International Journal of Genetic Engineering 2026, 14(11): 400-407
DOI: 10.5923/j.ijge.20261411.11

In Silico Characterization of Phytochrome Proteins in Potato (*Solanum Tuberosum*)

Khusanbaeva Sh. R.*, Usmanov D. E., Abdukarimov Sh. S., Jurakulov D. S., Sharifjonov A. A., Abdugafforov A. T., Ziyodov O. T., Buriev Z. T.

Center of Genomics and Bioinformatics of the Academy of Sciences of the Republic of Uzbekistan, Kibray District, Tashkent, Uzbekistan

Abstract Phytochromes are photoreceptors that regulate plant growth and development in response to red and far-red light. Here, we characterized five phytochrome genes of potato (*Solanum tuberosum*), *StPHYA*, *StPHYB*, *StPHYB2*, *StPHYE*, and *StPHYF*, using bioinformatic approaches. Gene structure, chromosomal distribution, phylogenetic relationships, conserved domains, sequence motifs, and predicted functional associations were investigated. The five genes were located on five different chromosomes, and each contained four coding exons. In the phylogenetic analysis, each potato phytochrome clustered with its corresponding tomato protein. Pairwise Ka/Ks ratios were below 0.3 for all potato-tomato ortholog pairs and for the *PHYB* paralogs of both species, indicating purifying selection. Sequence alignment showed that the chromophore-binding cysteine is conserved the GAF domain of all five proteins. Domain analysis showed the same architecture in all five proteins (PAS_2, GAF, PHY, two PAS repeats, HisKA, and HATPase_c), and the ten conserved motifs identified occurred in the same order in all five. AlphaFold-predicted structures of the five proteins matched the fold of the *A.thaliana* PHYB photosensory module (RMSD 1.91-2.80 Å; TM-score 0.69-0.70), with the chromophore-binding cysteine at the conserved position. STRING analysis identified six predicted functional associations among the five proteins, and the Gene Ontology enrichment (biological process) of this network highlighted terms related to detection of visible light, protein-tetrapyrrole linkage, and red far-red light phototransduction. Overall, this study provides a comparative overview of the structural features, evolutionary relationships, and predicted functional associations of the potato phytochromes. These findings provide a basis for further research on phytochrome-mediated light signaling in potato.

Keywords *Solanum tuberosum*, Phytochromes, Photoreceptor, Conserved domains, Phylogenetic analysis, Light signaling

1. Introduction

Light is an essential environmental signal that regulates plant growth, development, and adaptation to changing environmental conditions. In addition to providing energy for photosynthesis, light conveys information about its wavelength, intensity, direction, and duration, enabling plants to adjust their developmental processes to the surrounding environment. These responses are mediated by specialized photoreceptors, among which phytochromes primarily perceive red and far-red light [1,2]. Phytochromes regulate diverse physiological and developmental processes throughout the plant life cycle, including seed germination, seedling photomorphogenesis, shade avoidance, and flowering time [1]. They function as photoreversible molecular switches that interconvert between the red-light-absorbing Pr form and the far-red-light-absorbing Pfr form, allowing plants to detect changes in light quality and initiate appropriate signaling responses [3]. Phytochromes are therefore important components of plant environmental sensing and developmental regulation.

Plant phytochromes are multidomain proteins composed of an N-terminal photosensory module and a C-terminal regulatory region. The photosensory module contains three conserved domains, PAS (Per/ARNT/Sim), GAF (cGMP phosphodiesterase/adenylyl cyclase/Fh1A), and PHY (phytochrome-specific), which together contribute to light perception and signal transduction. The GAF domain provides the binding site for the tetrapyrrole chromophore phytochromobilin, which is covalently attached to a conserved cysteine residue. Absorption of red and far-red light induces chromophore photoisomerization and subsequent conformational changes in the protein, initiating signaling [3,4]. The C-terminal region contains two additional PAS domains and a histidine kinase-related domain, which contribute to

* Corresponding author:
shahnoza.husanboyeva@outlook.com (Khusanbaeva Sh. R.)
Received: Nov. 5; Accepted: Nov. 21, 2026; Published: Nov. 23, 2026
Published online at http://journal.sapub.org/ijge

regulatory functions and signal transmission. Conservation of these structural features is therefore important for understanding phytochrome function and provides a basis for comparative analyses of phytochromes across plant species [5].

Previous studies have highlighted the importance of light-signaling pathways in potato (*Solanum tuberosum L.*) Wang et al. [6] investigated the role of phytochrome F (*StPHYF*) in potato flowering and tuberization using transcriptomic analyses of leaves and stolons. Genome-wide studies of other potato light-signaling gene families have also shown the value of combining genomic, structural, phylogenetic, and functional analyses: Han et al. [7] characterized the potato phytochrome-interacting factor (PIF) gene family and functionally investigated *StPIF3* in shade–avoidance responses, and Gao et al. [8] examined the potato cryptochrome (CRY) gene family, including its genomic organization, conserved features, phylogenetic relationships, and response to abiotic stresses. Comparative research on light-regulated transcripts in Solanaceae has further contributed to understanding the diversity of light responses among related species [9].

