

# Advances in Phylogenomic Studies of the Fabaceae Family: Resolving Complex Evolutionary Relationships

Xuanjun Fang ✉

Hainan Institute of Tropical Agricultural Resources, Sanya, 572025, Hainan, China

✉ Corresponding email: [Xuanjun.fang@hitar.org](mailto:Xuanjun.fang@hitar.org)

Legume Genomics and Genetics, 2024 Vol.15, No.1 doi: [10.5376/lgg.2024.15.0004](https://doi.org/10.5376/lgg.2024.15.0004)

Received: 10 Jan., 2024

Accepted: 11 Feb., 2024

Published: 22 Feb., 2024

**Copyright © 2024** Fang, This is an open access article published under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Preferred citation for this article:

Fang X.J., 2024, Advances in phylogenomic studies of the fabaceae family: resolving complex evolutionary relationships, Legume Genomics and Genetics, 15(1): 27-36 (doi: [10.5376/lgg.2024.15.0004](https://doi.org/10.5376/lgg.2024.15.0004))

**Abstract** The objective of this study is to summarize recent advances in phylogenomic studies of the Fabaceae family, with a focus on resolving complex evolutionary relationships. The Fabaceae family, one of the most diverse and economically significant plant families, presents unique challenges in understanding its evolutionary history due to its extensive diversity and complex speciation events. This study aims to highlight key findings and methodological innovations that have advanced our understanding of the phylogeny of this family. Recent phylogenomic studies have provided new insights into the evolutionary history of the Fabaceae family. High-throughput sequencing technologies and advanced bioinformatics tools have enabled the identification of major clades, divergence times, and lineage-specific adaptations. These studies have resolved previously unclear relationships and have offered a more comprehensive view of the evolutionary trajectories within the family. In conclusion, the study underscores that recent phylogenomic studies have significantly enhanced our understanding of the evolutionary relationships within the Fabaceae family. Continued research, coupled with the integration of emerging technologies, is crucial for further resolving complex phylogenetic relationships and supporting conservation efforts. The findings from these studies have important implications for plant genomics, agriculture, and biodiversity conservation.

**Keywords** Fabaceae; Phylogenomics; Evolutionary relationships; High-throughput sequencing; Bioinformatics

## 1 Introduction

The Fabaceae family, also known as Leguminosae, is the third largest family of flowering plants, comprising approximately 19 500 species across 765 genera (Zhao et al., 2021). This family is of immense ecological and economic importance, contributing significantly to global agriculture through crops such as soybean, chickpea, and common bean (Kagale and Close, 2021). Fabaceae species are known for their ability to fix atmospheric nitrogen through symbiotic relationships with rhizobia, which enhances soil fertility and reduces the need for synthetic fertilizers (Zhao et al., 2021). This nitrogen-fixing ability, along with their nutritional value, makes legumes crucial for food security, environmental sustainability, and human health (Kagale and Close, 2021).

Phylogenomics, the intersection of phylogenetics and genomics, involves the analysis of genome-scale data to infer evolutionary relationships among species. This approach has become increasingly important for resolving complex phylogenetic relationships that are difficult to decipher using traditional methods (Koenen et al., 2019). In the context of the Fabaceae family, phylogenomic studies have provided insights into deep-branching relationships, polyploidization events, and the evolutionary history of nitrogen-fixing symbiosis (Koenen et al., 2019; Zhao et al., 2021). Advances in next-generation sequencing technologies have facilitated the generation of large-scale genomic data, enabling more accurate and comprehensive phylogenetic analyses (Kagale and Close, 2021).

This study aims to provide a detailed overview of recent advances in phylogenomic research of the Fabaceae family. By synthesizing data from multiple sources, we will present a comprehensive picture of the current understanding of evolutionary relationships within this diverse family. The study will emphasize key findings from recent research, such as the resolution of deep-branching relationships among subfamilies, the identification

of polyploidization events, and the reconstruction of the evolutionary history of nitrogen-fixing symbiosis. We will also review the methodologies used in these studies, including nuclear and chloroplast gene alignments, maximum likelihood, Bayesian inference, and multispecies coalescent summary methods. Additionally, this study will identify gaps in the current knowledge of Fabaceae phylogenomics and propose directions for future research. Despite significant advancements, challenges persist in resolving specific phylogenetic relationships and understanding the evolutionary mechanisms driving trait diversity in this family. Future research should aim to generate more comprehensive genomic datasets, improve phylogenetic resolution, and investigate the functional implications of evolutionary events such as polyploidization and hybridization. By addressing these objectives, this study will enhance the understanding of the evolutionary history and diversity of the Fabaceae family, providing a foundation for future research and applications in plant breeding and conservation.

