

Phylogenetic Relationships and Genetic Diversity among Domesticated Legumes

Xiaoxi Zhou¹, Shengyu Chen² ✉

¹ Institute of Life Sciences, Jiyang College, Zhejiang AandF University, Zhuji, 311800, Zhejiang, China

² Modern Agricultural Research Center, Cuixi Academy of Biotechnology, Zhuji, 311800, Zhejiang, China

✉ Corresponding email: shengyu.chen@cuixi.org

Legume Genomics and Genetics, 2024 Vol.15, No.4 doi: [10.5376/lgg.2024.15.0019](https://doi.org/10.5376/lgg.2024.15.0019)

Received: 08 Jul., 2024

Accepted: 09 Aug., 2024

Published: 20 Aug., 2024

Copyright © 2024 Zhou and Chen, 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:

Zhou X.X., and Chen S.Y., 2024, Phylogenetic relationships and genetic diversity among domesticated legumes, Legume Genomics and Genetics, 15(4): 187-198 (doi: [10.5376/lgg.2024.15.0019](https://doi.org/10.5376/lgg.2024.15.0019))

Abstract The study of domesticated legumes holds significant importance due to their role in global food security and sustainable agriculture. The study aims to elucidate the phylogenetic relationships and genetic diversity among major domesticated legume species. Initially, the study presents a historical perspective on legume domestication, drawing on archaeological and evolutionary evidence. Subsequently, the study explores phylogenetic relationships using molecular markers and computational approaches, focusing on species such as the common bean (*Phaseolus vulgaris*), chickpea (*Cicer arietinum*), lentil (*Lens culinaris*), and soybean (*Glycine max*). The genetic diversity within and between these species is examined, highlighting sources of variation and their implications for crop improvement and conservation. A detailed case study on the common bean underscores the practical applications of phylogenetic and genetic diversity insights. Additionally, the study discusses the molecular tools and techniques employed in these studies, including high-throughput sequencing and bioinformatics analysis. The environmental and agricultural implications of genetic diversity and phylogenetics are considered, emphasizing their impact on crop resilience and sustainable agricultural practices. Finally, the study outlines future research directions and challenges, advocating for an integrative approach that combines traditional knowledge with modern scientific techniques. This comprehensive study underscores the critical role of understanding phylogenetic relationships and genetic diversity in advancing legume crop productivity and sustainability.

Keywords Domesticated legumes; Phylogenetics; Genetic diversity; Crop improvement; Sustainable agriculture

1 Introduction

Legumes, belonging to the Fabaceae family, are of paramount importance both economically and nutritionally. They represent the second most significant family of crop plants after the Poaceae family, contributing substantially to global food security and agricultural sustainability. Grain legumes alone account for 27% of the world's crop production and provide 33% of the dietary protein consumed by humans. Additionally, pasture and forage legumes are vital for animal feed, underscoring their multifaceted role in agriculture (Smýkal et al., 2015). The domestication of legumes dates back to ancient civilizations, with crops such as soybean, chickpea, lentil, and common bean being among the earliest cultivated plants. These crops have been integral to human diets and agricultural systems across various regions, including China, the Fertile Crescent, Africa, and the Americas (Raina et al., 2019).

Understanding the phylogenetic relationships and genetic diversity among domesticated legumes is crucial for several reasons. Phylogenetics provides insights into the evolutionary history and relationships among different legume species, which is essential for tracing their domestication pathways and identifying genetic traits that have been selected over time (Jansen et al., 2008; Schmutz et al., 2014). Genetic diversity, on the other hand, is a key factor in crop improvement programs. It allows for the identification of valuable traits such as disease resistance, drought tolerance, and nutritional quality, which can be harnessed to develop superior legume varieties through breeding programs (Varshney et al., 2013; Gujaria-Verma et al., 2014; Varshney, 2016). Advances in genomic technologies have facilitated the sequencing of complete genomes of several legume species, providing a wealth of data that can be used to explore genetic diversity and phylogenetic relationships in greater detail (Kreplak et al., 2019).

This study aims to combine phylogenetic and genetic diversity approaches to better illustrate the origin, domestication history, and preserved germplasm of major legume crops. Specifically, the objectives are to elucidate the phylogenetic relationships among domesticated legumes using comprehensive genomic data, to assess the genetic diversity within and between different legume species, and to identify key genetic traits that have been selected during domestication. Additionally, the study seeks to provide insights into the potential for future crop improvement through the integration of phylogenetic and genetic diversity data, thereby enhancing the resilience and nutritional quality of legume crops. By achieving these objectives, this study will contribute to a deeper understanding of the evolutionary and genetic foundations of domesticated legumes, paving the way for more effective and sustainable agricultural practices.