Zhang et al. [10] previously identified five phytochrome genes in the potato genome (*StPHYA*, *StPHYB*, *StPHYB2*, *StPHYC*, and *StPHYE*) and analyzed their expression under red light and its correlation with glycoalkaloid accumulation. However, the nomenclature of these genes differs between studies: the gene designated *StPHYC* in [10] is the same locus as *StPHYF* in [6], the tomato RefSeq annotation. Conservation of the chromophore-binding residue in relation to domain and motif positions, and the predicted functional associations among the five proteins remain to be examined in the potato phytochrome family.

We therefore characterized the potato phytochrome family in silico on the DM v6.1 annotation. We resolved the gene nomenclature by phylogenetic comparison with the tomato and *Arabidopsis* phytochromes, described chromosomal distribution and gene structure, assessed the conservation of domains, motifs, and the chromophore-binding cysteine, estimated selection pressure by pairwise Ka/Ks analysis, compared AlphaFold-predicted structures with the *A.thaliana PHYB* photosensory module, and predicted functional associations by protein interaction network and Gene Ontology enrichment. These analyses give a structural and evolutionary framework for later experimental work on phytochrome-mediated light signaling in potato.

2. Materials and Methods

2.1. Identification of Phytochrome Genes in Potato

Phytochrome genes were identified by BLASTP [11] in SpudDB [12] against *Solanum tuberosum* Group Phureja DM v6.1 annotation [13]. The Arabidopsis PHYB (AT2G18790.1) was used as the query, with an E-value threshold of 1e-5, word size 3, the BLOSUM62 matrix, no low-complexity filter, and 20 reported alignments. The search recovered five loci with full-length phytochrome sequences: Soltu.DM.10G011330 (*StPHYA*), Soltu.DM.01G019510 (*StPHYB*), Soltu.DM.05G023390 (*StPHYB2*), Soltu.DM.02G014300 (*StPHYE*) and Soltu.DM.07G015300 (*StPHYF*). Four further hits, annotated as kinases and showing only partial similarity to phytochromes, were excluded. Where a locus had several transcript isoforms, the isoform with the highest BLAST score was used, with ties resolved in favor of isoform .1: Soltu.DM.10G011330.1, Soltu.DM.01G019510.1, Soltu.DM.05G023390.1, Soltu.DM.02G014300.1 and Soltu.DM.07G015300.2.

2.2. Chromosomal Location and Gene Structure

Chromosomal positions and exon-intron structures were taken from the DM v6.1 high-confidence gene model annotation. Positions were plotted with Gene Structure View (Advanced) in TBtools – II v2.566 [14]. Exon and intron lengths were calculated from GFF3 coordinates.

2.3. Sequence Retrieval

Protein sequences of tomato and *Arabidopsis* phytochromes were retrieved from NCBI and TAIR [15], respectively. Coding sequences used for Ka/Ks analysis were retrieved from Spud DB (DM v6.1) for potato and from NCBI for tomato.

2.4. Multiple Sequence Alignment

The protein sequences of the five potato, five tomato, and five *Arabidopsis* phytochromes were aligned with MUSCLE [16] in MEGA12 [17] using default parameters. The chromophore-binding cysteine of each potato protein was identified in the GAF domain region of the alignment and confirmed against the MEME motif 5 sites.

2.5. Phylogenetic Analysis and Selection Pressure

A maximum-likelihood tree was inferred from the alignment in MEGA12 using the Jones-Taylor-Thornton model with empirical amino acid frequencies [18,19]. Among-site rate variation was modeled with a discrete Gamma distribution and a proportion of invariant sites. Branch support was assessed with 1000 bootstrap replicates [20].

Ka and Ks substitution rates were calculated for the five potato-tomato orthologous pairs and for the PHYB paralogous pairs for each species (*StPHYB/StPHYB2* and *SlPHYB1/SlPHYB2*) with the Simple Ka/Ks Calculator (NG) in TBtools-II v2.566 [14]. A Ka/Ks ratio below 1 was interpreted as purifying selection.

2.6. Conserved Domain and Motif Analysis

Conserved domains were identified with the NCBI Batch CD-search tool [21] against the Pfam database [22] and visualized with TBtools-II v2.566.

Conserved motifs were identified in the five full-length protein sequences with MEME v5.5.9 [23] in classic mode. The search was set to find up to 10 motifs of 6-50 amino

acids, allowing any number of repetitions per sequence, with no E-value cutoff.

2.7. Comparison of Predicted Protein Structures

AlphaFold models of the five potato phytochromes were retrieved from the AlphaFold Protein Structure Database [24,25] by UniProt accession (StPHYA, M1BAD9; StPHYB, M1C7W5; StPHYB2, M1CLI4; StPHYE, M0ZTS8; StPHYF, M1BGQ7). Residues 1-600 of each model were aligned to the photosensory module of *Arabidopsis thaliana* PHYB (PDB, 4OUR, Chain A) [26] with the RCSB Pairwise Structure Alignment tool [27].

