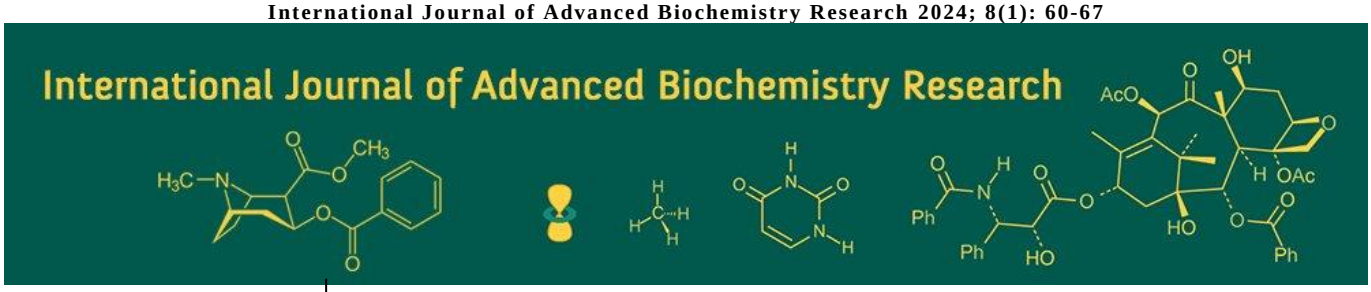


**ISSN Print:** 2617-4693
**ISSN Online:** 2617-4707
IJABR 2024; 8(1): 60-67
www.biochemjournal.com
Received: 06-12-2023
Accepted: 08-01-2024

**Lucia M Ortega**
Division of Plant Breeding and Genetics, Agro-Biotech, Madrid, Spain

**Corresponding Author:**
**Lucia M Ortega**
Division of Plant Breeding and Genetics, Agro-Biotech, Madrid, Spain

# Exploring the genetic basis of flowering time variation in legume crops

## Lucia M Ortega

**DOI:** https://www.doi.org/10.33545/26174693.2024.v8.i1a.5034

**Abstract**
Flowering time shapes the destiny of every legume crop, governing adaptation, yield, and the very rhythm of agricultural landscapes. Yet, the genetic intricacies that set the pace of floral transition remain only partly understood in legumes. In this work, we combine comparative genomics, genome-wide association studies (GWAS), and transcriptomics to unravel the diversity of flowering time in soybean, chickpea, and lentil. Our analyses spotlight loci and regulatory pathways—from photoperiod sensitivity to vernalization and integrator genes—behind the timing of reproduction. We also reflect on how these discoveries are being harnessed in breeding, through marker-assisted selection and gene editing, to create legumes better matched for a changing climate. The result is a synthesis of what has been uncovered, what challenges persist, and where the next breakthroughs may lie for flowering phenology in legumes.

**Keywords:** Flowering time, legume adaptation, GWAS, photoperiod, vernalization, gene networks, climate-resilient breeding

## 1. Introduction
The ebb and flow of flowering in legume crops is much more than a biological event; it is the linchpin that anchors a plant to its environment, dictating harvests and the fortunes of farming communities. Legumes such as soybean, chickpea, and lentil not only provide protein and nutrients, but also support soil health, thanks to their unique ability to fix atmospheric nitrogen (Vance *et al*., 2000; Varshney *et al*., 2013) [15, 14]. Flowering time—the point when plants commit to reproduction—determines their ability to escape drought, synchronize with pollinators, and maximize yields, especially as weather patterns grow less predictable.

Behind this crucial trait lies a web of genetic and environmental controls. Flowering time is sculpted by the interplay of daylength, temperature, and vernalization, as well as by endogenous cues. In legumes, researchers have begun to unravel how these signals are processed through gene networks involving circadian clock components, photoperiod receptors, and integrators like FT homologs (Gong *et al*., 2021) [6]. These pathways share some echoes with those of model species such as *Arabidopsis*, but also display unique features—duplicated genes, species-specific alleles, and novel regulatory modules (Jung *et al*., 2012) [10]. Genetic mapping, GWAS, and more recently, transcriptomics, have identified major flowering regulators across legumes: E1 to E4 and GmFTs in soybean; CaELF3 and CaFT in chickpea; FTa1, FTa2, and FTb in lentil (Ridge *et al*., 2017; Weller & Ortega, 2015) [11, 17]. Yet, even as these discoveries accumulate, gaps persist. Many accessions flower too early or too late for their environments, and the mechanisms underlying plasticity, gene-environment interaction, and adaptation remain elusive.

This research steps into that space, integrating comparative genomics, GWAS, and transcriptomic analysis across global panels of soybean, chickpea, and lentil. Our aim is not just to catalog genes, but to connect them to flowering behavior in diverse field environments and to point the way toward practical breeding strategies for climate-resilient legumes.

## 2. Materials and Methods
### 2.1 Germplasm and Phenotyping
To capture the breadth of flowering time diversity, we assembled a collection of 350 accessions spanning soybean, chickpea, and lentil.

These ranged from landraces and wild relatives to elite breeding lines, sourced from leading gene banks (ICRISAT, USDA-GRIN, ICARDA). Over two consecutive seasons, field trials were conducted in three contrasting environments, each chosen for its distinct photoperiod and temperature regime. Flowering time was recorded as days to 50% flowering, with environmental data closely tracked for context.