2 Historical Perspective on Legume Domestication

2.1 Early domestication events

The domestication of legumes began around 12 000 years ago, coinciding with the advent of Neolithic agriculture. This period marked the transformation of wild species into cultivated crops through human selection for desirable traits such as taste, yield, and ease of cultivation (Ambika et al., 2022; Bohra et al., 2022). Legumes were among the earliest domesticated plants, paralleling the domestication of cereals. Key legume crops such as soybean, lentil, chickpea, and pea were first cultivated in regions like the Fertile Crescent, China, and the Indian subcontinent (Smýkal et al., 2015). The domestication process involved significant genetic changes, including the loss of seed dormancy and shattering resistance, which were crucial for the development of stable and high-yielding crops (Rendón-Anaya et al., 2017).

2.2 Archaeological and historical evidence of legume cultivation

Archaeological evidence provides critical insights into the history of legume domestication. For instance, lentils were first cultivated in Southwest Asia around 8 000~10 000 years ago, although the exact number of independent domestication events remains unclear (Liber et al., 2021). Similarly, the domestication of common beans in Mesoamerica and the Andes has been traced back through genetic and archaeological studies, revealing distinct domestication events in these regions (Nanni et al., 2011; Rendón-Anaya et al., 2017). The Indus Valley civilization also played a significant role in the domestication of various legumes, including pigeon pea and black gram (Figure 1) (Ambika et al., 2022). These findings are supported by radiocarbon dating of plant remains and the geographical distribution of wild relatives, which help reconstruct the domestication timeline and migration patterns of these crops (Fuller, 2007).

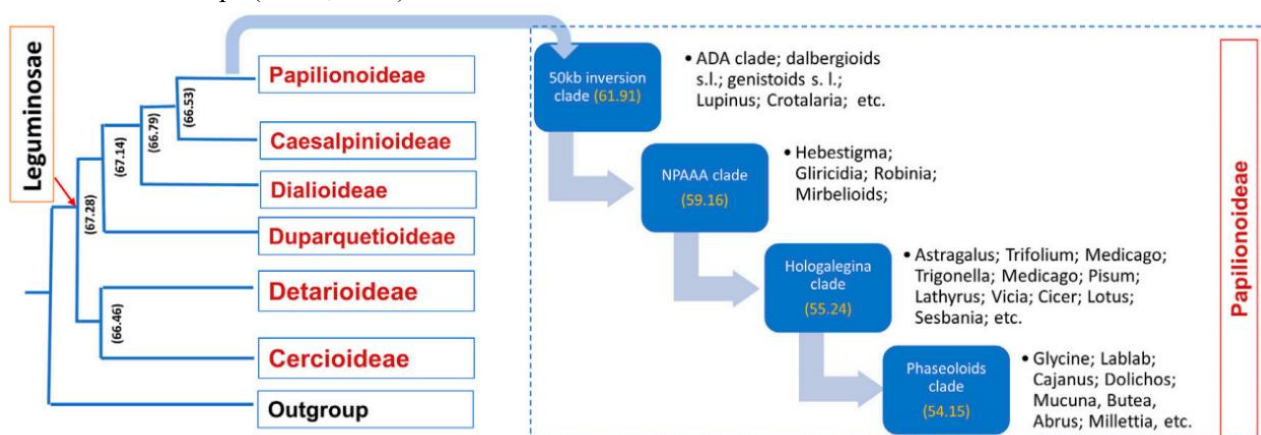


Figure 1 Time divergence of major legume clades (Adopted from Ambika et al., 2022)

Image caption: Values inside the brackets are the estimated time of divergence based on phylogenomic studies (Adopted from Ambika et al., 2022)

2.3 Evolutionary implications of domestication

The domestication of legumes has had profound evolutionary implications. The process led to a severe genetic bottleneck, resulting in a significant loss of genetic diversity in domesticated varieties compared to their wild ancestors (Smýkal et al., 2015; Ambika et al., 2022). This reduction in genetic variability has implications for crop

resilience and adaptability. However, the identification of domestication-responsive loci and the use of modern genomic tools, such as whole genome sequencing and gene-editing technologies, offer opportunities to reintroduce genetic diversity and improve crop traits (Bohra et al., 2022). Additionally, the study of phylogenetic relationships among legume species has revealed complex evolutionary histories, including instances of polyploidy and hybridization, which have contributed to the diversification and success of legume crops (Delêtre et al., 2017; Koenen et al., 2020). Understanding these evolutionary processes is crucial for developing strategies to enhance the genetic diversity and resilience of modern legume cultivars.