2.8. Protein-Protein Association Network and Gene Ontology Enrichment

Functional associations among the five proteins were predicted with the STRING web server v12.5 [28] for *S.tuberosum*, with a minimum required interaction score of 0.400. Each potato protein was matched to its STRING node by UniProt alias and confirmed by identical protein length. Gene Ontology enrichment of the network was obtained from STRING enrichment function against the *S.tuberosum* whole-genome background, and terms with a false discovery rate <0.05 were considered enriched.

3. Results

3.1. Identification of Five Phytochrome Genes in Potato

BLASTP of the *Arabidopsis* PHYB protein (AT2G18790.1) against the DM v6.1 annotation recovered five loci that encode full-length phytochromes. The selected isoforms had BLAST scores of 3112-4788 with query coverage of 93,86-97,78% and identity of 53,02-79,29% to the query. Four additional hits, annotated as kinases and aligning only partially with the query, were excluded. The five genes are located on five different chromosomes (chromosomes 10, 1, 5, 2 and 7, respectively), and their proteins are 1120-1135 amino acids long (Table 1).

3.2. Chromosomal Distribution and Gene Structure

The five phytochrome genes are distributed on five different chromosomes: *StPHYA* on chromosome 10, *StPHYB* on chromosome 1, *StPHYB2* on chromosome 5, *StPHYE* on chromosome 2 and *StPHYF* on chromosome 7 (Figure 1). Each gene contains four coding exons separated by three introns (Figure 2). The coding sequence of the first exon is the longest of all five genes (2041-2071 bp, about 60% of the coding sequence), followed by exons of 802-817 bp, 291-294 bp and 208-250 bp.

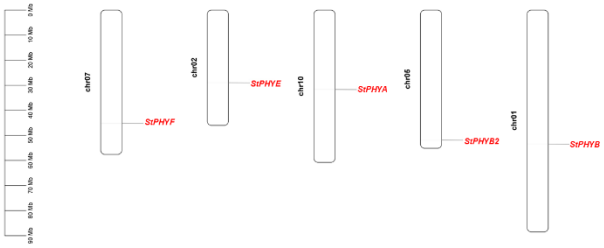

Figure 1. Chromosomal distribution of the five phytochrome genes

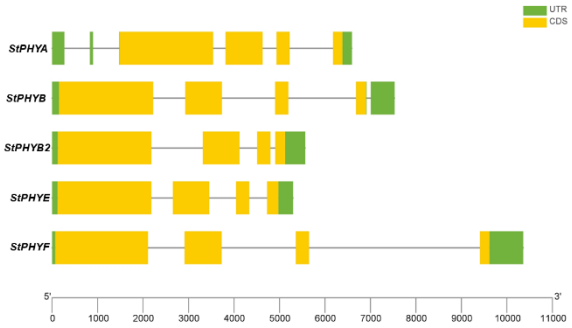

Figure 2. Gene structure of the five potato phytochrome genes

3.3. Sequence Alignment and Conservation of the Chromophore-Binding Cysteine

Alignment of the five potato phytochromes with tomato and Arabidopsis phytochromes (Figure 3) showed a conserved cysteine in the GAF domain of all five potato proteins, at positions 323 (StPHYA), 332 (StPHYB), 329 (StPHYB2), 338 (StPHYE) and 323 (StPHYF). This cysteine was present in all 15 aligned sequences and is followed by invariant histidine, in a conserved stretch (STLR-x-PH-x-CH). It corresponds to the cysteine to which the phytochromobilin chromophore is covalently attached in plant phytochromes. In the potato proteins, it lies within the Pfam GAF domain (for example, residues 218-401 in StPHYA and 227-406 in StPHYB) and at the same position in MEME motif 5 (residue 37 of the motif). These observations indicate that the chromophore-binding site is conserved in all five potato phytochromes.

Table 1. Phytochrome genes identified in potato by BLASTP of *Arabidopsis* PHYB (AT2G18790.1) against the DM v6.1 annotation

Gene	Locus	Chr	Protein (aa)	Blast score	Query coverage (%)	Identity (%)
StPHYA	Soltu.DM.10G011330.1	10	1123	3112	93.86	53.02
StPHYB	Soltu.DM.01G019510.1	1	1130	4788	94.28	79.29
StPHYB2	Soltu.DM.05G023390.1	5	1125	4403	97.78	70.79
StPHYE	Soltu.DM.02G014300.1	2	1135	3578	95.99	58.82
StPHYF	Soltu.DM.07G015300.2	7	1120	3175	93.86	53.48

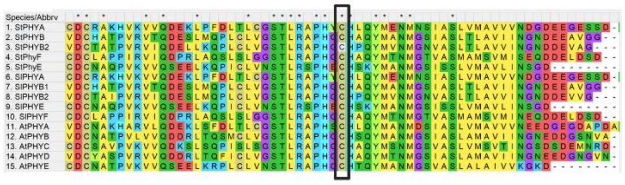

Figure 3. Alignment of the region surrounding the conserved chromophore-binding cysteine in the five potato phytochromes and in the tomato and Arabidopsis phytochromes