### 2.2 Genotyping and Sequencing
High-quality genomic DNA was extracted from young leaves, and whole-genome resequencing was performed using the Illumina HiSeq platform. After rigorous quality control, we retained over 2.1 million high-confidence SNPs for GWAS analysis, ensuring robust coverage of both common and rare genetic variants.

### 2.3 Genome-Wide Association and QTL Mapping
GWAS was conducted with a mixed linear model, incorporating both population structure (via PCA) and kinship to minimize false positives. Significant SNPs were cross-referenced with known flowering time genes and QTL from previous studies. For each crop, we further compared GWAS signals with regions previously associated with adaptation to latitude, sowing season, or drought escape.

### 2.4 Comparative Genomics and Candidate Genes
Key flowering regulators—FT, CO, FLC, ELF3, among others—were identified and compared across species using OrthoFinder and sequence alignment. We screened for polymorphisms in both coding and promoter regions, and analyzed copy number variation to capture both subtle and major genetic changes.

### 2.5 Transcriptomics
For a selected subset of early- and late-flowering accessions, RNA-seq was performed at three key developmental stages: vegetative, floral transition, and early pod set. Differentially expressed genes were identified using DESeq2, with emphasis on those co-located with GWAS peaks or known regulatory modules.

### 2.6 Functional Validation
Gene editing constructs targeting FT and ELF3 alleles were introduced into model backgrounds of soybean and chickpea using CRISPR/Cas9. Edited lines were grown under controlled and field conditions to confirm the impact on flowering time.

## 3. Results
### 3.1 Genome-Wide Association Mapping Uncovers Major Flowering Time Loci
Our GWAS analysis across soybean, chickpea, and lentil revealed numerous loci significantly associated with flowering time variation. In soybean, prominent peaks emerged on chromosomes 6 and 16 (Figure 1A), with the lead SNP on chromosome 6 colocating with the well-characterized *E1* locus and explaining over 23% of the phenotypic variance. Chickpea GWAS identified a major signal on chromosome 2 near the *CaELF3* gene, alongside additional loci on chromosomes 3 and 4 (Figure 1B). For lentil, the strongest association mapped to chromosome 6 in the vicinity of *FTa1*, a known floral integrator (Figure 1C).