## **2 Historical Context of Fabaceae Phylogeny**

### **2.1 Early studies**

Early studies on the phylogeny of the Fabaceae family primarily relied on morphological and anatomical characteristics to classify and understand the evolutionary relationships among different species. These traditional methods, while foundational, often faced limitations due to convergent evolution and phenotypic plasticity, which could obscure true phylogenetic relationships. The advent of molecular techniques, particularly DNA barcoding, marked a significant advancement in the study of Fabaceae phylogeny. DNA barcoding, which involves the use of specific gene regions to identify species, provided a more objective and accurate method for species identification compared to morphological approaches. For instance, the MatK gene of the chloroplast has been utilized effectively for barcoding in various legume species, demonstrating high accuracy and simplicity in species identification (Abdelsalam et al., 2021). These early molecular studies laid the groundwork for more comprehensive phylogenetic analyses by providing initial insights into the genetic relationships within the family.

### **2.2 Development of phylogenomics**

The introduction of high-throughput sequencing technologies revolutionized phylogenetic studies by enabling the analysis of large-scale genomic data. This technological advancement allowed researchers to move beyond single-gene analyses to more comprehensive whole-genome approaches. High-throughput sequencing facilitated the acquisition of complete plastome sequences and mitochondrial protein-coding gene sequences, which provided deeper insights into the evolutionary history and relationships within plant families (McManus et al., 2018). The transition from single-gene to whole-genome analyses marked a significant leap in resolving complex phylogenetic relationships. In the case of the Fabaceae family, large-scale genomic sequence data have been instrumental in resolving deep divergences and understanding the evolutionary origins of its subfamilies. For example, a study utilizing alignments of 72 chloroplast genes and 7 621 homologous nuclear-encoded proteins successfully resolved the deepest divergences in the legume phylogeny, highlighting the near-simultaneous evolutionary origin of all six subfamilies (Figure 1) (Koenen et al., 2019). This comprehensive approach provided a robust phylogenetic framework, overcoming the limitations of earlier methods that relied on fewer genetic markers. Overall, the development of phylogenomics has significantly advanced our understanding of the Fabaceae family's evolutionary history, enabling researchers to resolve complex relationships that were previously unattainable with traditional and early molecular methods.

## **3 Methodological Advances in Phylogenomics**

### **3.1 High-throughput sequencing technologies**

High-throughput sequencing (HTS) technologies have revolutionized phylogenomic studies by enabling the rapid and cost-effective generation of large-scale sequence data. Platforms such as Illumina and PacBio are at the forefront of these advancements. Illumina sequencing, known for its high accuracy and throughput, is widely used for generating short-read data, which is essential for various genomic applications, including phylogenetic analyses (Mandel et al., 2015; Uribe-Convers et al., 2016). On the other hand, PacBio's

single-molecule real-time (SMRT) sequencing provides long-read data, which is particularly useful for resolving complex genomic regions and studying polyploid species (Rothfels et al., 2017; Lyu et al., 2021).