3 Phylogenetic Relationships among Domesticated Legumes

3.1 Methods for constructing phylogenetic trees

Molecular markers and DNA sequencing are fundamental tools in constructing phylogenetic trees. Various types of molecular markers, such as microsatellites (SSR) and single nucleotide polymorphisms (SNPs), are used to assess genetic diversity and relationships among legume species. For instance, microsatellite markers have been effectively used to evaluate the genetic diversity of *Vigna stipulacea*, revealing significant variation and clustering accessions into distinct clades (Gore et al., 2022). Additionally, sequencing of specific genes, such as the plastid *matK* gene, has provided high-resolution phylogenetic trees for legumes, supporting well-resolved subclades within the family (Wojciechowski et al., 2004).

Computational approaches play a crucial role in analyzing large-scale genomic data to infer phylogenetic relationships. Methods such as maximum likelihood, Bayesian inference, and multispecies coalescent models are commonly employed. For example, a study utilizing 72 chloroplast genes and 7 621 nuclear-encoded proteins across 157 taxa employed these computational methods to resolve deep divergences in the legume phylogeny, despite challenges like incomplete lineage sorting (Figure 2) (Koenen et al., 2019). These approaches help in constructing robust phylogenetic frameworks and understanding evolutionary processes.

3.2 Phylogenetic analysis of major domesticated legume species

The common bean has been extensively studied for its phylogenetic relationships and genetic diversity. Genotyping-by-sequencing (GBS) has revealed the presence of two major gene pools, Mesoamerican and Andean, with evidence of hybridization events between them (Lioi et al., 2019). Additionally, the assembly of the common bean genome and comparison with the soybean genome have provided insights into the genetic changes associated with domestication (Schmutz et al., 2014). Chickpea phylogenetic studies often focus on its genetic diversity and domestication history. Molecular markers and sequencing data have been used to trace the evolutionary relationships and divergence times within the species, contributing to a better understanding of its domestication and adaptation processes (Smýkal et al., 2015). Lentil phylogenetic analysis involves the use of molecular markers to assess genetic diversity and relationships among different accessions. These studies help in identifying key traits and understanding the evolutionary history of lentils, which is crucial for breeding and conservation efforts. Soybean phylogenetic studies have utilized both molecular markers and whole-genome sequencing. Comparative genomic analyses between soybean and other legumes, such as *Medicago truncatula*, have highlighted the conservation and divergence of genomic regions, providing insights into the evolutionary history and domestication of soybean (Choi et al., 2004; Pfeil et al., 2005).

3.3 Phylogenetic insights into legume evolution

Phylogenetic studies have provided significant insights into the divergence times and evolutionary relationships among legume species. For example, the simultaneous origin of all six subfamilies of legumes suggests a rapid speciation event, challenging the traditional view of basal and early-diverging subfamilies (Koenen et al., 2019). Additionally, the reconstruction of ancestral genomes has revealed the chromosomal evolution history of legumes, indicating a common ancestor with nine chromosomes (Ren et al., 2019).

The phylogenetic insights gained from these studies have important implications for the classification and taxonomy of legumes. For instance, the reestablishment of the tribe Diocleae within the Millettioideae clade based on molecular phylogenetic analysis has clarified the higher-level phylogeny and generic relationships within this

group (Queiroz et al., 2015). These findings contribute to a more accurate and robust classification system for legumes, aiding in the understanding of their evolutionary history and diversity.

