

International Journal of Advanced Biochemistry Research



ISSN Print: 2617-4693
ISSN Online: 2617-4707
NAAS Rating (2026): 5.29
IJABR 2026; SP-10(4): 216-221
www.biochemjournal.com
Received: 08-02-2026
Accepted: 10-03-2026

Soumyanetra Saha
Department of Genetics and
Plant Breeding, N. M. College
of Agriculture, Navsari
Agricultural University,
Navsari, Gujarat, India

VB Rana
Megaseed, Pulses and Castor
Research Station, Navsari
Agricultural University,
Navsari, Gujarat, India

Kaushal Modha
Department of Genetics and
Plant Breeding, N. M. College
of Agriculture, Navsari
Agricultural University,
Navsari, Gujarat, India

VB Patel
Forest Biotechnology
laboratory, College of Forestry,
Navsari Agricultural
University, Navsari,
Gujarat, India

RK Patel
Department of Genetics and
Plant Breeding, N. M. College
of Agriculture, Navsari
Agricultural University,
Navsari, Gujarat, India

DA Chauhan
Megaseed, Pulses and Castor
Research Station, Navsari
Agricultural University,
Navsari, Gujarat, India

Naresh Chaudhary
Department of Genetics and
Plant Breeding, N. M. College
of Agriculture, Navsari
Agricultural University,
Navsari, Gujarat, India

Corresponding Author:
Soumyanetra Saha
Department of Genetics and
Plant Breeding, N. M. College
of Agriculture, Navsari
Agricultural University,
Navsari, Gujarat, India

Deciphering the evolutionary dynamics of CONSTANS (CO) gene homologs in legumes

Soumyanetra Saha, VB Rana, Kaushal Modha, VB Patel, RK Patel, DA Chauhan and Naresh Chaudhary

DOI: <https://www.doi.org/10.33545/26174693.2026.v10.i4Sc.8088>

Abstract

Legume crops belonging to the family Fabaceae play a vital role in global agriculture due to their nutritional value, ecological benefits, and ability to fix atmospheric nitrogen through symbiotic associations with *Rhizobium* species. Among the genes regulating plant development, the *CONSTANS* (*CO*) and *CONSTANS-LIKE* (*COL*) gene family is known to play a crucial role in controlling photoperiod-dependent flowering and circadian rhythm pathways. Understanding the evolutionary relationships of these genes can provide valuable insights into their functional conservation and diversification across plant species. The present study aimed to investigate the phylogenetic relationships of *CONSTANS* (*CO*) and *CONSTANS-LIKE* (*COL*) genes in legumes. Three Indian bean (*Lablab purpureus*) genotypes (GNIB21, GNIB22, and GP189) were analyzed by amplifying the *LprCO* gene using primers designed from *Vigna unguiculata* as a reference. The amplified fragments were sequenced using the Sanger sequencing method and processed to obtain consensus sequences. Homologous *CO/COL* gene sequences from various plant species were retrieved from the National Center for Biotechnology Information (NCBI) database. Multiple sequence alignment was performed using the CLUSTAL W algorithm in MEGA11 software, and the best-fit nucleotide substitution model (T92+G) was selected based on Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) values. Phylogenetic trees were constructed using bootstrap analysis with 1000 replicates to evaluate statistical support. The phylogenetic analysis based on both nucleotide and protein sequences revealed the presence of several distinct clades, indicating evolutionary divergence within the *CO/COL* gene family. The *LprCO* sequences clustered closely with known *CO* homologs from related legume species, suggesting a conserved evolutionary origin. The presence of multiple *COL*-related genes within specific clades indicates possible gene duplication events contributing to functional diversification. Overall, the results highlight the evolutionary conservation and diversification of *CO/COL* genes across plant species and provide insights into their potential roles in flowering time regulation and photoperiodic responses in legumes.