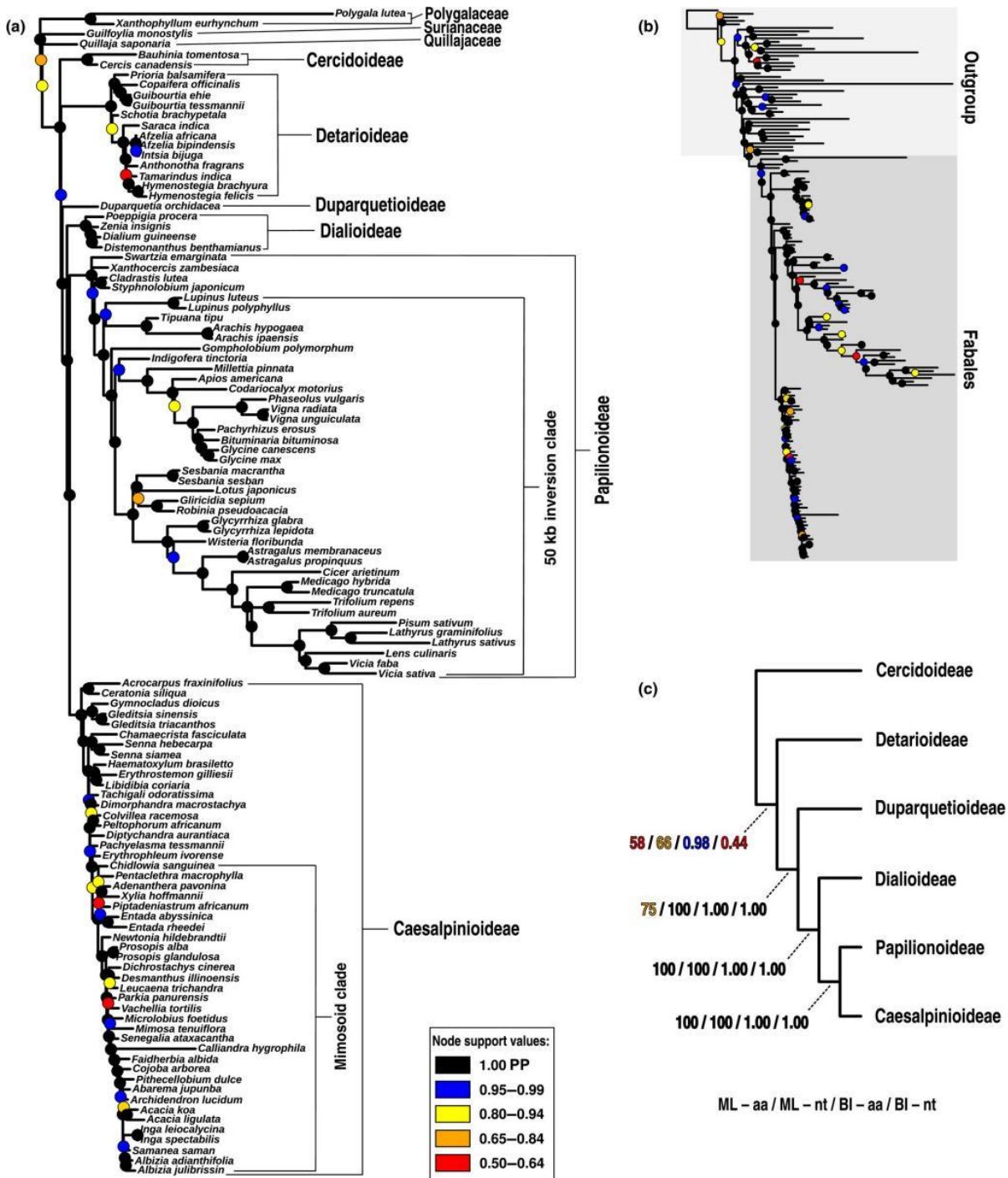


Figure 1 Phylogeny of legumes based on Bayesian analyses of 72 protein-coding chloroplast genes under the CATGTR model in Phylobayes (Adopted from Koenen et al., 2019)

Image caption: a: Majority-rule consensus tree of the amino acid alignment, displaying only the Fabales portion, with outgroup taxa removed; b: Complete tree including outgroup taxa; c: Simplified tree showing support for subfamily relationships using different inference methods (ML: Maximum Likelihood; BI: Bayesian Inference) and sequence types (aa: Amino acids; nt: Nucleotides); Majority-rule consensus trees for both the amino acid and nucleotide alignments, with tip labels for all taxa and support values indicated. In panels a and b, colored circles indicate node support in posterior probabilities (PP) (Adapted from Koenen et al., 2019)