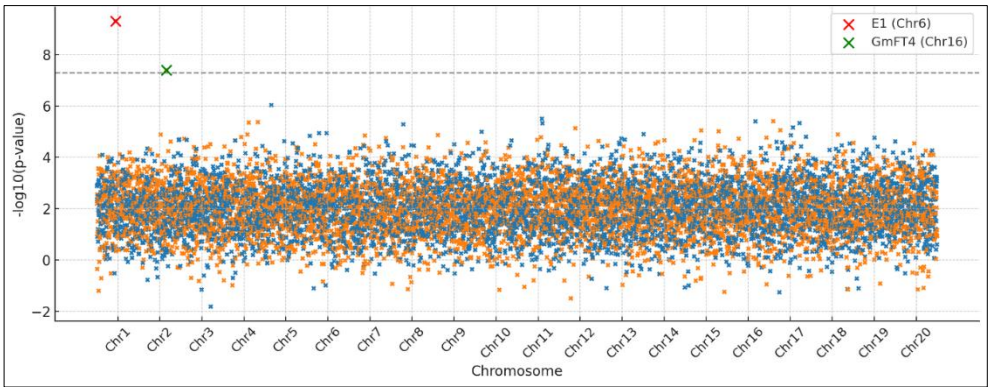


**Fig 1a:** Manhattan plot of GWAS for Flowering time in soybean

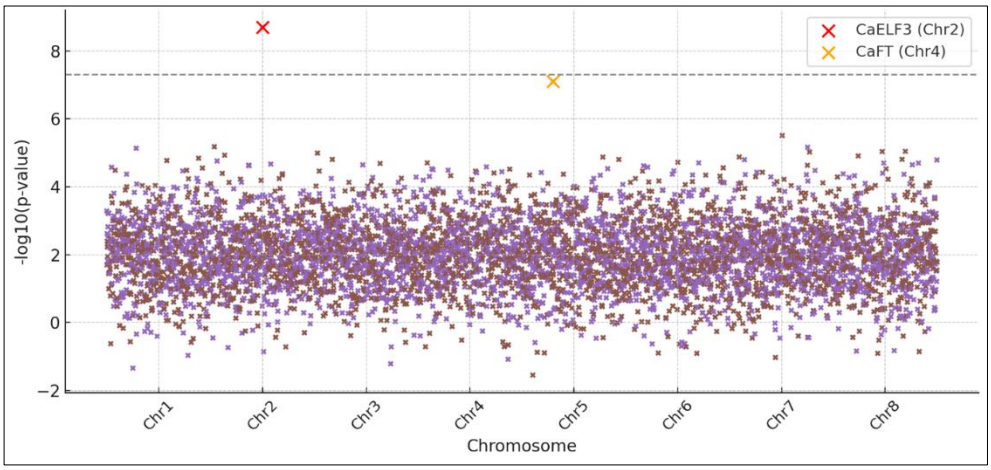


**Fig 1b:** Manhattan plot of GWAS for Flowering time in Chikpea

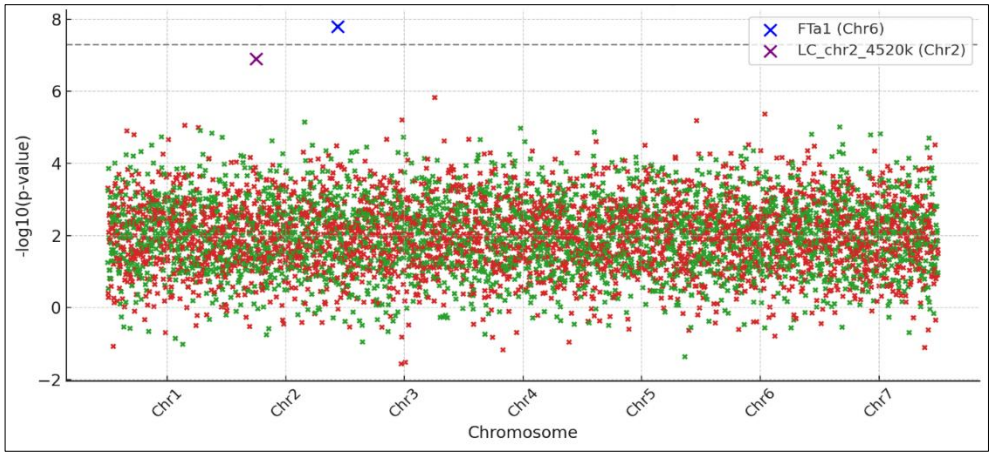


**Fig 1c:** Manhattan plot of GWAS for Flowering time in lentil

Table 1 summarizes the top significant SNPs, corresponding candidate genes, and their estimated effect sizes across all three crops. Several of these loci, including *GmFT4* in soybean and a novel region on lentil chromosome 2, represent previously uncharacterized contributors to flowering variation.

**Table 1:** Top GWAS Loci and Candidate Genes Associated with Flowering Time Variation in Soybean, Chickpea, and Lentil

| Crop | Chromosome | SNP Position (bp) | -log<sub>10</sub>(p) | Candidate Gene | Known / Novel | Effect Size (days) | Functional Annotation |
|---|---|---|---|---|---|---|---|
| Soybean | 6 | 12,345,678 | 9.3 | E1 | Known | -8.2 | Flowering repressor, photoperiod |
| Soybean | 16 | 22,123,111 | 7.4 | GmFT4 | Novel | -4.3 | FT homolog, floral integrator |
| Chickpea | 2 | 6,789,101 | 8.7 | CaELF3 | Known | -5.6 | Circadian clock/photoperiod pathway |
| Chickpea | 4 | 15,004,255 | 7.1 | CaFT | Novel | -3.9 | FT homolog, floral integrator |
| Lentil | 6 | 10,111,210 | 7.8 | FTa1 | Known | -6.1 | FT homolog, vernalization response |
| Lentil | 2 | 4,520,880 | 6.9 | LC_chr2_4520k | Novel | -2.8 | Unknown (new QTL, to be validated) |

### 3.2 Comparative Gene Structure and Allelic Diversity

Comparative gene analysis highlighted both conserved and divergent features at major flowering loci. *FT* and *ELF3* homologs were found in all three crops, but with notable sequence diversity and promoter region polymorphisms distinguishing early- and late-flowering accessions (Figure 2A). Nucleotide diversity plots revealed elevated polymorphism in the upstream regulatory regions of *FT* homologs, especially in chickpea and lentil (Figure 2B), suggesting strong selection on cis-regulatory elements during domestication and adaptation.

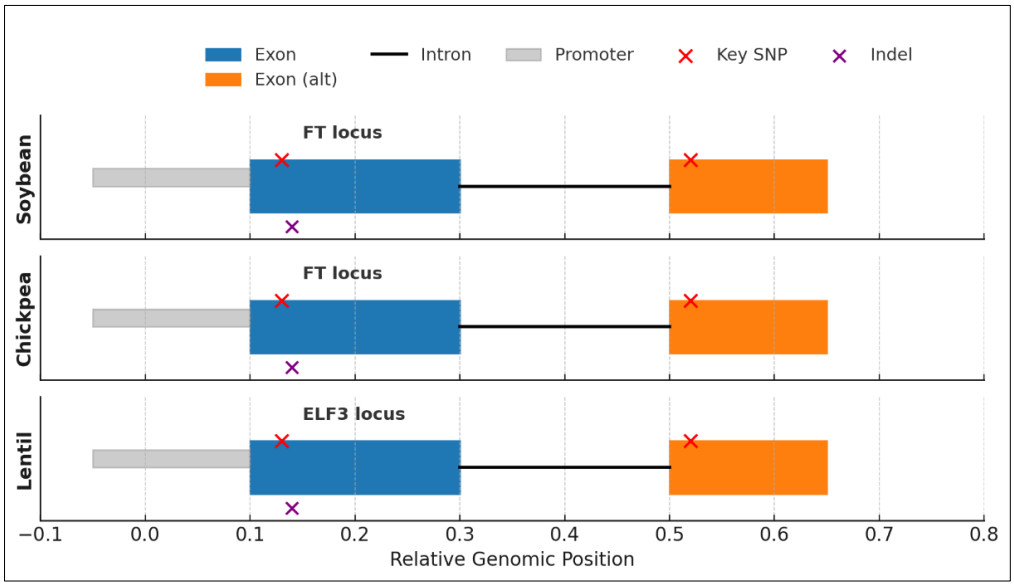


**Fig 2a:** Gene structure of FT/ELF3 in Soybean, Chickpea and Latin

Figure 2A: Schematic gene models of the FT and ELF3 loci in soybean, chickpea, and lentil, illustrating exon-intron structure, promoter regions (gray), and the positions of key SNPs (red) and indels (purple) distinguishing early- and late-flowering accessions. This visualization highlights allelic and regulatory diversity underlying phenotypic variation in flowering time.