3.4. Phylogenetic Relationships and Selection Pressure

The maximum-likelihood tree of the five potato, five tomato and five *Arabidopsis* phytochromes (Figure 4) separated into three major groups, each with 100% bootstrap support: a PHYB/PHYD group containing StPHYB and StPHYB2, a PHYE group containing the PHYA and PHYC/PHYF lineages. Each potato protein is grouped with a single tomato counterpart: StPHYA with SlPHYA, StPHYB with SlPHYB1, StPHYB2 with SlPHYB2, StPHYE with SlPHYE, and StPHYF with SlPHYB2. All five potato-tomato pairs received 100% support.

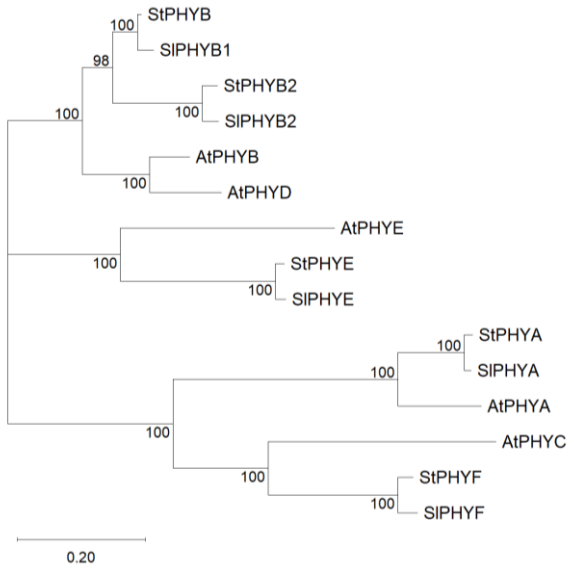

Figure 4. Maximum-likelihood phylogeny of the potato, tomato, and Arabidopsis phytochromes. The tree was inferred in MEGA12 from a MUSCLE alignment of 15 protein sequences using the JTT+F+G+I model

The Ka/Ks ratios of the five potato-tomato orthologous pairs ranged from 0.176 to 0.289 (Ka 0.014-0.024; Ks 0.064-0.096). The PHYB paralogs were considerably more divergent than the orthologous pairs in both species (potato *StPHYB/StPHYB2*: Ka 0.096, Ks 0.630, Ka/Ks 0.152; tomato *SlPHYB1/SlPHYB2*: Ka 0.106, Ks 0.675, Ka/Ks 0.156). All seven Ka/Ks ratios were below 0.3, indicating strong purifying selection on the phytochrome genes.

3.5. Conserved Domains and Motifs

Pfam domain analysis identified the same seven domains, in the same order, in all five potato phytochromes (Figure 5). The N-terminal photosensory region contains a PAS_2 domain, a GAF domain, and a PHY domain, and the C-terminal region contains two PAS domains followed by histidine kinase-related domains HisKA and HATPase_c. The start and end positions of each domain differed by at most 22 residues among the five proteins. The conserved cysteine described above lies within the GAF domain in all five proteins.

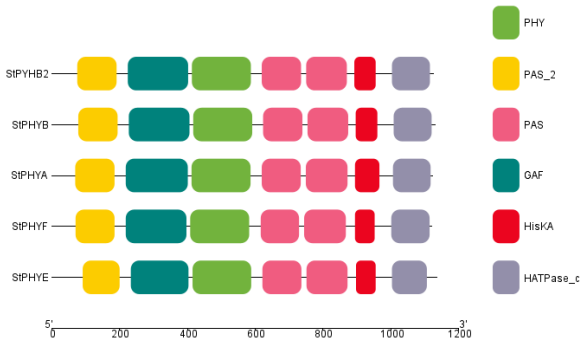

Figure 5. Domain architecture of the five potato phytochromes

MEME identified ten motifs of 49-50 amino acids that occurred in the same order in all five potato phytochromes (Figure 6). Each of the ten motifs overlapped a Pfam domain in every protein. Motifs 9 and 7 lay in the N-terminal PAS_2 domain (motif 7 extending about 17 residues beyond its end), motifs 3, 5 and 6 in the GAF domain, and motifs 4, 8 and 2 in the PHY domain. Motif 1 occurred twice per protein, in the first and the second C-terminal PAS domains, and motif 10 lay in the first PAS domain between the two copies of motif 1. Motif 5 includes the conserved chromophore-binding cysteine. Together, the ten motifs covered 548 residues in each protein (about 48-49% of the sequence), up to residues 794-804, and none fell within the C-terminal HisKA and HATPase_c domains.