Keywords: CONSTANS, CONSTANS like, phylogeny, legume, photoperiod dependent flowering

Introduction

A significant category of plants grown globally for food, fodder, oil, and improving soil fertility are legume crops, which are members of the Fabaceae (Leguminosae) family (Smykal *et al.*, 2015) [9]. Through symbiotic relationships with *Rhizobium* species, they exhibit a distinctive ability to fix atmospheric nitrogen. Since root nodules contain microbes that enhance soil fertility and lessen reliance on artificial nitrogen fertilizers, they are crucial to sustainable agriculture (Graham and Vance, 2003) [5]. Legumes are widely produced in India's various agroclimatic zones and are essential to the country's food and nutritional security. Climate-resilient crops, including legumes, have drawn closer consideration due to stagnant production trends of common staple crops and rising concerns about environmental sustainability (Cowling *et al.*, 2016) [3]. Legumes are desirable targets for genomic and proteomic studies targeting abiotic stress tolerance due to their notable genetic diversity and adaptive features (Varshney *et al.*, 2013) [12]. A vital element of evolutionary biology, phylogenetic analysis assists in discovering genotypes that are resilient to stress and evolutionarily conserved in addition to helping illustrate the genetic links across species (Wendel and Doyle, 1998) [13]. This developmental knowledge supports molecular breeding operations aimed at crop improvement when integrated with high-throughput genomics

(Varshney *et al.*, 2005) [11]. Darwin originally suggested phylogenetic trees, which are now codified using molecular data and are powerful instruments for tracing evolutionary paths, detecting gene flow, and anticipating functional conservation (Gregory, 2008; Weyenberg and Yoshida, 2016) [6, 14].

The CONSTANS-like (COL) protein family, that is well-known for its ability to sense photoperiod, is vital for controlling plant flowering, which is a vital stage in the life cycle of plants (Wickland and Hanzawa, 2015; Liu *et al.*, 2021) [15, 8]. Homologous COL genes have been discovered in a number of plant species over the last twenty years, which has substantially improved our comprehension of their function in the flowering procedure and reaction to abiotic stress (Wong *et al.*, 2015) [16]. Cloning a number of flowering loci in pea (*Pisum sativum* L.) has shown that the *Arabidopsis* circadian clock genes GI, ELF4, and ELF3 have conserved functions in the control of FT genes and the control of photoperiod-specific flowering (Abbas *et al.*, 2023). GI in soybeans has also been shown to play a similar impact. Nevertheless, there has been no direct testing of the endogenous activity of CO-like (COL) genes in legumes, and it is still unclear whether they could be involved in photoperiod measurement (Flores *et al.*, 2018). This study focused on phylogeny analysis of *CONSTANS* (CO) and its homologs in legumes.

Materials and Methods

1. Data extraction

The first and most crucial phase in any bioinformatics project is gathering precise data. In order to collect entire sequence data instead of partial sequence data, choosing the right sequences is a crucial step.

Three Indian bean genotype has been used to amplify *LprCO* by using primers designed from *Vigna unguiculata* as a reference gene. Amplified fragments were sequenced via Sanger sequencing method. Sequencing data have been processed and final consensus sequence was obtained. The National Center for Biotechnology Information (NCBI) provided comparable *CONSTANS* gene sequence data for other legume species, which were obtained in FASTA format (Anonymous, 2025) [2].

2. Statistical analysis

Validated computational techniques were used for statistical analysis, guaranteeing consistent results and precise biological data interpretation.

Molecular Evolutionary Genetics Analysis version 11 (MEGA11) software was used for phylogenetic analysis in the current study (Hall, 2013) [7]. The following methodical process was adhered to:

- a) **Data collection:** Creating a high-quality dataset using sequences of proteins and DNA obtained from GenBank (NCBI) as a main source in FASTA format was the initial stage in building the phylogenetic tree.
- b) **Homology search:** To find homologous sequences in the NCBI GenBank database, the retrieved sequences were utilized as queries in BLASTn (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). Significantly comparable sequences derived from the Fabaceae family were chosen for additional examination.
- c) **Sequence curation:** To guarantee quality and completeness, a selection of sequences was examined.

Full-length or nearly full-length sequences were the only ones kept.