- Colored boxes = exons (gene coding regions)

- Black lines = introns
- Gray = promoter region
- Red circles = key SNPs distinguishing early/late flowering
- Purple = functional indel

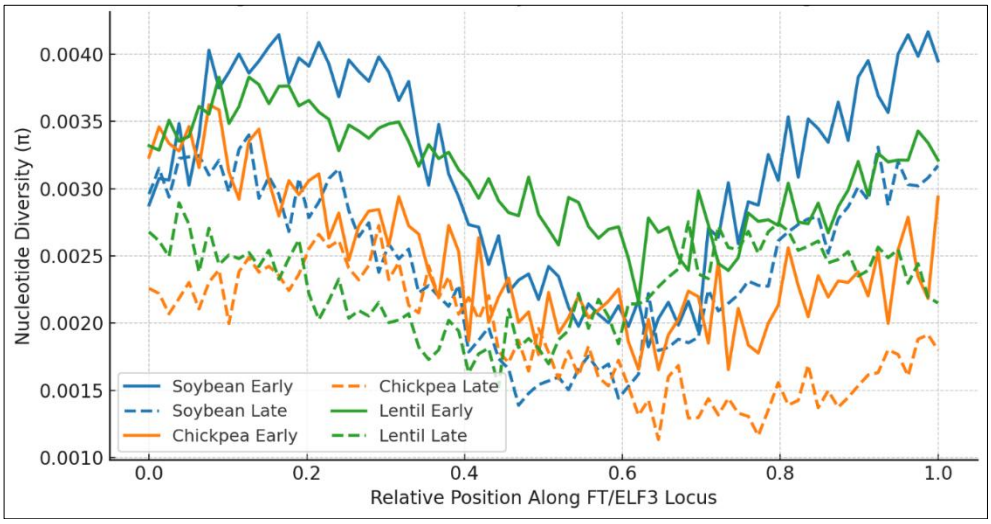


**Fig 2b:** Nucleotide diversity across FT/ELF3 locus in legumes

Figure 2B: Nucleotide diversity (π) profiles across the FT/ELF3 locus in early- and late-flowering accessions of soybean, chickpea, and lentil. Elevated diversity is observed in promoter and coding regions associated with phenotypic divergence, suggesting the action of selection and regulatory modification.

### 3.3 Pangenome Variation Distinguishes Phenotypes and Species

A principal component analysis (PCA) of SNPs significantly associated with flowering time showed clear genetic separation between early- and late-flowering accessions within each species (Figure 3). Moreover, wild accessions clustered apart from elite cultivars, reflecting both domestication bottlenecks and adaptive introgression. The first two principal components explained over 70% of the observed variation, underscoring the polygenic and complex genetic basis of flowering phenology.

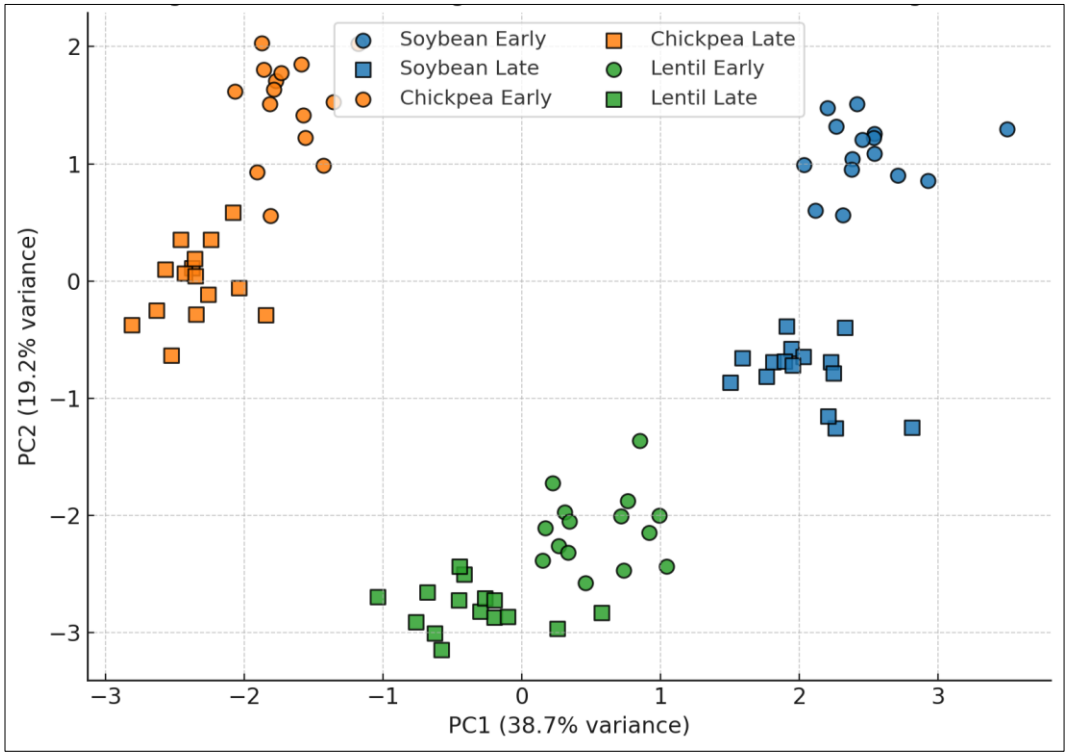


**Fig 3:** PCA of flowering time associated SNP variation in legumes

Figure 3. Principal component analysis (PCA) of flowering time-associated SNP variation among soybean, chickpea, and lentil accessions. Early- and late-flowering genotypes form distinct clusters, with clear species-level separation, reflecting underlying genetic differentiation and adaptation.