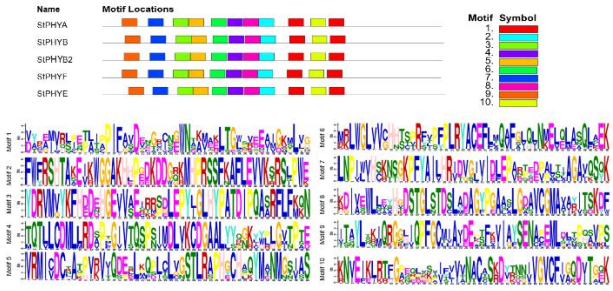

Figure 6. Conserved motifs in the five potato phytochromes

3.6. Comparison of Predicted Protein Structures

The AlphaFold models of five potato phytochromes had the same fold as the photosensory module of *A.thaliana* PHYB (PDB 4OUR). Over residues 1-600, the RMSD to the reference was 1.91-2.80 Å for 435-441 aligned residues, and the TM-scores were 0.69-.70 (Figure 8A). Sequence identity to the reference was highest for StPHYB (89%) and StPHYB2 (83%), and lower for StPHYE (72%), StPHYF (67%), and StPHYA (63%). The conserved chromophore-binding cysteine was at the same position as Cys 357 of *Arabidopsis* PHYB (PDB 4OUR; Figure 8B)

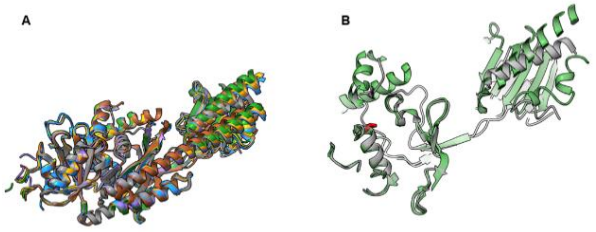

Figure 7. Predicted structures of the potato phytochrome photosensory modules. (A) Superposition of the AlphaFold models of StPHYA (blue), StPHYB (green), StPHYB2 (orange), StPHYE (purple) and StPHYF (brown). (B) The same superposition for StPHYB (green) and *Arabidopsis* PHYB (grey), showing conserved chromophore-binding cysteine

3.7. Functional Association Network and Gene Ontology Enrichment

The STRING network of the five potato phytochromes (minimum required interaction score 0.400) contained 6 edges and formed a single connected component (average node degree 2.4; Figure 7). StPHYB and StPHYB2 had the highest degree (three connections each), whereas StPHYA, StPHYE and StPHYF had two. Each of StPHYA, StPHYE and StPHYF was connected to both StPHYB and StPHYB2, with combined scores of 0.801 and 0.746 (StPHYA), 0.560 and 0.560 (StPHYE), and 0.772 and 0.707 (StPHYF), respectively; there were no edges among StPHYA, StPHYE and StPHYF or between StPHYB and StPHYB2. The number of edges exceeded that expected by chance (PPI enrichment $p=2.31\times10^{-11}$), although this comparison is of limited meaning for proteins that all belong to one family. All six edges are predicted associations. None was supported by interaction experiments in potato, curated databases or co-expression; they rest mainly on experimental evidence transferred from other species (STRING scores 0.558-0.712), with a smaller contribution from text mining (0.047-0.494). Four edges (those of StPHYA and StPHYF) had combined scores ≥ 0.700, whereas the two edges of StPHYE (0.560) rested almost entirely on transferred evidence. Gene Ontology (biological process) enrichment of the network identified detection of visible light (GO:0017006; FDR 3.1×10^{-16}) and red, far-red light phototransduction (GO:0009585; FDR 8.4×10^{-13}). These terms describe the annotation of the five proteins and are consistent with the known roles of phytochromes in light perception.

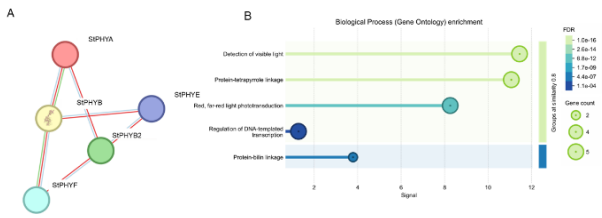

Figure 8. Functional association network and Gene Ontology enrichment of the five potato phytochromes. (A) STRING network (v12.5). Edges represent predicted functional associations mainly supported by experimental evidence transferred from other species and by text mining, not by interaction experiments in potato (B) Enrichment of Gene Ontology biological process terms