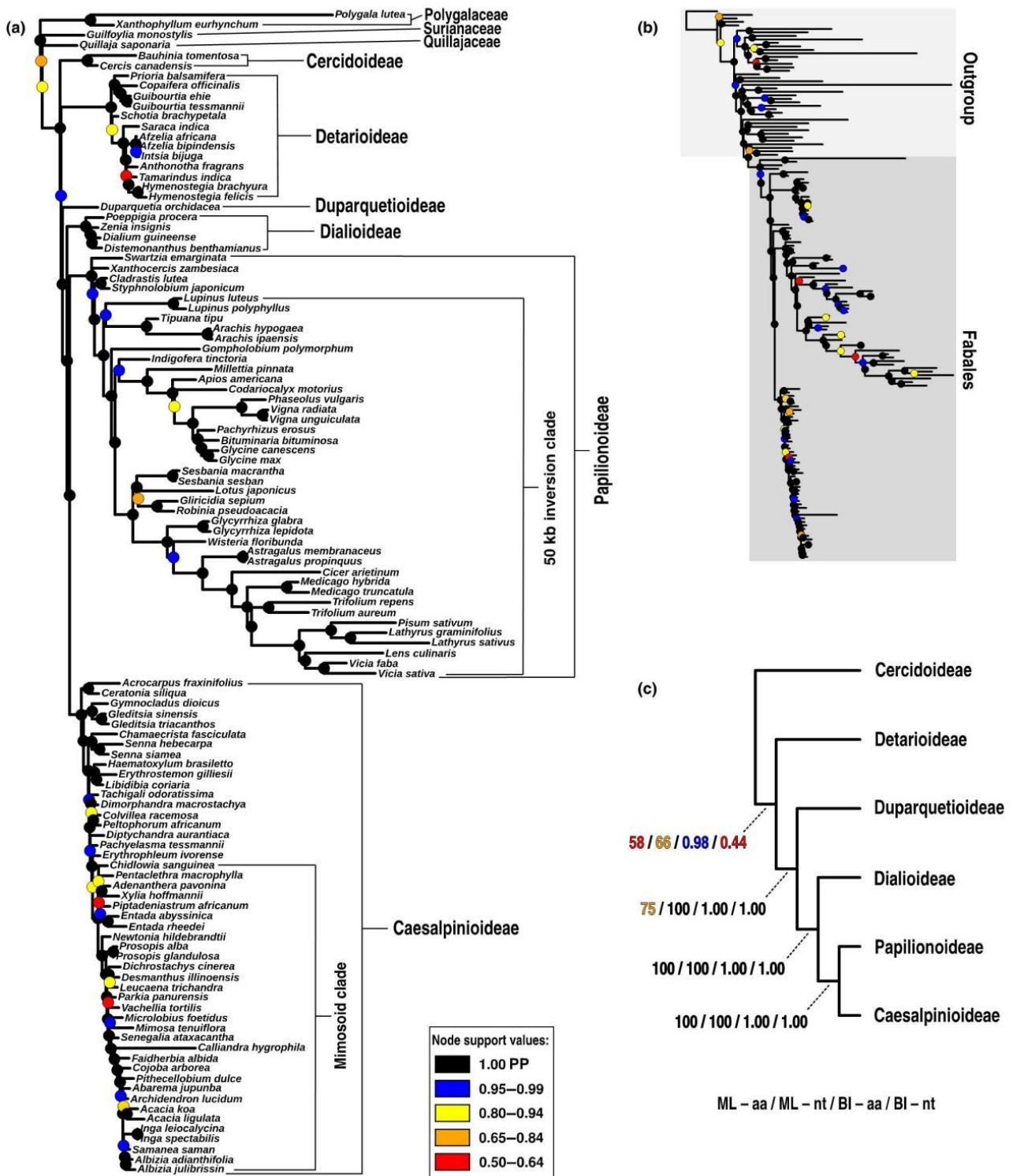


Figure 2 Phylogeny of legumes based on Bayesian analyses of 72 protein-coding chloroplast genes under the CATGTR model in PHYLOBAYES (Adopted from Koenen et al., 2019)

Image caption: (a) Majority-rule consensus tree of the amino acid alignment, showing only the Fabales portion of the tree, outgroup taxa pruned; (b) Complete tree including outgroup taxa; (c) Simplified tree showing support for subfamily relationships with different inference methods (ML, maximum likelihood; BI, Bayesian inference) and sequence types (aa, amino acids; nt, nucleotides). In (a) and (b), coloured circles indicate node support in posterior probabilities (PP) (Adopted from Koenen et al., 2019)

4 Genetic Diversity in Domesticated Legumes

4.1 Genetic variation within and between species

Genetic diversity in domesticated legumes arises from various sources, including wild relatives, landraces, and exotic germplasm. Wild relatives, in particular, harbor a wealth of genetic variation that has been shaped by natural selection in diverse environments. This genetic reservoir is crucial for introducing novel traits into domesticated species (Coyne et al., 2020; Pratap et al., 2021). Additionally, the use of molecular markers such as microsatellites has revealed significant allelic diversity within legume species, indicating a relatively wide genetic base (Smýkal et al., 2015).

Population genetics studies have been instrumental in assessing the genetic diversity within and between legume species. For instance, the use of SSR markers in *Vigna stipulacea* has shown high genetic diversity among accessions, which is essential for breeding programs. Similarly, genome-wide analyses of common bean have identified distinct genetic pools resulting from independent domestication events, highlighting the genetic differentiation within the species (Schmutz et al., 2014).