- d) **Multiple sequence alignment (MSA):** *Arabidopsis thaliana* (L.) Heynh is one of the curated sequences used in multiple sequence alignment (MSA) were aligned using MEGA11's CLUSTAL W method as an outgroup.
- e) **Model selection:** Using MEGA11, go to Models and choose Find Best DNA/Protein Models to determine the best-fit nucleotide or protein substitution model (Tamura *et al.*, 2011). Run the model test after choosing the alignment file. The model that had the lowest values for both the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) was chosen.
- f) **Phylogenetic tree construction:** The chosen substitution model was used to build the phylogenetic tree. Select Phylogeny and, depending on the approach chosen, choose Construct/Test Neighbor-Joining (NJ), Maximum Likelihood (ML), or UPGMA Tree. Open the file containing the aligned sequence. For statistical support, set the bootstrap replications to 1000. Create the tree and publish it for documentation in picture format (Felsenstein, 1985). Distance matrix approaches (also called clustering or algorithmic methods) and character-based methods (sometimes called discrete data methods) were the two main categories into which phylogenetic tree construction techniques were divided. The phylogenetic tree was constructed using the most suitable approach of these.

Results

The processed *Lablab purpureus Constans* (*LprCO*) nucleotide sequence was used as a query in BLASTn (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) to identify homologous sequences in the NCBI nucleotide database. We recovered homologous sequences from species belonging to the Fabaceae family. The CLUSTAL W technique in MEGA 11 (Molecular Evolutionary Genetics Analysis) was used to perform multiple sequence alignment of *Arabidopsis thaliana* and numerous Fabaceae PhyA species. The optimal substitution model was determined using the "Find Best DNA/Protein Models" function in MEGA 11. Sequences with all gaps and missing data removed were used to examine nucleotide substitutions. By striking a balance between model complexity and goodness-of-fit, the findings of nucleotide substitution model selection for the CO gene, assessed based on Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) values, aid in the identification of the best evolutionary model.

A better-fitting model is indicated by lower BIC and AIC scores. The best-fitting model for the CO sequencing data was T92+G (Tamura 3-parameter model with Gamma distribution), which had the lowest BIC (66116.14) and AIC (65783.66) values among the investigated models. This model is appropriate for describing the evolutionary processes of the gene because it takes into consideration uneven nucleotide frequencies and variation in substitution rates among sites. The form parameter of the Gamma distribution, which is used to model rate heterogeneity among sites, is represented by gamma (Γ) values. The majority of models produced a Gamma value between 0.49 and 0.51, indicating a moderate rate variation across PhyA sequence locations.

For every model, nucleotide frequencies (G, A, C & T) were also computed. With adenine (A) and thymine (T) frequencies typically at 0.27-0.30 and cytosine (C) and guanine (G) frequencies around 0.21, the observed base composition was comparatively balanced in the majority of models. Notably, the K2+G model has a more balanced nucleotide distribution (0.25 for each base), but its low performance in terms of BIC and AIC suggests that it fits the data less accurately. Overall, the T92+G model's ideal ratio of biological significance to statistical support led to its selection for downstream phylogenetic analysis.

The rates at which one nucleotide transforms into another nucleotide are displayed in the substitution matrix. A particular evolutionary model used in genetics to explain the rate of nucleotide substitutions is the Tamura 3-parameter model. It considers variations in the rates of transversions (purine to pyrimidine or vice versa) and transitions pyrimidine to pyrimidine (Cytosine and Thymine) or (purine to purine (Adenine and Guanine)). The figures show the likelihood that one nucleotide will change into another in a specific amount of time. A greater degree of substitution is indicated by larger numbers. The phylogeny tree illustrates the evolutionary relationships among the proteins and reveals their classification into several distinct clades. Bootstrap values (based on 1000 replicates) are indicated at each node and represent the statistical support for the branching pattern. Proteins and nucleic acid from related species cluster together, suggesting conserved evolutionary relationships and possible functional similarity within each subgroup.