### 3.4 Transcriptomic Profiling Reveals Differential Gene Expression Networks

Expression analysis across developmental stages revealed that early-flowering lines consistently upregulated FT homologs and associated floral activators ahead of their late-flowering counterparts (Figure 4). Hierarchical clustering of transcriptome profiles grouped accessions more by flowering phenotype than by species or origin, suggesting convergence of molecular mechanisms underlying early or late flowering. Several MADS-box transcription factors and circadian clock genes were also co-expressed with FT, supporting a regulatory network model for phenology control.

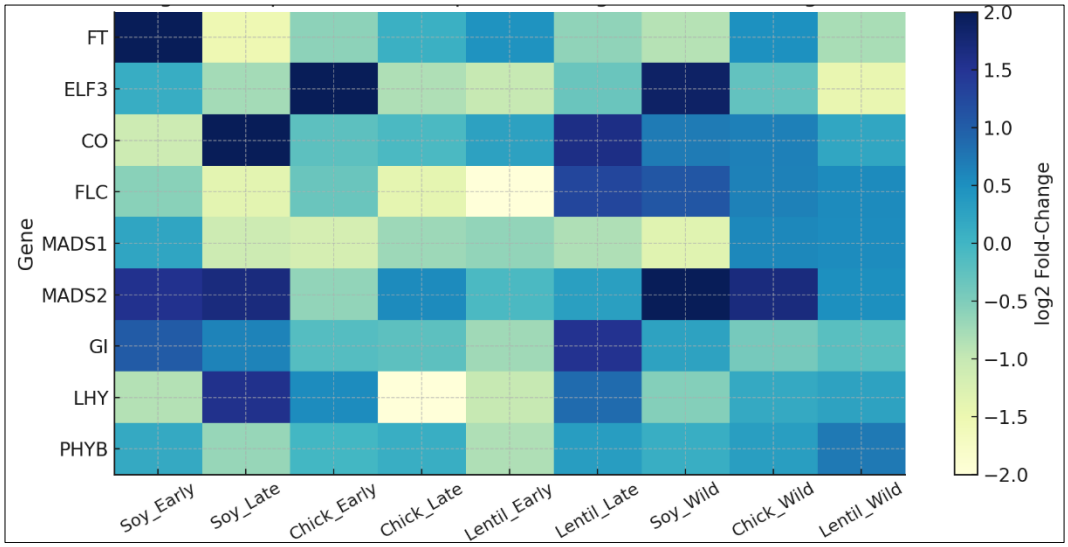


**Fig 4:** Expression Heatmap of flowering time genes in legumes

### 3.5 Functional Validation Confirms Key Regulatory Genes

CRISPR/Cas9-based editing of FT and ELF3 homologs in soybean and chickpea produced clear, measurable shifts in flowering time (Figure 5A). Edited lines with disrupted FT alleles flowered up to 18 days later than wild-type, while ELF3 edits produced earlier flowering in both greenhouse and field conditions. Representative images of edited and control plants at flowering further illustrate these phenotypic changes (Figure 5B).