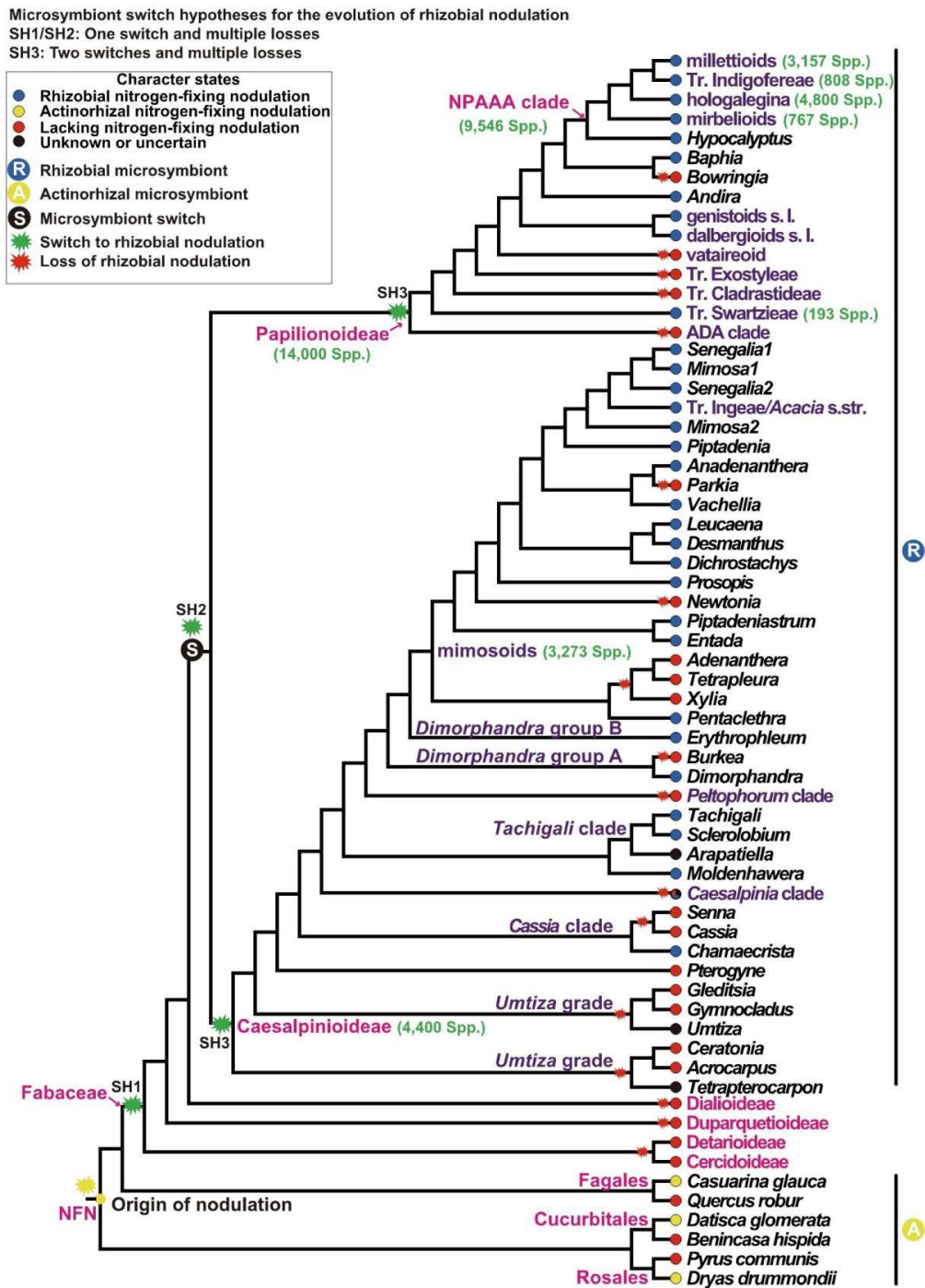


Figure 2 An overview of proposed evolutionary histories of the rhizobial nitrogen-fixing symbiosis in Fabaceae (Adopted from Zhao et al., 2021)

Image caption: Microbiont switch hypotheses for the evolution of rhizobial nodulation in Fabaceae. The phylogenetic tree shows different clades and the state of nitrogen-fixing nodulation; Character states: Blue circles represent rhizobial nitrogen-fixing nodulation; Yellow circles indicate actinorhizal nitrogen-fixing nodulation; Red circles signify a lack of nitrogen-fixing nodulation; Black circles denote unknown or uncertain states; Symbols: "R" indicates a rhizobial microsymbiont; "A" denotes an actinorhizal microsymbiont; "S" signifies a microsymbiont switch; Green stars represent switches to rhizobial nodulation; Red stars indicate losses of rhizobial nodulation; Key clades and their species counts are highlighted in magenta. SH1/SH2 and SH3 indicate different hypotheses of nodulation evolution, with SH1/SH2 suggesting one switch and multiple losses, and SH3 indicating two switches and multiple losses (Adopted from Zhao et al., 2021)

In the context of Fabaceae, HTS technologies have been instrumental in resolving phylogenetic relationships and understanding evolutionary processes. For instance, the use of Illumina sequencing has facilitated the generation of extensive nuclear and chloroplast genomic data, enabling the resolution of deep-branching relationships within the family (Koenen et al., 2019). Additionally, PacBio sequencing has been employed to study the transcriptomes of Fabaceae species, providing insights into gene expression and stress response mechanisms (Lyu et al., 2021). These technologies have also supported the identification of whole-genome duplication events and the reconstruction of the evolutionary history of nitrogen-fixing symbiosis in Fabaceae (Figure 2) (Zhao et al., 2021).

### **3.2 Bioinformatics tools and pipelines**

The vast amount of data generated by HTS requires robust bioinformatics tools and pipelines for processing and assembly. Tools such as the Fluidigm Access Array System and associated bioinformatics pipelines have been developed to handle large datasets efficiently. These pipelines facilitate the generation of consensus sequences and the recovery of allelic information, which are crucial for phylogenetic analyses (Uribe-Convers et al., 2016). Additionally, specialized pipelines like Purc have been designed to manage the complexities of polyploid data, ensuring accurate sequence inference from raw reads (Rothfels et al., 2017).