To investigate the evolutionary relationships among *CONSTANS* (CO) and *CONSTANS LIKE* (COL) genes, a phylogenetic tree was constructed using nucleotide sequences from multiple plant species (Figure 1). The analysis revealed that the CO/COL genes were grouped into three major clades, indicating evolutionary divergence within this gene family.

The first clade consisted of *LprCO* GNIB21, *LprCO* GNIB22, *LprCO* GP189, *GmCO*, *LaCO*, and *VrCO*. Within this group, *LprCO* GNIB21 and *LprCO* GNIB22 clustered closely together and further grouped with *LprCO* GP189 with strong bootstrap support. These sequences were also closely associated with *GmCO*, *LaCO*, and *VrCO*, suggesting a high level of nucleotide sequence similarity among these genes. This clustering indicates that these genes may represent conserved CO homologs sharing a common evolutionary origin.

The second clade included *GmCOL2a*, *VuCOL-2L*, *VuCOL2*, and *VuCOL4*, forming a distinct subgroup with relatively high bootstrap support values. Within this group, *VuCOL2* and *VuCOL4* clustered together, indicating close nucleotide sequence similarity and possible functional relatedness. The presence of several COL genes within this clade suggests that gene duplication and subsequent evolutionary divergence may have contributed to the expansion of this subgroup.

The third clade comprised *MiCO*, *CaCOL-5*, *OsHdl*, *AiCO*, *AtCOL1*, *GmCOL9*, and *MtCOL-2*. This clade showed several internal groupings with varying bootstrap support values. For example, *MiCO* and *CaCOL-5* clustered together with strong statistical support, while *AiCO* and *AtCOL1* formed another well-supported subgroup.

Additionally, *GmCOL9* and *MtCOL-2* were grouped together, indicating evolutionary conservation of these COL genes across different plant species.

Overall, the phylogenetic analysis based on nucleotide sequences demonstrated that CO/COL genes are organized into three distinct evolutionary groups, reflecting sequence divergence and potential functional specialization among members of this gene family.

To investigate the evolutionary relationship among *CONSTANS* (CO) and *CONSTANS-LIKE* (COL) proteins, a phylogenetic tree was constructed using protein sequences from various plant species (Figure 2). The analysis revealed that CO/COL proteins are grouped into several well-defined clades, indicating evolutionary divergence within this gene family.

The phylogenetic tree showed that protein CO GNIB22, protein CO GP189, protein CO GNIB21, and PvCOL2-like clustered together with strong bootstrap support (99%), suggesting that these proteins share a close evolutionary relationship and may represent orthologous members of the CO/COL family. The close association among these proteins indicates a high degree of sequence conservation.

Another cluster consisted of *VuCO*, *VaCOL2-like*, *VrCOL2*, and *VmCO*, forming a separate subgroup supported by bootstrap values ranging from 44 to 85. Although some internal nodes displayed moderate support, the grouping of these proteins suggests that they belong to a conserved COL2-related lineage within leguminous species.

A third clade included *CcCOL-2*, *GmCOL2*, *GsCOL2*, and *GmCOL-2a*, which showed strong bootstrap support (82-100). The presence of two homologous genes (*GmCOL2* and *GmCOL-2a*) within the same species suggests possible gene duplication events during evolution. Such duplication events may have contributed to functional diversification within the CO/COL gene family.

The fourth clade contained *ApCOL2l*, *PpCOL2*, *MdCOL-2*, *PaCOL2*, *GhCOL-1*, and *GbCOL-1*, which also formed a well-supported cluster with bootstrap values ranging from 75 to 100. The grouping of these proteins indicates that the CO/COL gene family is conserved across diverse plant species and has undergone evolutionary divergence while maintaining structural similarities.

Overall, the phylogenetic analysis revealed that CO/COL proteins are distributed into multiple subgroups, reflecting their evolutionary diversification and potential functional specialization among plant species.

Discussion

The *CONSTANS* (CO) and *CONSTANS-LIKE* (COL) gene family plays an important role in regulating photoperiodic flowering and circadian rhythm pathways in plants. The phylogenetic analysis based on nucleotide and protein sequences provides insights into the evolutionary relationships and diversification of these genes among different plant species.