4. Discussion

We identified five phytochrome genes in the potato reference genome (DM v6.1): *StPHYA*, *StPHYB*, *StPHYB2*, *StPHYE* and *StPHYF*. They are located on chromosomes 10, 1, 5, 2 and 7, respectively. Seed plant phytochromes diversified into three lineages (PHYA, PHYB/E, and PHYC) before the origin of the crown group, and angiosperm PHYE arose as a duplicate within the PHYB lineage[29]. The five phytochrome genes therefore represent all three lineages. Because our tree is unrooted, the grouping of PHYA with PHYC/PHYF and the separate position of PHYE (Figure 4) do not conflict with this scheme. The tomato phytochrome family has the same five members [30], and each potato gene is paired with a tomato counterpart with 100% bootstrap support (Figure 4). This suggests that the family was already established before the two species diverged and has been maintained since. *StPHYF* corresponds to the gene named *StPHYC* in [10]. We use StPHYF because the sequence is the same (1200 aa, chromosome 7) and the gene groups with *SlPHYF* and *AtPHYC*. The tree separates StPHYB and StPHYB2 into two clades, each paired with a tomato PHYB sequence, and the node joining the two clades has 98% support. AtPHYB and AtPHYD form their own clade (100%). This matches the earlier finding that tomato has two B-type phytochrome genes, neither of which is an ortholog of *Arabidopsis* PHYB [31]. It suggests that the Solanaceae PHYB duplication arose independently of the PHYB/PHYD split in *Arabidopsis*. StPHYB (79% identity) and StPHYB2 (71%) were closest to AtPHYB, while StPHYE (59%), StPHYA (53%) and StPHYF (53%) were more divergent (Table 1).

Gene structure was conserved across the family. All five genes have four coding exons and three introns (Figure 2), as reported for tomato *PHYA* [32]. The coding sequence in each exon varied little among the genes (by up to about 40 bp), whereas intron lengths ranged from 103 bp to 3755 bp, with the longest introns in *StPHYF*. This pattern of constrained coding regions and variable introns suggests that selection acts mainly on protein-coding exons.

Protein-level conservation was higher still. All five proteins contain seven Pfam domains in the same order (Figure 5). The photosensory module (PAS_2, GAF and PHY) is at the N-terminus. It is followed by two PAS domains, then the HisKA and HATPase_c domains at the C-terminus, which are related to histidine kinases. Domain boundaries differed by at most 22 residues among the proteins. This architecture matches that of phytochromes in other plants [2,3], and [10] reported the same Pfam domains. The Cys residue that attaches the chromophore lies in the GAF domain of each protein (positions 323, 332, 329, 338 and 323; Figure 3). It was conserved in all 15 aligned sequences and is followed by an invariant His. This is consistent with the established role of the GAF Cys in covalent binding of the chromophore [3,5]. The MEME analysis agrees. Motif 5 contains the Cys, and motif 1 occurs twice in each protein, in the two PAS repeats (Figure 6).

The low Ka/Ks ratios agree with the conservation of the domains and of the chromophore-binding cysteine described above. The nearly identical divergence of the PHYB paralogs in potato and tomato (Ks 0.630 and 0.675) indicates that the duplication predates the split of the two species, consistent with the phylogeny (Figure 4). Both copies evolve under purifying selection, with no sign of relaxed constraint. Because the species are closely related, the numbers of substitutions are small, and a gene-wide ratio cannot detect selection at individual sites. We therefore do not compare strength among genes, and possible functional differences between the PHYB copies remain a hypothesis. Selection was assessed only by pairwise Ka/Ks; site-specific tests were not performed.

The STRING network connected all five proteins (Figure 8A). StPHYB and StPHYB2 were hubs, each linked to the other three proteins (degree 3). StPHYA, StPHYE, and StPHYF were each linked to the two PHYB proteins, and no edges connected A, E, and F. The network showed significantly enriched interactions (p=2.31x10-11). The enriched biological process terms were detection of visible light, protein-tetrapyrrole linkage, and red/far-red light phototransduction (Figure 8). In our network, none of the edges was supported by potato experiments, databases or co-expression, so the network should be read as a hypothesis that is consistent with the PHYB-centred heterodimer data from *Arabidopsis* [33] and tomato [34].

Beyond sequence data, specific potato phytochromes are already linked to light-dependent development. *StPHYF* is required for photoperiodic tuberization [35] and affects flowering and tuberization [6]. *PHYB-deficient* antisense potato plants can tuberize under long days, which wild-type plants cannot [36]. Red light changes phytochrome gene expression and glycoalkaloid accumulation [10]. Together with the duplication of *StPHYB*, the conserved domains we found suggest that the two *PHYB* genes could have partly distinct roles. This is a hypothesis and needs experimental testing.

This study has limits. All results are in silico and come from the DM v6.1 reference [13], a doubled monoploid. Cultivated potato is tetraploid, and its alleles may differ in sequence and structure. We did not analyze expression, and the STRING network is predicted.

5. Conclusions

We identified five phytochrome genes in the potato genome (*StPHYA*, *StPHYB*, *StPHYB2*, *StPHYE* and *StPHYF*), located on chromosomes 10, 1, 5, 2 and 7. Each potato gene was paired with a tomato counterpart with 100% bootstrap support, and the two *PHYB-type genes* formed separate clades from *AtPHYB* and *AtPHYD*. Pairwise Ka/Ks ratios were below 0.3 for all potato-tomato ortholog pairs and for the PHYB paralogs of both species, which indicates purifying selection. All five genes have four coding exons and three introns, and all five proteins contain the same seven Pfam domains. The chromophore-binding Cys in the GAF domain is conserved in all of them, and the MEME analysis placed it in motif 5. AlphaFold-predicted structures of the five proteins match the fold of the Arabidopsis PHYB photosensory module, with this cysteine at the conserved position. The STRING network placed StPHYB and StPHYB2 as hubs linked to the other three proteins. These associations are predicted, not experimentally shown. These results give a sequence-, evolution-, and - structure based reference for the potato phytochrome family. Expression analysis under different light and photoperiod regimes, together with functional tests of *StPHYB2* and *StPHYF*, is needed to establish the roles of individual genes in light signaling and tuberization.