Phylogenetic tree construction is a critical step in phylogenomics, with methods such as maximum likelihood (ML) and Bayesian inference (BI) being commonly used. These methods provide robust frameworks for inferring evolutionary relationships based on genomic data. For example, ML and BI have been employed to analyze nuclear and chloroplast gene alignments, resolving the deepest divergences in the legume phylogeny (Koenen et al., 2019). Additionally, multispecies coalescent methods have been applied to evaluate support for alternative topologies across gene trees, addressing issues such as incomplete lineage sorting and hybridization (Vargas et al., 2017; Koenen et al., 2019).

### **3.3 Comparative genomics**

Comparative genomics allows for the study of gene family evolution, providing insights into the processes of gene duplication and divergence. Phylogenetic approaches are used to assess gene orthology and paralogy, trace evolutionary changes in gene families, and predict structure-function relationships (Thornton and DeSalle, 2000). In Fabaceae, phylogenomic analyses have identified numerous whole-genome duplication events, shedding light on the evolutionary history and diversification of gene families within the family (Zhao et al., 2021).

Genome synteny and collinearity analyses are essential for understanding the structural organization and evolutionary history of genomes. These analyses involve comparing the order and orientation of genes across different species to identify conserved genomic regions. In Fabaceae, synteny analyses have revealed patterns of genome evolution and provided evidence for ancient polyploidy events (Zhao et al., 2021). Such studies are crucial for reconstructing the evolutionary trajectories of complex plant genomes and understanding the genetic basis of key traits.

By leveraging these methodological advances, researchers can gain a deeper understanding of the evolutionary relationships and genomic architecture of the Fabaceae family, ultimately contributing to the broader field of plant phylogenomics.

## **4 Major Findings in Fabaceae Phylogenomics**

### **4.1 Phylogenetic relationships**

Recent advances in phylogenomic studies have significantly enhanced our understanding of the phylogenetic relationships within the Fabaceae family. One of the major findings is the resolution of the deepest divergences among the subfamilies of Fabaceae. By analyzing large-scale genomic sequence data, researchers have been able to generate robust phylogenetic frameworks that clarify the relationships among the six subfamilies. This study utilized alignments of 72 chloroplast genes and 7 621 homologous nuclear-encoded proteins across 157

and 76 taxa, respectively, employing methods such as maximum likelihood, Bayesian inference, and multispecies coalescent summary methods. The results indicate that all six subfamilies originated nearly simultaneously, challenging the traditional view of some subfamilies as 'basal' or 'early-diverging' (Koenen et al., 2019).

#### **4.2 Evolutionary insights**

The phylogenomic analyses have also provided profound evolutionary insights into the Fabaceae family. One significant insight is the recognition of incomplete lineage sorting as a prevalent evolutionary process within the family. The study found strongly supported conflicts in the nuclear genes, which suggest that incomplete lineage sorting has played a crucial role in the evolutionary history of Fabaceae. This finding underscores the complexity of evolutionary processes and highlights the limitations of phylogenetic resolution when dealing with rapid successive speciation events (Koenen et al., 2019). Additionally, the simultaneous origin of all six subfamilies has important implications for understanding the evolution of legume diversity and traits, suggesting a more complex evolutionary history than previously thought.

#### **4.3 Case studies**

Case studies within the Fabaceae family have further illustrated the utility of phylogenomic approaches in resolving complex evolutionary relationships. For instance, the comparative and phylogenetic analyses of plastid genomes in the Fagaceae family, which is closely related to Fabaceae, have demonstrated the effectiveness of plastid phylogenomics in resolving phylogenetic relationships. This study highlighted the impact of codon composition bias on phylogenetic inference, revealing that the first two codon sites dataset recovered nearly all relationships with high support. Such findings emphasize the importance of considering codon composition bias in phylogenomic studies and provide a methodological framework that can be applied to Fabaceae phylogenomics (Yang et al., 2018).

### **5 Implications for Understanding Evolutionary Processes**

#### **5.1 Gene flow and hybridization**

Gene flow and hybridization are critical processes in the evolution of the Fabaceae family. The extensive phylogenomic analyses conducted on over 1 500 nuclear genes from 391 species have revealed numerous polyploidization events within the family, which are indicative of historical gene flow and hybridization events. These polyploidization events, including whole-genome duplications (WGD) and whole-genome triplications (WGT), have been identified at various ancestral nodes within the Fabaceae, suggesting that hybridization has played a significant role in the diversification and adaptation of this family (Zhao et al., 2021). The study also supports the hypothesis of multiple switches to rhizobial nodulation followed by several losses, which could be attributed to gene flow and hybridization events that facilitated the spread and subsequent loss of this trait (Zhao et al., 2021).