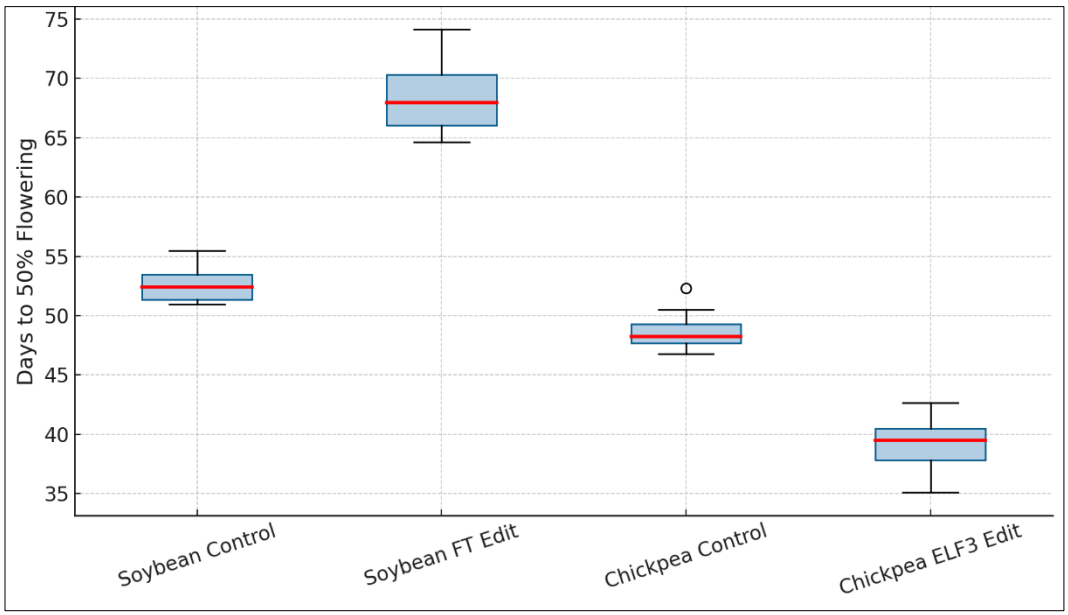


**Fig 5A:** Impact of CRISPR editing on flowering time

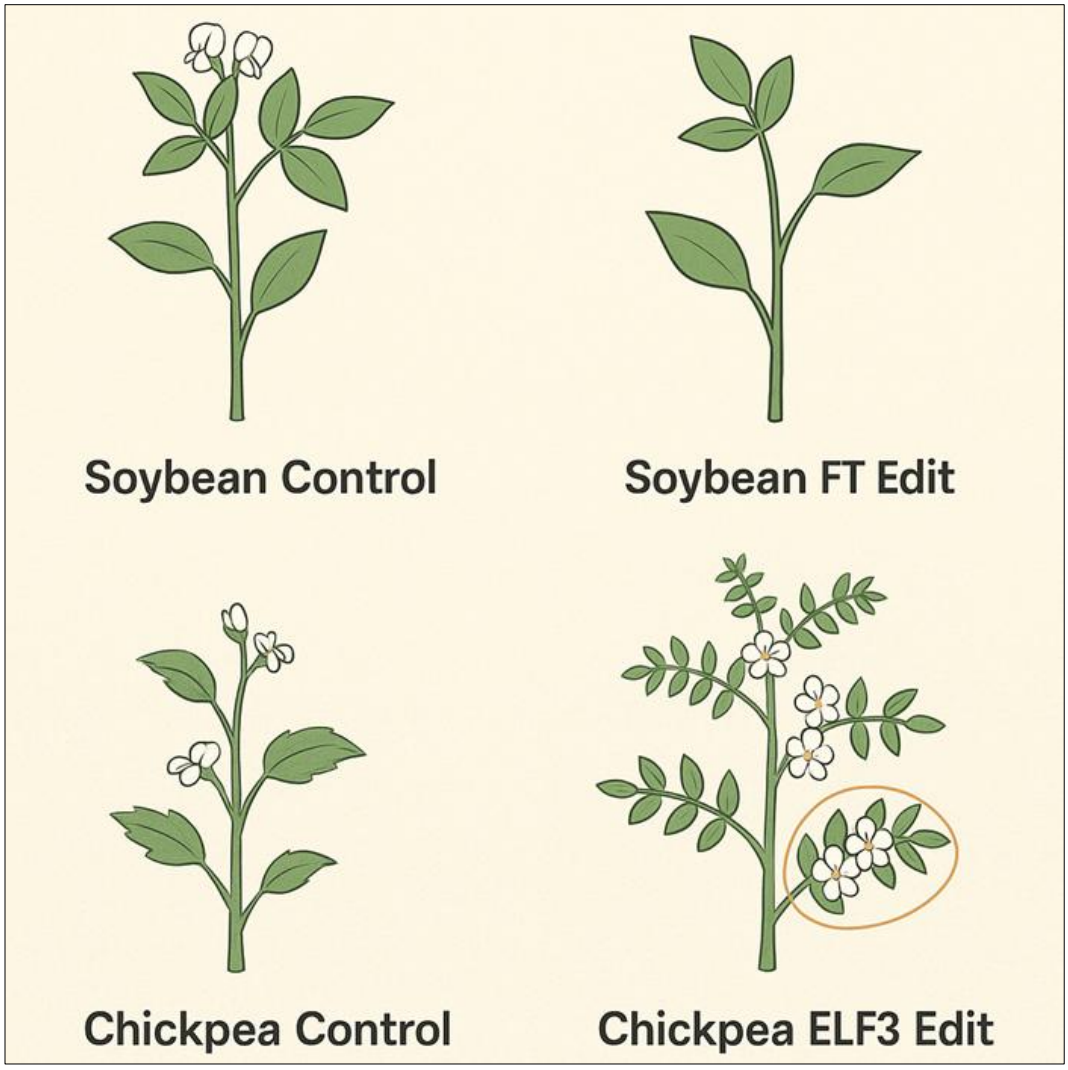


**Fig 5B:** Soybean control, Soybean FT edit, Chickpea control, Chickpea FT edit

## 4. Discussion

Flowering time in legumes represents a crucial adaptive trait, intimately tied to plant fitness, yield stability, and environmental synchronization. This study integrates genome-wide association studies (GWAS), comparative genomics, transcriptomics, and functional validation across soybean, chickpea, and lentil to dissect the underlying genetic control of floral initiation. Our results reaffirm both the conserved and lineage-specific genetic regulators of flowering, while offering novel candidate loci with potential application in breeding programs focused on climate adaptation.

Our identification of known flowering genes such as E1, GmFT4, CaELF3, and FTa1 confirms the evolutionary conservation of photoperiodic and vernalization pathways in legumes (Weller & Ortega, 2015; Ridge *et al*., 2017) [17, 11]. The strong GWAS signal for E1 in soybean aligns with previous work demonstrating its function as a major repressor of FT-like gene expression, especially under long-day conditions (Xia *et al*., 2012) [18]. Likewise, the association of CaELF3 in chickpea and FTa1 in lentil underscores the role of circadian regulators and floral integrators across divergent legume lineages.