4.2 Role of genetic diversity in crop improvement

Breeding strategies that leverage genetic diversity are vital for crop improvement. Introgression breeding, which involves the incorporation of genes from wild relatives into domesticated species, has been successful in developing improved cultivars with enhanced traits such as disease resistance and stress tolerance (Figure 3) (Pratap et al., 2021). The use of pre-breeding programs to develop new gene pools with a broad genetic base is also crucial for the genetic enhancement of grain legumes (Sharma et al., 2013).

Conserving genetic diversity is essential for sustainable crop improvement. Ex situ and in situ conservation methods are employed to preserve the genetic resources of wild relatives and landraces. These conserved genetic resources serve as a reservoir for future breeding efforts, ensuring the availability of diverse alleles for crop improvement (Coyne et al., 2020; Ambika et al., 2022).

4.3 Comparative analysis of genetic diversity in wild and domesticated legumes

Domestication bottlenecks have led to a significant reduction in genetic diversity in domesticated legumes compared to their wild counterparts. This loss of diversity is attributed to the selection of specific traits during domestication, which narrows the genetic base of cultivated species (Ambika et al., 2022). For example, the domestication of common bean involved two independent events, each resulting in a distinct genetic pool with limited shared genetic material (Schmutz et al., 2014).

Gene flow between wild and cultivated populations can introduce new genetic variation into domesticated species, enhancing their adaptability and resilience. Hybrid populations created through the introgression of wild alleles into cultivated backgrounds provide valuable genetic resources for breeding programs (Coyne et al., 2020; Pratap et al., 2021). This gene flow is facilitated by the use of modern genomics and systematic phenotyping, which help identify and harness beneficial alleles from wild relatives. In summary, understanding and leveraging genetic diversity in domesticated legumes is crucial for crop improvement. By utilizing genetic resources from wild relatives and employing advanced breeding strategies, it is possible to develop cultivars with enhanced traits and ensure the sustainability of legume crops (Kreplak et al., 2019; Gore et al., 2022).

5 Case Study: Genetic Diversity and Phylogenetics of Common Bean (*Phaseolus vulgaris*)

5.1 Importance of common bean as a global food source

The common bean (*Phaseolus vulgaris* L.) is a crucial legume crop globally, serving as a fundamental source of protein, particularly in rural societies. It is the most important grain legume for human consumption and plays a significant role in sustainable agriculture due to its ability to fix atmospheric nitrogen (Scarano et al., 2014; Schmutz et al., 2014; Assefa et al., 2019). In regions like Southern Italy and Brazil, common bean cultivation is integral to local economies and diets, highlighting its importance as a staple food (Burle et al., 2010).