#### **5.2 Molecular evolution**

Molecular evolution within the Fabaceae family has been profoundly influenced by polyploidization events. The phylogenomic analyses have provided a highly resolved phylogeny that supports the monophyly of the subfamilies and reveals the early radiation of these subfamilies near the K/Pg boundary, a period marked by mass extinction. This timing suggests that the molecular evolution of Fabaceae was significantly impacted by global environmental changes (Zhao et al., 2021). The study's molecular clock estimation indicates that most tribe-level clades diverged within approximately 15 million years, highlighting a rapid evolutionary radiation that could be driven by molecular adaptations to changing environments (Zhao et al., 2021). Additionally, the comparative analysis of plastid genomes in related families, such as Fagaceae, underscores the importance of codon composition bias in phylogenetic inference, which can affect our understanding of molecular evolution in these complex families (Yang et al., 2018).

### **5.3 Biogeography**

The biogeographical history of the Fabaceae family is intricately linked to its evolutionary processes. The resolved phylogeny and divergence times provide insights into the historical biogeography of the family. The early radiation of subfamilies near the K/Pg boundary suggests that the Fabaceae family may have originated and diversified in response to the mass extinction event, which could have opened up new ecological niches (Zhao et al., 2021). The subsequent divergences of tribe-level clades within a relatively short time frame indicate that the family rapidly adapted to various geographical regions, leading to its current widespread distribution. The study's findings on the evolution of rhizobial nitrogen-fixing nodulation also have biogeographical implications, as the ability to fix nitrogen would have allowed Fabaceae species to colonize and thrive in nutrient-poor soils across different continents (Zhao et al., 2021).

## **6 Challenges and Limitations**

### **6.1 Technical challenges**

Phylogenomic studies in the Fabaceae family face several technical challenges that can impact the accuracy and reliability of results. One significant issue is the quality and quantity of DNA extracted from herbarium specimens, which are often the only available sources for rare or extinct species. Degraded DNA from these specimens can lead to incomplete or biased genomic data. Additionally, the complexity of the Fabaceae genome, characterized by high levels of polyploidy and large genome sizes, poses challenges in sequencing and assembly. The need for high-throughput sequencing technologies requires substantial financial resources and technical expertise, which may not be readily available in all research settings.

Another technical challenge is the accurate annotation of gene sequences. The Fabaceae family includes species with diverse genomic structures, making it difficult to identify homologous genes across different taxa accurately. This complexity is further compounded by the presence of paralogous genes resulting from genome duplications, which can confuse phylogenomic analyses. Furthermore, the computational demands of analyzing large genomic datasets require access to advanced bioinformatics tools and powerful computing resources, which can be limiting for some research groups.

### **6.2 Analytical challenges**

Analytical challenges in phylogenomic studies of the Fabaceae family are primarily related to the interpretation and integration of complex data sets. One of the main issues is the accurate reconstruction of phylogenetic trees, particularly when dealing with incomplete lineage sorting (ILS) and hybridization events. These phenomena can obscure true evolutionary relationships and lead to conflicting phylogenetic signals. To address these issues, researchers often need to employ sophisticated models and methods that can accommodate such complexities, but these methods are not always straightforward to implement or interpret.

Another analytical challenge is the alignment of large-scale genomic data. Multiple sequence alignments are critical for identifying conserved regions and inferring evolutionary relationships, but the presence of highly divergent sequences can complicate alignment processes and lead to errors. Moreover, the selection of appropriate outgroup species, which is crucial for rooting phylogenetic trees, can be problematic if suitable candidates are not available or if their evolutionary histories are not well-resolved.

### **6.3 Sampling biases**

Sampling biases present a significant limitation in phylogenomic studies of the Fabaceae family. The uneven representation of species within the family can skew results and lead to incomplete or biased conclusions about evolutionary relationships. For instance, certain clades within the Fabaceae are well-studied due to their economic importance or ease of access, while others are underrepresented in genomic databases. This imbalance can result in a distorted view of the family's phylogeny and may overlook important evolutionary events.

Geographic biases in sampling can also affect phylogenomic studies. Many Fabaceae species are endemic to specific regions, and limited sampling from these areas can hinder our understanding of the family's global

diversity and evolutionary history. Additionally, sampling strategies that focus predominantly on modern species without considering extinct or ancient lineages can lead to incomplete phylogenetic reconstructions. This is particularly problematic in the Fabaceae family, where historical diversification events have played a crucial role in shaping current diversity.