The grouping of *LprCO* GNIB21, *LprCO* GNIB22, and *LprCO* GP189 with known CO genes such as *GmCO*, *LaCO*, and *VrCO* suggests that these genes share a close evolutionary relationship. The high nucleotide sequence similarity within this group indicates that they may have originated from a common ancestral gene and may perform similar regulatory functions in flowering pathways.

The second group, which includes GmCOL2a, VuCOL-2L, VuCOL2, and VuCOL4, likely represents a lineage of COL2-related genes within the CO/COL gene family. The presence of multiple homologous genes within this group suggests that gene duplication events may have contributed to the expansion and diversification of COL genes in plant genomes.

The third group contains several CO/COL homologs from different plant species, including MiCO, CaCOL-5, OsHd1, AiCO, AtCOL1, GmCOL9, and MtCOL-2. The presence of OsHd1, a well-characterized CO homolog involved in photoperiodic flowering regulation in rice, highlights the evolutionary conservation of CO-related genes across different plant taxa. This conservation suggests that the fundamental regulatory mechanisms controlled by CO/COL genes have been maintained throughout plant evolution.

The clustering of protein CO GNIB22, protein CO GP189, and protein CO GNIB21 with PvCOL2-like suggests that these proteins share a common evolutionary origin. The strong bootstrap support observed within this clade indicates a high level of sequence conservation, which is characteristic of genes involved in essential regulatory processes such as flowering time control. The conservation of these proteins across species implies that they may retain similar biological functions.

The second clade, which includes VuCO, VaCOL2-like, VrCOL2, and VmCO, likely represents a conserved lineage of COL2 proteins within related plant species. The moderate bootstrap support observed for some nodes may reflect evolutionary divergence resulting from species-specific adaptations. Such divergence may influence gene expression patterns or regulatory functions in different plant systems.

The clustering of CcCOL-2, GmCOL2, GsCOL2, and GmCOL-2a suggests that these proteins belong to a conserved COL2 subgroup. The occurrence of multiple homologs within the same species, such as GmCOL2 and GmCOL-2a, indicates that gene duplication events may have contributed to the expansion of the CO/COL gene family. Gene duplication is an important evolutionary mechanism that provides opportunities for functional diversification while maintaining essential regulatory roles.

In addition, the grouping of ApCOL2l, PpCOL2, MdCOL-2, PaCOL2, GhCOL-1, and GbCOL-1 into a distinct clade further highlights the evolutionary conservation of COL proteins across diverse plant taxa. The strong bootstrap values supporting this cluster indicate a high degree of sequence similarity, suggesting that these genes may perform conserved functions in flowering regulation and photoperiodic signaling pathways.

Taken together, the phylogenetic analysis indicates that the CO/COL gene family has undergone evolutionary diversification while maintaining conserved structural and functional characteristics. These results provide valuable insights into the evolutionary history of CO/COL proteins and may contribute to understanding their roles in plant developmental regulation and flowering time control. Two B-box domains within an N-terminal Zn finger region and a conserved C-terminal (CCT) domain, which is also present in the circadian clock-related pseudo-response regulator gene TOC1 and related PRR, are characteristics of this set of genes, which includes Arabidopsis CO.

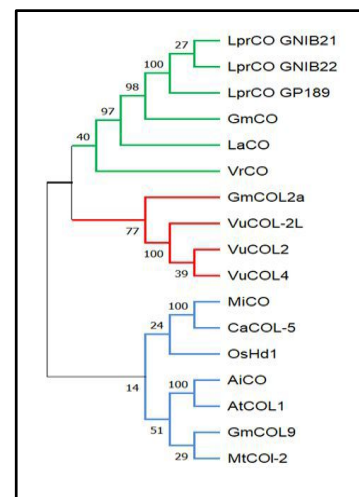


Fig 1: Phylogeny analysis of nucleotide sequence belongs to different CO homologs