REFERENCES

[1] Franklin KA, Quail PH. Phytochrome functions in Arabidopsis development. Journal of Experimental Botany. 2010; 61(1): 11–24. doi: 10.1093/jxb/erp304.

[2] Sharrock RA. The phytochrome red/far-red photoreceptor superfamily. Genome Biology. 2008; 9(8): 230. doi: 10.1186/gb-2008-9-8-230.

[3] Rockwell NC, Su Y-S, Lagarias JC. PHYTOCHROME STRUCTURE AND SIGNALING MECHANISMS. Annual Review of Plant Biology. 2006; 57(1): 837–858. doi: 10.1146/annurev.arplant.56.032604.144208.

[4] Bae G, Choi G. Decoding of Light Signals by Plant Phytochromes and Their Interacting Proteins. Annual Review of Plant Biology. 2008; 59(1): 281–311. doi: 10.1146/annurev.arplant.59.032607.092859.

[5] Burgie ES, Vierstra RD. Phytochromes: An Atomic Perspective on Photoactivation and Signaling. The Plant Cell. 2014; 26(12): 4568–4583. doi: 10.1105/tpc.114.131623.

[6] Wang E, Zhou T, Jing S, Dong L, Sun X, Fan Y, Shen Y, Liu T, Song B. Leaves and stolons transcriptomic analysis provide insight into the role of *phytochrome F* in potato flowering and tuberization. The Plant Journal. 2023; 113(2): 402–415. doi: 10.1111/tpj.16056.

[7] Han Y, Yang J, Zhang N, Gong Y, Liu M, Qiao R, Jiao X, Zhu F, Li X, Si H. Genome-Wide Identification of Phytochrome-Interacting Factor (PIF) Gene Family in Potatoes and Functional Characterization of StPIF3 in Regulating Shade-Avoidance Syndrome. Agronomy. 2024; 14(4): 873. doi: 10.3390/agronomy14040873.

[8] Gao Y, Yang X, Lv X, Li Y, Li K, Gao Y. Genome-Wide Identification of the CRY Gene Family in Solanum tuberosum and Response to Abiotic Stresses. Genes. 2025; 16(10): 1234. doi: 10.3390/genes16101234.

[9] Rutitzky M, Ghiglione HO, Cur áJA, Casal JJ, Yanovsky MJ. Comparative genomic analysis of light-regulated transcripts in the Solanaceae. BMC Genomics. 2009; 10(1): 60. doi: 10.1186/1471-2164-10-60.

[10] Zhang X, Jiang H, Liu W, Wang Y, Zeng F. Effect of Red Light on the Expression of the Phytochrome Gene Family and

the Accumulation of Glycoside Alkaloids in Potatoes. Foods. 2023; 12(23): 4194. doi: 10.3390/foods12234194.

[11] Altschul S. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research. 1997; 25(17): 3389–3402. doi: 10.1093/nar/25.17.3389.

[12] Hamilton JP, Brose J, Buell CR. SpudDB: a database for accessing potato genomic data Harris T, editor. GENETICS. 2025; 229(3): iyae205. doi: 10.1093/genetics/iyae205.

[13] Pham GM, Hamilton JP, Wood JC, Burke JT, Zhao H, Vaillancourt B, Ou S, Jiang J, Buell CR. Construction of a chromosome-scale long-read reference genome assembly for potato. GigaScience. 2020; 9(9): giaa100. doi: 10.1093/gigascience/giaa100.

[14] Feng J, Chen C, Wu Y, Fan L, Zhang Z, Chen M, Tang H, Yu G, Ren J, Luo J, et al. How to make big data accessible to plant biologists and beyond: Ten years of lessons from TBtools. Molecular Plant. 2026; 19(7): 1378–1391. doi: 10.1016/j.molp.2026.05.015.

[15] Reiser L, Bakker E, Subramaniam S, Chen X, Sawant S, Khosa K, Prithvi T, Berardini TZ. The Arabidopsis Information Resource in 2024 Harris T, editor. GENETICS. 2024; 227(1): iyae027. doi: 10.1093/genetics/iyae027.

[16] Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research. 2004; 32(5): 1792–1797. doi: 10.1093/nar/gkh340.

[17] Kumar S, Stecher G, Suleski M, Sanderford M, Sharma S, Tamura K. MEGA12: Molecular Evolutionary Genetic Analysis Version 12 for Adaptive and Green Computing Battistuzzi FU, editor. Molecular Biology and Evolution. 2024; 41(12): msae263. doi: 10.1093/molbev/msae263.

[18] Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. Bioinformatics. 1992; 8(3): 275–282. doi: 10.1093/bioinformatics/8.3.275.