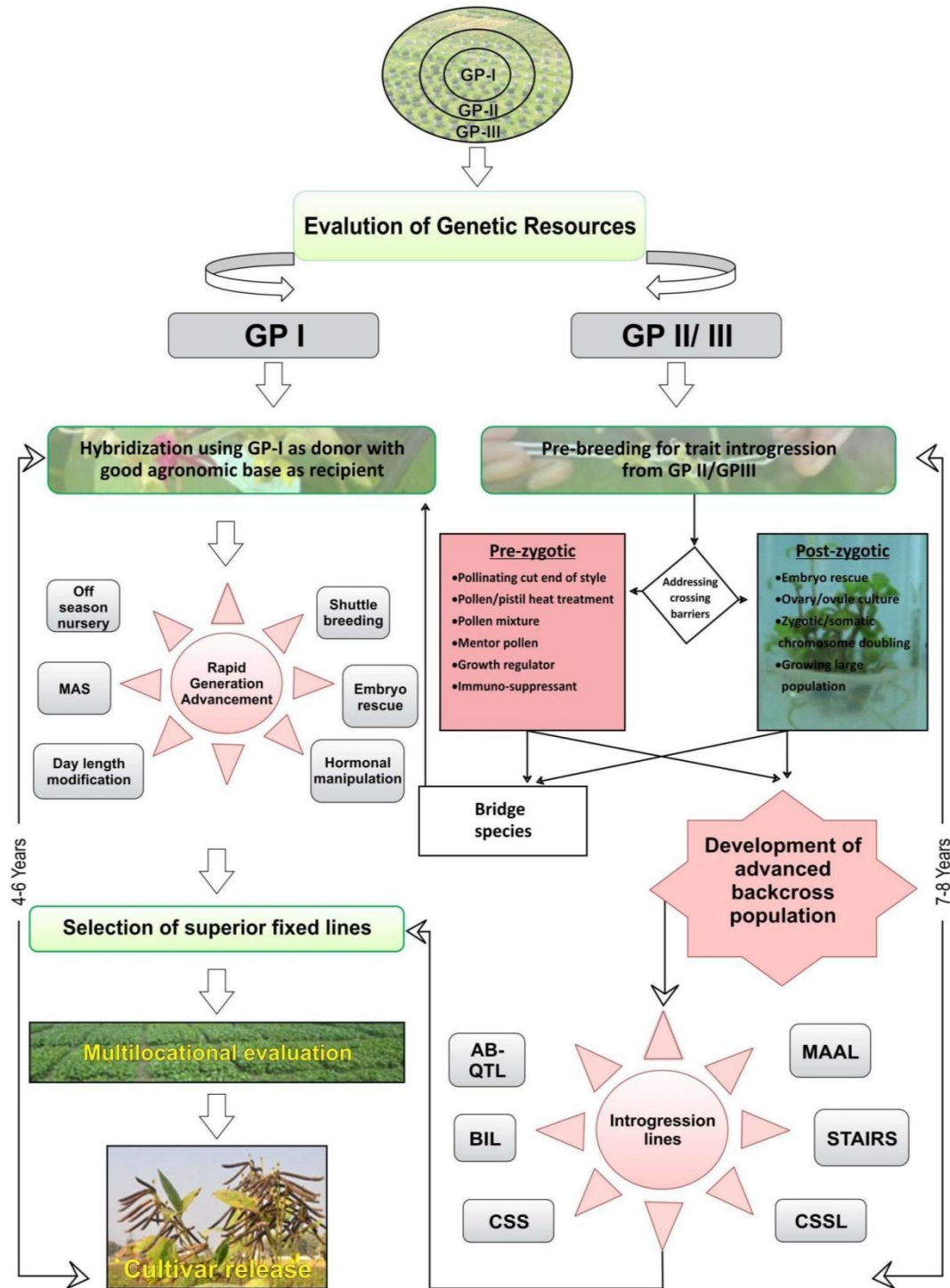


Figure 3 Scheme for gene introgression for improvement of food legumes (Adopted from Pratap et al., 2021)

Image caption: Left, Hybridization using GP-I as donor with good agronomic base as recipient. Right, Pre-breeding for trait introgression using GP-II and/or GP-III (Adopted from Pratap et al., 2021)

5.2 Phylogenetic relationships within the *Phaseolus* genus

Phylogenetic studies have revealed that the common bean has undergone two independent domestication events in geographically distinct areas: Mesoamerica and the Andes. These events have led to the formation of two major gene pools, each with unique genetic characteristics (Mamidi et al., 2011; Bitocchi et al., 2013; Schmutz et al., 2014). The genetic differentiation between these gene pools is significant, with less than 10% of the

domestication-related genetic sequences being shared between them. This dual domestication has provided a rich framework for studying the evolutionary history and genetic diversity of the species.

5.3 Analysis of genetic diversity in domesticated and wild common bean

Genetic diversity in common bean has been extensively studied using various molecular markers, including SSRs and SNPs. Studies have shown a reduction in genetic diversity from wild to domesticated populations, with a more pronounced bottleneck effect observed in the Mesoamerican gene pool compared to the Andean gene pool (Mamidi et al., 2011; Bitocchi et al., 2013; Gioia et al., 2019). In regions like Southern Italy and Brazil, significant phenotypic and genetic diversity has been observed among local landraces, indicating the presence of a wide-ranging variation that is crucial for breeding programs (Burle et al., 2010; Scarano et al., 2014). Additionally, the use of SNP genotyping platforms has revealed genetic similarities and differences among germplasm entries, aiding in the identification of duplicated lines and ensuring germplasm purity (Raatz et al., 2019).

5.4 Implications for breeding and conservation

The genetic diversity observed in common bean landraces and wild populations has important implications for breeding and conservation. The presence of significant intra-varietal differences and the identification of genes linked with desirable traits such as increased leaf and seed size provide valuable resources for genomics-enabled crop improvement (Scarano et al., 2014; Schmutz et al., 2014; Gioia et al., 2019). Conservation efforts, such as on-farm management and the establishment of core collections, are essential to preserve this genetic variability and ensure the sustainability of common bean cultivation (Rivera et al., 2018). Moreover, the development of marker-assisted breeding methods and the use of genomic resources can enhance the efficiency and effectiveness of breeding programs, addressing both biotic and abiotic stresses (Assefa et al., 2019; Raatz et al., 2019).