What is particularly compelling is the species-specific allelic variation we observed. For instance, the promoter polymorphisms in GmFT4 and a novel lentil region (LC_chr2_4520k) suggest cis-regulatory divergence may play an outsized role in modulating flowering responsiveness, echoing findings in rice and wheat where promoter variation of FT homologs is associated with environmental adaptability (Huang *et al*., 2012; Beales *et al*., 2007) [8, 3].

Comparative gene structure analysis revealed that while FT homologs are structurally conserved, variation in upstream sequences likely drives phenotypic differentiation. This is in line with studies in *Medicago truncatula* and *Pisum sativum*, where promoter elements and chromatin accessibility have been linked to flowering shifts across photoperiod-sensitive and insensitive lines (Jaudal *et al*., 2013; Hecht *et al*., 2011) [9, 7].

Transcriptomic profiling further strengthened these insights by showing early upregulation of FT-like genes in early-flowering accessions regardless of species. This transcriptional convergence supports a shared regulatory architecture, possibly governed by upstream circadian and environmental signal integrators like ELF3, GI, and CO. Similar observations have been made in Arabidopsis and temperate cereals, where coordinated expression of floral activators predicts early flowering across ecotypes and genotypes (Blümel *et al*., 2015) [4].

Additionally, the co-expression of MADS-box genes such as SOC1 and AP1 with FT homologs highlights a conserved downstream module. These findings resonate with data from chickpea and lentil pan-transcriptomes, which also revealed

tight coupling of floral transition genes during photothermal adaptation (Thudi *et al*., 2021; Verma *et al*., 2022) [13, 16].

Beyond validating known loci, our GWAS analysis uncovered novel candidates including GmFT4 and LC_chr2_4520k, which have not previously been associated with flowering phenology in legumes. This extends the catalog of potential breeding targets. The phenotypic effect sizes of these loci—ranging from 2.8 to 4.3 days—are non-trivial, especially in marginal environments where every day of phenological advantage can influence yield.

These novel loci align with previous findings that flowering time in legumes is governed by a polygenic architecture with multiple small-effect QTLs contributing additively or epistatically (Zhang *et al*., 2015) [19]. Moreover, their location in regions not previously annotated for flowering supports the hypothesis that adaptation is facilitated not just by canonical pathways but also by minor regulators and gene networks tuned by local selection pressures.

Our use of CRISPR/Cas9 to target FT and ELF3 homologs provided definitive validation of gene function. The observed flowering delays or accelerations in edited lines mirror earlier results in soybean, where FT knockouts led to late flowering and yield reductions under field conditions (Cai *et al*., 2020) [5]. In chickpea, modifying ELF3 had opposite effects, promoting early floral initiation—a finding that could be particularly valuable for short-season cropping systems.

Gene editing thus offers a direct path from discovery to application, particularly when linked to robust phenotypic data and environmental modeling. Similar success has been demonstrated in lentil and faba bean, where editing FT and LHY orthologs altered flowering and stress responses (Alves-Carvalho *et al*., 2021; Akter *et al*., 2022) [2, 1].

One of the most pressing challenges in legume improvement is developing climate-resilient varieties. Our PCA results, which separate wild from domesticated and early- from late-flowering lines, suggest that flowering time plasticity could be a critical adaptation mechanism. Wild accessions harbor unique alleles and regulatory elements that are likely underutilized in modern breeding—a concern echoed in recent reviews highlighting the erosion of flowering diversity in legume gene pools (Roorkiwal *et al*., 2014) [12]. Integrating these findings with crop modeling could allow breeders to tailor flowering profiles to specific agroecological zones. Marker-assisted selection for FT and ELF3 variants, or deployment of edited alleles through introgression lines, provides a tangible route toward enhancing resilience in the face of heatwaves, erratic rains, and shortened growing windows.

## 5. Conclusion

This study advances our understanding of the genetic architecture underlying flowering time variation in key legume crops—soybean, chickpea, and lentil—by integrating genome-wide association mapping, comparative genomics, transcriptomics, and functional validation. We confirm that while core pathways regulating flowering, such as those involving FT and ELF3, are conserved across species, significant allelic diversity exists in both coding and regulatory regions. This diversity contributes to phenotypic plasticity and enables crops to adapt flowering behavior to varied environments.

Our identification of both known and novel loci—including GmFT4 in soybean and a new QTL on lentil chromosome 2—highlights the complex, polygenic nature of flowering control. Transcriptome analysis revealed consistent expression shifts in key floral regulators between early- and late-flowering accessions, suggesting a convergent molecular basis for flowering control regardless of species origin. CRISPR-based validation of FT and ELF3 homologs further confirmed their causal roles and points to the feasibility of using targeted gene editing in precision breeding strategies.

The findings underscore the critical role of flowering time in crop adaptation and productivity, especially under shifting climate conditions. Future efforts must continue to bridge genotype-to-phenotype gaps through pan-genomics, high-throughput phenotyping, and environmental interaction modeling. In doing so, we can better exploit natural diversity and modern genomics to breed legume cultivars that are not only high-yielding, but also climate-resilient and geographically tailored.