In conclusion, while advances in phylogenomics have significantly enhanced our understanding of the Fabaceae family's evolutionary relationships, several challenges and limitations remain. Addressing these technical, analytical, and sampling biases is essential for achieving a more comprehensive and accurate phylogenetic framework for this diverse and ecologically important plant family.

## **7 Future Directions**

### **7.1 Emerging technologies**

The advent of long-read sequencing technologies, such as those provided by PacBio and Oxford Nanopore, offers significant potential for resolving complex evolutionary relationships within the Fabaceae family. These technologies can generate much longer reads compared to traditional short-read sequencing, which helps in assembling more complete and accurate genomes, particularly in regions with high repeat content or structural variations (Mandel et al., 2015). Additionally, single-cell genomics and transcriptomics are emerging as powerful tools to study gene expression and genetic variation at the cellular level, providing insights into the functional aspects of phylogenomic relationships (Zhao et al., 2021). These approaches can help in understanding the heterogeneity within tissues and the evolutionary dynamics of gene expression.

### **7.2 Integrative approaches**

Integrating phylogenomics with other scientific disciplines such as ecology and physiology can provide a more comprehensive understanding of the evolutionary processes in the Fabaceae family. For instance, combining phylogenomic data with ecological information can help in understanding how environmental factors influence evolutionary trajectories and species diversification (Mandel et al., 2015). Similarly, integrating physiological data can shed light on how different species adapt to their environments at a functional level. This multidisciplinary approach can lead to more robust phylogenetic inferences and a deeper understanding of the evolutionary history and adaptive strategies of the Fabaceae family (Zhao et al., 2021).

### **7.3 Conservation applications**

Phylogenomic studies can play a crucial role in conservation planning by providing detailed insights into the genetic diversity and evolutionary history of species within the Fabaceae family. Understanding the genetic diversity and structure of populations can inform conservation strategies aimed at preserving genetic resources and enhancing the resilience of species to environmental changes (Zhao et al., 2021). Additionally, phylogenomic data can help identify evolutionary significant units (ESUs) and prioritize species or populations for conservation based on their unique genetic makeup and evolutionary history. This information is vital for developing effective conservation policies and ensuring the long-term survival of Fabaceae species in their natural habitats (Mandel et al., 2015; Zhao et al., 2021).

By leveraging these emerging technologies, integrative approaches, and conservation applications, future research can significantly advance our understanding of the complex evolutionary relationships within the Fabaceae family and contribute to their preservation and sustainable use.

## **8 Concluding Remarks**

Recent advances in phylogenomic studies have significantly enhanced our understanding of the complex evolutionary relationships within the Fabaceae family. Key findings include:

**Deep Phylogenetic Resolution:** Large-scale genomic sequence data have resolved the deepest divergences in the legume phylogeny, revealing that all six subfamilies originated nearly simultaneously. This challenges the traditional view of some subfamilies being 'basal' or 'early-diverging'.



**Polyploidization Events:** Phylogenomic analyses have identified numerous whole-genome duplication (WGD) and whole-genome triplication (WGT) events across the Fabaceae family. These events are crucial for understanding the evolutionary history and diversification of the family.

**Mitochondrial Genome Variability:** Studies have shown significant fluctuations in mitochondrial genome size and content within Fabaceae, driven by both ancient and recent expansions. These variations are attributed to intercellular gene transfer, horizontal gene transfer, and duplication of native mitochondrial sequences.

**Phylogenetic Methods and Challenges:** The use of next-generation sequencing and advanced phylogenomic methods has been instrumental in resolving contentious relationships within large plant families. However, the resolution of some branches may depend on a small subset of genes, highlighting the need for careful evaluation of phylogenetic signal distribution.

The continued research in Fabaceae phylogenomics is of paramount importance for several reasons:

**Ecological and Economic Significance:** The Fabaceae family includes many species that are vital for ecosystems and agriculture, particularly due to their nitrogen-fixing capabilities. Understanding their evolutionary relationships can inform conservation strategies and agricultural practices.

**Complex Evolutionary Processes:** The family exhibits complex evolutionary processes such as incomplete lineage sorting, hybridization, and polyploidization. Continued research can provide deeper insights into these processes and their implications for plant evolution.

**Call for Collaboration and Data Sharing:** The complexity and scale of phylogenomic studies necessitate collaboration among researchers and institutions. Sharing genomic data and resources can accelerate discoveries and enhance the robustness of phylogenetic inferences.