Lpr: *Lablab purpureus*

Gm: *Glycine max*

Vr: *Vigna radiata*

Gm: *Glycine max CO-like2a*

Vu: *Vigna unguiculata*

Mi: *Mangifera indica*

Ca: *Cicer arietinum*

Os: *Oryza sativa*

Ai: *Arabidopsis iryta*

At: *Arabidopsis thaliana*

Mt: *Medicago truncatula*

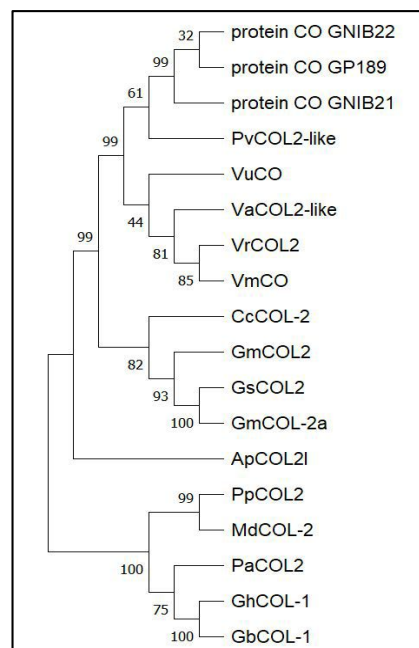


Fig 2: Phylogeny analysis of protein sequence belongs to different CO homologs

Ap: *Abrus precatorious*

Tp: *Trifolium pratense*

Md: *Malus domestica*

Gb: *Gossipium barbadense*

Gh: *Gossipium hirsutum*

Pp: *Populus alba*

Vm: *Vigna mungo*

Vu: *Vigna unguiculata*

Vr: *Vigna radiata*

Gm: *Glycine max*

Gs: *Glycine soja*

Cc: *Cajanus cajan*

Lj: *Lotus japonicus*

Conclusion

The present study provides a comprehensive phylogenetic analysis of the *CONSTANS* (*CO*) and *CONSTANS-LIKE* (*COL*) gene family, highlighting their evolutionary relationships and diversification across different plant species. The clustering of *LprCO* GNIB21, *LprCO* GNIB22, and *LprCO* GP189 with previously characterized *CO* genes such as *GmCO*, *LaCO*, and *VrCO* suggests that these genes share a close evolutionary origin and may perform similar regulatory roles in photoperiod-dependent flowering pathways. The identification of several distinct clades containing *COL2*-related genes further indicates that gene duplication and subsequent divergence have played a significant role in the expansion and functional diversification of the *CO/COL* gene family in plant genomes.

The presence of well-characterized homologs, including OsHd1 and other *COL* proteins, within the phylogenetic groups highlights the evolutionary conservation of this gene family across diverse plant taxa. Such conservation suggests that the fundamental molecular mechanisms controlling flowering time and photoperiodic responses have been maintained during plant evolution. Moreover, strong bootstrap support observed within several clades reflects a high degree of sequence similarity and structural conservation among proteins involved in essential regulatory processes such as circadian rhythm regulation and flowering time control.

In addition to phylogenetic relationships, the conserved structural features of *CO/COL* proteins, including the N-terminal B-box zinc finger domains and the C-terminal CCT domain, further support their functional conservation. These domains are known to play critical roles in transcriptional regulation, protein-protein interactions, and photoperiodic signaling pathways that regulate flowering in plants.

Overall, the findings of this study provide valuable insights into the evolutionary history, structural conservation, and potential functional roles of *CO/COL* genes. This research contributes to a deeper understanding of the molecular mechanisms underlying flowering time regulation and photoperiodic responses in plants. Furthermore, the identified genes may serve as potential candidates for future functional characterization and genetic studies aimed at improving flowering regulation, adaptation, and crop productivity in different plant species.

Author's Contributions

Soumyanetra Saha, Kaushal Modha, Vipul Patel and V. B. Rana designed the study. Kaushal Modha and Ritesh Patel helped in funding. Kaushal Modha framed the laboratory experimentation. D. A. Chauhan and Kaushal Modha provided the seed material and agronomic practices. Soumyanetra Saha, V. B. Rana and Naresh Chaudhary collected the data. Soumyanetra Saha and V. B. Rana organized the data, performed the analysis and prepared the manuscript.