6 Molecular Tools and Techniques for Studying Phylogenetics and Genetic Diversity

6.1 High-throughput sequencing technologies

Next-generation sequencing (NGS) has revolutionized the field of legume genomics by enabling the rapid and cost-effective sequencing of entire genomes and transcriptomes. This technology has been instrumental in uncovering the genetic basis of important traits and understanding the evolutionary history of legumes. For instance, NGS has facilitated the analysis of the soybean transcriptome, revealing insights into seed development and nutrient utilization (O'Rourke et al., 2014). Additionally, NGS has been used to generate large-scale genomic datasets that resolve deep phylogenetic relationships within the legume family, despite challenges such as incomplete lineage sorting (Koenen et al., 2019). The integration of NGS data from different germplasm collections has also provided valuable information on the genetic diversity and evolutionary history of crops like peas (Pavan et al., 2022).

Whole-genome sequencing (WGS) provides a comprehensive view of the genetic makeup of an organism, allowing for detailed studies of genome structure, function, and evolution. In legumes, WGS has been used to assemble high-quality reference genomes for species such as common bean and soybean, which are essential for comparative genomics and breeding programs (Schmutz et al., 2014). The availability of these reference genomes has enabled researchers to identify genomic regions associated with domestication and important agronomic traits, thereby facilitating crop improvement efforts (Kumar et al., 2014). Furthermore, WGS has revealed the extent of genome conservation and structural variation among different legume species, highlighting the evolutionary processes that have shaped their genomes (Choi et al., 2004).

6.2 Bioinformatics tools for data analysis

Phylogenetic tree construction is a critical step in understanding the evolutionary relationships among legume species. Software tools such as maximum likelihood and Bayesian inference methods have been employed to analyze large genomic datasets and construct robust phylogenetic trees. These tools have been used to resolve the deepest divergences in the legume phylogeny and to test alternative evolutionary hypotheses (Koenen et al., 2019). Additionally, multispecies coalescent methods have been applied to account for gene tree discordance and incomplete lineage sorting, providing a more accurate picture of legume evolution (Crameri et al., 2021).

Population genetics tools are essential for analyzing genetic diversity and structure within and between legume populations. Techniques such as genotyping-by-sequencing (GBS) and single nucleotide polymorphism (SNP) analysis have been used to study the genetic variation in crops like peas and to infer their evolutionary history (Pavan et al., 2022). These tools have also been employed to identify selective sweeps and genomic regions associated with domestication and breeding, providing insights into the genetic basis of important traits (Egan and Vatanparast, 2019). Moreover, population genomic analyses have revealed the existence of cryptic species and highlighted the need for conservation efforts in economically significant legume species.

6.3 Integrative approaches combining phylogenetics and genomics

Integrative approaches that combine phylogenetic and genomic data are increasingly being used to study the evolution and genetic diversity of legumes. By integrating data from whole-genome sequencing, transcriptomics, and population genetics, researchers can gain a comprehensive understanding of the genetic basis of important traits and the evolutionary processes that have shaped legume genomes. For example, the integration of genomic and phylogenetic data has provided insights into the role of whole-genome duplication events in legume evolution and the genetic changes associated with domestication (Kumar et al., 2014). Additionally, these approaches have been used to study the genetic diversity and phylogenetic relationships of economically important legume species, informing breeding and conservation strategies (Smýkal et al., 2015).

7 Environmental and Agricultural Implications

7.1 Impact of genetic diversity on crop resilience

Genetic diversity plays a crucial role in enhancing the resilience of legume crops to various environmental stresses. For instance, the ability of legumes to fix atmospheric nitrogen not only contributes to soil fertility but also helps them adapt to nutrient-poor conditions (Schmutz et al., 2014). Additionally, the genetic diversity found in wild relatives of legumes offers a reservoir of traits that can be harnessed to improve stress tolerance in domesticated varieties. This includes resistance to drought, salinity, and extreme temperatures, which are becoming increasingly important due to climate change (Abdelrahman et al., 2018; Coyne et al., 2020; Yung et al., 2022). Epigenetic mechanisms, such as DNA methylation and histone modification, also play a significant role in regulating gene expression in response to abiotic stresses, further enhancing the adaptability of legumes.

The genetic diversity within legume species is also vital for developing resistance to pests and diseases. Modern breeding programs often face challenges due to the narrow genetic base of cultivated varieties, making them more susceptible to biotic stresses. However, the incorporation of genetic material from wild relatives can introduce new resistance genes, thereby enhancing the overall resilience of the crop (Coyne et al., 2020). For example, studies have shown that specific genes linked to disease resistance can be identified and utilized in breeding programs to develop more robust legume varieties (Schmutz et al., 2014; Smýkal et al., 2015).