In conclusion, the advances in phylogenomic studies of the Fabaceae family have provided a more comprehensive understanding of their evolutionary history. However, ongoing research and collaboration are essential to fully unravel the complexities of their phylogeny and to leverage this knowledge for ecological and agricultural benefits.

## **Acknowledgments**

The author extends sincere thanks to two anonymous peer reviewers for their invaluable feedback on the manuscript.

## **Conflict of Interest Disclosure**

The author affirms that this research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

## **References**

- Abdelsalam N., Hasan M., Rabie S., El-wakeel H., Zaitoun A., Ghareeb R., Abdelsalam A., Aly H., Ibrahim A., and Hemeida A., 2021, Endorsement and phylogenetic analysis of some Fabaceae plants based on DNA barcoding, *Molecular Biology Reports*, 49: 5645-5657.  
<https://doi.org/10.1101/2021.07.27.454001>
- Kagale S., and Close T., 2021, Legumes: embracing the genome era, *Legume Science*, 3(3): e113.  
<https://doi.org/10.1002/leg3.113>
- Koenen E., Ojeda D., Steeves R., Migliore J., Bakker F., Wieringa J., Kidner C., Hardy O., Pennington R., Bruneau, A., and Hughes C., 2019, Large-scale genomic sequence data resolve the deepest divergences in the legume phylogeny and support a near-simultaneous evolutionary origin of all six subfamilies, *The New Phytologist*, 225: 1355-1369.  
<https://doi.org/10.1111/nph.16290>
- Lyu J., Ramekar R., Kim J., Hung N., Seo J., Kim J., Choi I., Park K., and Kwon S., 2021, Unraveling the complexity of faba bean (*Vicia faba* L.) transcriptome to reveal cold-stress-responsive genes using long-read isoform sequencing technology, *Scientific Reports*, 11(1): 21094.  
<https://doi.org/10.1038/s41598-021-00506-0>
- Mandel J., Dikow R., and Funk V., 2015, Using phylogenomics to resolve mega-families: an example from compositae, *Journal of Systematics and Evolution*, 53(5): 391-402.  
<https://doi.org/10.1111/jse.12167>

- McManus H., Fučíková K., Lewis P., Lewis L., and Karol K., 2018, Organellar phylogenomics inform systematics in the green algal family Hydrodictyaceae (Chlorophyceae) and provide clues to the complex evolutionary history of plastid genomes in the green algal tree of life, *American Journal of Botany*, 105(3): 315-329.  
<https://doi.org/10.1002/ajb2.1066>
- Rothfels C., Pryer K., and Li F., 2017, Next-generation polyploid phylogenetics: rapid resolution of hybrid polyploid complexes using PacBio single-molecule sequencing, *The New Phytologist*, 213(1): 413-429.  
<https://doi.org/10.1111/nph.14111>
- Uribe-Convers S., Settles M., and Tank D., 2016, A Phylogenomic approach based on PCR target enrichment and high throughput sequencing: resolving the diversity within the south american species of *Bartsia* L. (Orobanchaceae), *PLoS One*, 11(2): e0148203.  
<https://doi.org/10.1371/journal.pone.0148203>
- Vargas O., Ortiz E., and Simpson B., 2017, Conflicting phylogenomic signals reveal a pattern of reticulate evolution in a recent high-Andean diversification (Asteraceae: Astereae: *Diplostephium*), *The New Phytologist*, 214(4): 1736-1750.  
<https://doi.org/10.1111/nph.14530>
- Yang Y., Zhu J., Feng L., Zhou T., Bai G., Yang J., and Zhao G., 2018, Plastid genome comparative and phylogenetic analyses of the key genera in fagaceae: highlighting the effect of codon composition bias in phylogenetic inference, *Frontiers in Plant Science*, 9: 82.  
<https://doi.org/10.3389/fpls.2018.00082>
- Zhao Y., Zhang R., Jiang K., Qi J., Hu Y., Guo J., Zhu R., Zhang T., Egan A., Yi T., Huang C., and Ma H., 2021, Nuclear phylotranscriptomics/phylogenomics support numerous polyploidization events and hypotheses for the evolution of rhizobial nitrogen-fixing symbiosis in Fabaceae, *Molecular Plant*, 14(5): 748-773.  
<https://doi.org/10.1016/j.molp.2021.02.006>



#### **Disclaimer/Publisher's Note**

The statements, opinions, and data contained in all publications are solely those of the individual authors and contributors and do not represent the views of the publishing house and/or its editors. The publisher and/or its editors disclaim all responsibility for any harm or damage to persons or property that may result from the application of ideas, methods, instructions, or products discussed in the content. Publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

---