Acknowledgment

The authors acknowledge the support provided by the Department of Genetics and Plant Breeding, High tech Omics Experimentation Laboratory (HOME), Navsari Agricultural University, Navsari. We are also grateful to the faculty members of the Department of Genetics and Plant

Breeding, NAU, Navsari for their valuable guidance and encouragement.

Disclaimer (AI)

The authors declare that generative AI technologies, including OpenAI ChatGPT (GPT-5 model) were used during the writing and editing of this manuscript to assist in language refinement and formatting. All intellectual and analytical contributions are the authors' own, and the AI was used solely as a tool for improving readability.

Competing Interests

The authors declare that no competing interests exist.

Reference

1. Abbas A, Ali A, Hussain A, Ali A, Alrefaei AF, Naqvi SAH, Baloch FS. Assessment of genetic variability and evolutionary relationships of *Rhizoctonia solani* inherent in legume crops. *Plants*. 2023;12(13):2515.
2. Anonymous. NCBI (National Center for Biotechnology Information) [Internet]. 2025 [cited 2025 Jan 18]. Available from: <https://www.ncbi.nlm.nih.gov/>
3. Cowling W, Bramley H, Mori TA, Hodgson JM, Cooper JW. Neglecting legumes has compromised human health and sustainable food production. *Nat Plants*. 2016;2(8):10.
4. Flores AC, Via VD, Savy V, Villagra UM, Zanetti ME, Blanco F. Comparative phylogenetic and expression analysis of small GTPases families in legume and non-legume plants. *Plant Signal Behav*. 2018;13(2):1432956.
5. Graham PH, Vance CP. Legumes: importance and constraints to greater use. *Plant Physiol*. 2003;131(3):872-877.
6. Gregory TR. Understanding evolutionary trees. *Evol Educ Outreach*. 2008;1:121-137.
7. Hall BG. Building phylogenetic trees from molecular data with MEGA. *Mol Biol Evol*. 2013;30(5):1229-1235.
8. Liu C, Zhang Q, Zhu H, Cai C, Li S. Characterization of mungbean *CONSTANS-LIKE* genes and functional analysis of *CONSTANS-LIKE 2* in the regulation of flowering time in *Arabidopsis*. *Front Plant Sci*. 2021;12:608603.
9. Smýkal P, Coyne CJ, Ambrose MJ, Maxted N, Schaefer H, Blair MW, Varshney RK. Legume crops phylogeny and genetic diversity for science and breeding. *Crit Rev Plant Sci*. 2015;34(1-3):43-104.
10. Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. *Mol Biol Evol*. 2021;38(7):3022-3027.
11. Varshney RK, Graner A, Sorrells ME. Genomics-assisted breeding for crop improvement. *Trends Plant Sci*. 2005;10(12):621-630.
12. Varshney RK, Roorkiwal M, Nguyen T. Legume genomics: from genomic resources to molecular breeding. *Plant Genome*. 2013;6(3):1-7.
13. Wendel JF, Doyle JJ. Phylogenetic incongruence: window into genome history and molecular evolution. In: Soltis DE, Soltis PS, Doyle JJ, editors. *Molecular systematics of plants II: DNA sequencing*. Boston: Springer; 1998. p. 265-296.

14. Weyenberg G, Yoshida R. Phylogenetic tree distances. In: Kaliman RM, editor. Encyclopedia of Evolutionary Biology. Oxford: Academic Press; 2016. p. 285-290.
15. Wickland DP, Hanzawa Y. The FLOWERING LOCUS T/TERMINAL FLOWER 1 gene family: functional evolution and molecular mechanisms. Mol Plant Biol. 2015;8(7):983-997.
16. Wong AC, Hecht VF, Picard K, Diwadkar P, Laurie RE, Wen J, Weller JL. Isolation and functional analysis of CONSTANS-LIKE genes suggests that a central role for CONSTANS in flowering time control is not evolutionarily conserved in *Medicago truncatula*. Front Plant Sci. 2014;5:486.