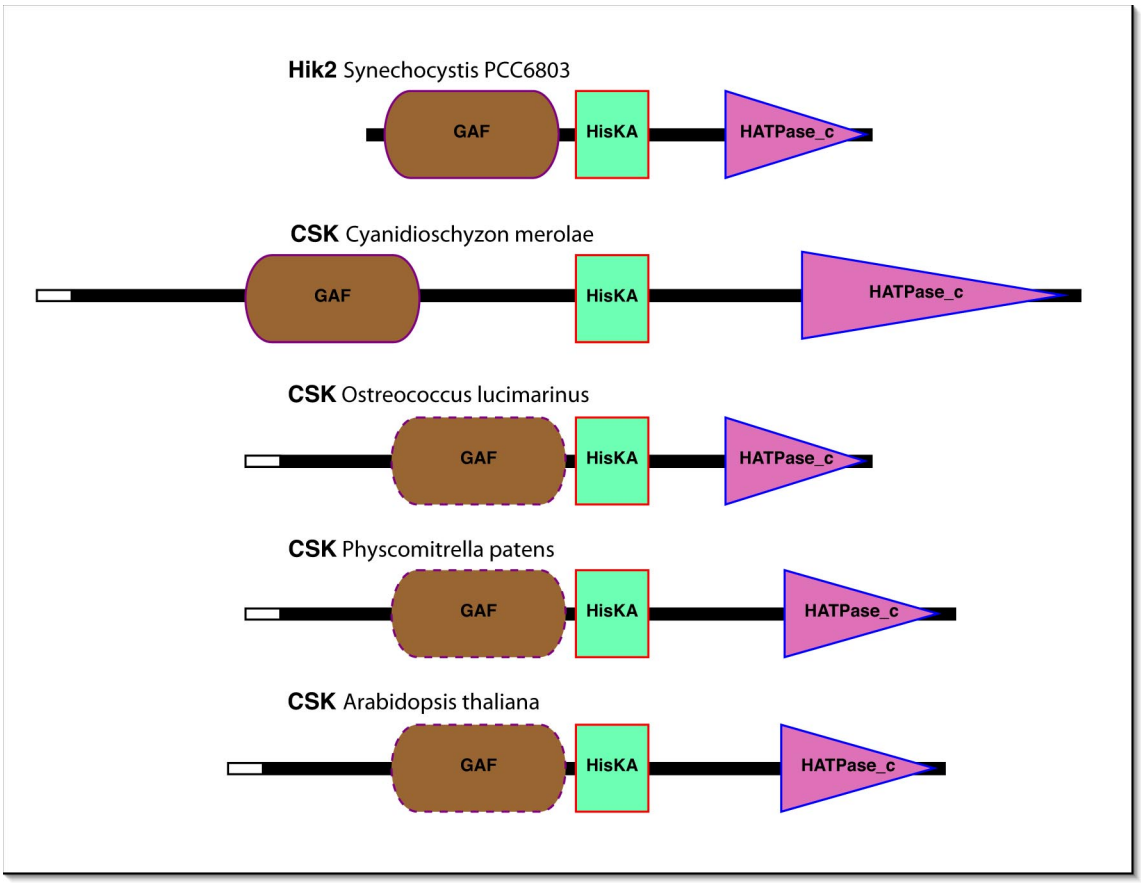

Fig. S3. Conserved functional domains in CSK and representative algal and cyanobacterial homologues (redrawn from SMART database predictions). For the *Ostreococcus* and the *Physcomitrella* CSKs, the predicted GAF domain is below the curated threshold score of SMART database. For the *Arabidopsis* CSK, SMART database does not predict a GAF domain, even though the *Arabidopsis* CSK sensor domain shows significant sequence homology to the GAF domain of cyanobacterial Hik2. Domain boundaries of GAF domains which fall below the curated threshold score or which are not predicted by SMART database are shown in broken lines. The chloroplast targeting signal is depicted as a white rectangle at the N terminus. Domain denotations: HATPase_c, ATP-binding domain; HisKA, site of histidine auto phosphorylation and dimerisation; GAF, sensor domain.