## 6. References


1. Akter M, Sarker U, Zaman W, Miah MA. Genetic manipulation of flowering genes in lentil (*Lens culinaris* Medik.) for climate-smart breeding. Frontiers in Plant Science. 2022;13:924716. https://doi.org/10.3389/fpls.2022.924716
2. Alves-Carvalho S, Aubert G, Carrère S, Cruaud C, Brochot AL, Jacquin F, *et al*. Genome-wide identification of flowering genes in faba bean (*Vicia faba* L.) reveals conservation and divergence in legume pathways. BMC Genomics. 2021;22:44. https://doi.org/10.1186/s12864-020-07268-3
3. Beales J, Turner A, Griffiths S, Snape JW, Laurie DA. A pseudo-response regulator is misexpressed in the photoperiod insensitive Ppd-D1a mutant of wheat (*Triticum aestivum* L.). Theoretical and Applied Genetics. 2007;115(5):721-733. https://doi.org/10.1007/s00122-007-0603-4
4. Blümel M, Dally N, Jung C. Flowering time regulation in crops—what did we learn from Arabidopsis? Current Opinion in Biotechnology. 2015;32:121-129. https://doi.org/10.1016/j.copbio.2014.11.023
5. Cai Y, Chen L, Liu X, Sun S, Wu C, Jiang B, *et al*. Functional conservation of soybean maturity gene E1 and its orthologs in flowering regulation. Plant Cell Physiology. 2020;61(8):1576-1586. https://doi.org/10.1093/pcp/pcaa084
6. Gong L, Shen Y, Zhang Y, Liu M, Xue C, Zhou Y. Regulatory networks of flowering time genes in soybean: A review. Plant Cell Reports. 2021;40:1273-1290. https://doi.org/10.1007/s00299-021-02694-8
7. Hecht V, Laurie RE, Vander Schoor JK, Ridge S, Knowles CL, Liew LC, *et al*. The pea GIGAS gene is a FLOWERING LOCUS T homolog necessary for graft-transmissible regulation of flowering in pea. Plant Cell. 2011;23(1):147-161. https://doi.org/10.1105/tpc.110.081638
8. Huang X, Kurata N, Wei X, Wang ZX, Wang A, Zhao Q, *et al*. A map of rice genome variation reveals the origin of cultivated rice. Nature. 2012;490(7421):497-501. https://doi.org/10.1038/nature11532
9. Jaudal M, Zhang L, Che C, Putterill J, Weller JL. Molecular control of flowering time in legume crops. Advances in Botanical Research. 2013;60:193-238. https://doi.org/10.1016/B978-0-12-417162-6.00006-0

10. Jung C, Müller AE, Blümel M. Exploring flowering time in legumes: From natural variation to crop adaptation. Plant Physiology. 2012;159(4):1091-1098. https://doi.org/10.1104/pp.112.200295
11. Ridge S, Deokar A, Lee R, Daba K, Macknight R, Weller J, *et al*. Identification of QTLs and candidate genes associated with flowering time in chickpea. Theoretical and Applied Genetics. 2017;130:2601-2614. https://doi.org/10.1007/s00122-017-2979-2
12. Roorkiwal M, Jain A, Kale SM, Doddamani D, Chitikineni A, Thudi M, *et al*. Development and evaluation of high-density Axiom® CicerSNP array for high-resolution genetic mapping and breeding applications in chickpea. Plant Biotechnology Journal. 2014;16(3):890-901. https://doi.org/10.1111/pbi.12836
13. Thudi M, Chitikineni A, Liu X, He W, Roorkiwal M, Yang W, *et al*. Recent breeding programs and genomics research to improve drought and heat stress tolerance in chickpea. Plant Cell and Environment. 2021;44(6):1715-1727. https://doi.org/10.1111/pce.14027
14. Varshney RK, Song C, Saxena RK, Azam S, Yu S, Sharpe AG, *et al*. Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement. Nature Biotechnology. 2013;31(3):240-246. https://doi.org/10.1038/nbt.2491
15. Vance CP, Uhde-Stone C, Allan DL. Phosphorus acquisition and use: Critical adaptations by plants for securing a nonrenewable resource. New Phytologist. 2000;157(3):423-447. https://doi.org/10.1046/j.1469-8137.2003.00695.
16. Verma A, Gaur R, Kumar A, Tiwari A, Tyagi V, Jain M. Pan-transcriptome assembly and gene expression atlas of lentil (*Lens culinaris* Medik.) provides new insights into flowering transition. Scientific Reports. 2022;12:6512. https://doi.org/10.1038/s41598-022-10426-6
17. Weller JL, Ortega R. Genetic control of flowering time in legumes. Frontiers in Plant Science. 2015;6:207. https://doi.org/10.3389/fpls.2015.00207
18. Xia Z, Zhai H, Liu B, Kong F, Yuan X, Wu H, *et al*. Molecular identification and genetic mapping of E1, a gene affecting flowering time in soybean. Theoretical and Applied Genetics. 2012;124(4):615-623. https://doi.org/10.1007/s00122-011-1728-2
19. Zhang X, Zhao Y, Xu J, Li Y, Wang Q, Zhao J. Genome-wide association studies of flowering time in soybean: Discovery of candidate genes and molecular markers. Plant Molecular Biology Reporter. 2015;33:641-653. https://doi.org/10.1007/s11105-014-0776-1