[19] Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution. 1987 July [accessed 2026 Oct 29]. https://academic.oup.com/mbe/article/4/4/406/1029664/The-neighborjoining-method-a-new-method-for. doi: 10.1093/oxfordjournals.molbev.a040454.

[20] Felsenstein J. Confidence Limits on Phylogenies: An Approach Using the Bootstrap. Evolution. 1985; 39(4): 783. doi: 10.2307/2408678.

[21] Marchler-Bauer A, Bryant SH. CD-Search: protein domain annotations on the fly. Nucleic Acids Research. 2004; 32 (Web Server): W327–W331. doi: 10.1093/nar/gkh454.

[22] Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, Tosatto SCE, Paladin L, Raj S, Richardson LJ, et al. Pfam: The protein families database in 2021. Nucleic Acids Research. 2021; 49(D1): D412–D419. doi: 10.1093/nar/gkaa913.

[23] Bailey TL, Elkan C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. 1994. p. 28-36. (2:). PMID: 7584402.

[24] Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021; 596(7873): 583–589. doi: 10.1038/s41586-021-03819-2.

[25] Varadi M, Bertoni D, Magana P, Paramval U, Pidruchna I, Radhakrishnan M, Tsenkov M, Nair S, Mirdita M, Yeo J, et al. AlphaFold Protein Structure Database in 2024: providing structure coverage for over 214 million protein sequences. Nucleic Acids Research. 2024; 52(D1): D368–D375. doi: 10.1093/nar/gkad1011.

[26] Burgie ES, Bussell AN, Walker JM, Dubiel K, Vierstra RD. Crystal structure of the photosensing module from a red/far-red light-absorbing plant phytochrome. Proceedings of the National Academy of Sciences. 2014; 111(28): 10179–10184. doi: 10.1073/pnas.1403096111.

[27] Bittrich S, Segura J, Duarte JM, Burley SK, Rose Y. RCSB protein Data Bank: exploring protein 3D similarities via comprehensive structural alignments Elofsson A, editor. Bioinformatics. 2024; 40(6): btae370. doi: 10.1093/bioinformatics/btae370.

[28] Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, Gable AL, Fang T, Doncheva NT, Pyysalo S, et al. The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Research. 2023; 51(D1): D638–D646. doi: 10.1093/nar/gkac1000.

[29] Mathews S, Clements MD, Beilstein MA. A duplicate gene rooting of seed plants and the phylogenetic position of flowering plants. Philosophical Transactions of the Royal Society B: Biological Sciences. 2010; 365(1539): 383–395. doi: 10.1098/rstb.2009.0233.

[30] Alba R, Kelmenson PM, Cordonnier-Pratt M-M, Pratt LH. The Phytochrome Gene Family in Tomato and the Rapid Differential Evolution of this Family in Angiosperms. Molecular Biology and Evolution. 2000; 17(3): 362–373. doi: 10.1093/oxfordjournals.molbev.a026316.

[31] Pratt LeeH, Cordonnier-Pratt M-M, Hauser B, Caboche M. Tomato contains two differentially expressed genes encoding B-type phytochromes, neither of which can be considered an ortholog of Arabidopsis phytochrome B. Planta. 1995 [accessed 2026 Oct 19]; 197(1). http://link.springer.com/10.1007/BF00239958. doi: 10.1007/BF00239958.

[32] Lazarova GI, Kerckhoffs LHJ, Brandstädter J, Matsui M, Kendrick RE, Cordonnier-Pratt M, Pratt LH. Molecular analysis of *PHYA* in wild-type and phytochrome A-deficient mutants of tomato. The Plant Journal. 1998; 14(6): 653–662. doi: 10.1046/j.1365-313x.1998.00164.x.

[33] Sharrock RA, Clack T. Heterodimerization of type II phytochromes in *Arabidopsis*. Proceedings of the National Academy of Sciences. 2004; 101(31): 11500–11505. doi: 10.1073/pnas.0404286101.

[34] Balderrama D, Barnwell S, Carlson KD, Salido E, Guevara R, Nguyen C, Madlung A. Phytochrome F mediates red light responsiveness additively with phytochromes B1 and B2 in tomato. Plant Physiology. 2023; 191(4): 2353–2366. doi: 10.1093/plphys/kiad028.

[35] Zhou T, Song B, Liu T, Shen Y, Dong L, Jing S, Xie C, Liu J. Phytochrome F plays critical roles in potato photoperiodic tuberization. The Plant Journal. 2019; 98(1): 42–54. doi: 10.1111/tpj.14198.

[36] Jackson SD, James PE, Carrera E, Prat S, Thomas B. Regulation of Transcript Levels of a Potato Gibberellin 20-Oxidase Gene by Light and Phytochrome B. Plant Physiology. 2000; 124(1): 423–430. doi: 10.1104/pp.124.1.423.

Copyright © 2026 The Author(s). Published by Scientific & Academic Publishing
This work is licensed under the Creative Commons Attribution International License (CC BY). http://creativecommons.org/licenses/by/4.0/