7.2 Role of phylogenetics in sustainable agriculture

Phylogenetic studies provide valuable insights into the evolutionary relationships among legume species, which can be leveraged to design effective crop rotation and diversification strategies. By understanding the genetic and evolutionary backgrounds of different legume species, farmers can select complementary crops that enhance soil health and reduce pest and disease pressures (Smýkal et al., 2015). This approach not only improves yield stability but also contributes to sustainable agricultural practices by minimizing the need for chemical inputs (Duc et al., 2015).

Legumes play a significant role in enhancing ecosystem services, such as soil fertility and biodiversity. Their ability to fix atmospheric nitrogen reduces the need for synthetic fertilizers, thereby lowering greenhouse gas emissions and improving soil health (Schmutz et al., 2014). Additionally, the presence of diverse legume species in agricultural systems can support a wide range of beneficial soil microorganisms, which further contribute to soil fertility and plant health (Prudent et al., 2019). Phylogenetic studies can help identify legume species that are particularly effective in providing these ecosystem services, thereby guiding the development of more sustainable agricultural systems (Smýkal et al., 2015).

7.3 Future directions and challenges

One of the major challenges facing legume cultivation is the impact of climate change. Future research should focus on identifying and breeding legume varieties that are resilient to the changing climate. This includes developing cultivars that can withstand extreme weather conditions, such as drought and heat, as well as those that can adapt to new pest and disease pressures (Araújo et al., 2015; Thudi et al., 2020). Advances in genomic technologies, such as next-generation sequencing and marker-assisted selection, offer promising tools for accelerating the development of climate-resilient legume varieties.

Integrating traditional agricultural knowledge with modern scientific approaches can provide a holistic strategy for improving legume cultivation. Traditional knowledge often includes valuable insights into local crop varieties and farming practices that have been optimized over generations. Combining this knowledge with modern genomic and breeding techniques can lead to the development of legume varieties that are not only high-yielding and resilient but also culturally and regionally appropriate (Coyne et al., 2020; Bohra et al., 2022). This integrative approach can help address the complex challenges of modern agriculture, ensuring food security and sustainability for future generations.

8 Concluding Remarks

This study has explored the phylogenetic relationships and genetic diversity among domesticated legumes, revealing several critical insights. Legumes, belonging to the Fabaceae family, are economically significant, providing essential dietary proteins and animal feed. The phylogenetic analysis of various legume species, such as *Pisum* (peas), *Lens* (lentils), and *Phaseolus* (common beans), has highlighted their evolutionary trajectories and genetic diversity. The research underscores the importance of wild relatives as reservoirs of genetic variation, which can be harnessed for crop improvement. Additionally, the study has shown that the domestication of legumes involved multiple independent events and significant genetic bottlenecks, which have shaped the current genetic landscape of these crops.

The findings of this study have several implications for future research and agricultural practices. Firstly, the genetic diversity present in wild legume species should be further explored and conserved, as it holds potential for breeding programs aimed at improving yield, quality, and stress tolerance in domesticated varieties. Secondly, the use of advanced genomic tools and techniques, such as whole genome sequencing and gene editing, can accelerate the domestication and improvement of legume crops without significant loss of genetic diversity. Additionally, understanding the phylogenetic relationships among legume species can inform strategies for the conservation of genetic resources and the development of new cultivars with desirable traits.

Understanding the phylogenetic relationships and genetic diversity in domesticated legumes is crucial for several reasons. It provides insights into the evolutionary history and domestication processes of these vital crops, which can inform breeding strategies and conservation efforts. Moreover, the genetic diversity within and among legume species is a valuable resource for developing resilient and high-yielding cultivars, essential for ensuring food security and sustainable agriculture. As legumes play a significant role in human and animal nutrition, as well as in maintaining soil fertility through nitrogen fixation, preserving their genetic diversity and understanding their evolutionary relationships is of paramount importance for future agricultural advancements.

Acknowledgments

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

Conflict of Interest Disclosure

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

References

- Abdelrahman M., Jogaiah S., Burritt D., and Tran L., 2018, Legume genetic resources and transcriptome dynamics under abiotic stress conditions, *Plant, Cell and Environment*, 41(9): 1972-1983.
<https://doi.org/10.1111/pce.13123>

- Ambika, Aski M.S., Gayacharan, Hamwiah A., Talukdar A., Kumar Gupta S., Sharma B.B., Joshi R., Upadhyaya H.D., Singh K., and Kumar R., 2022, Unraveling origin, history, genetics, and strategies for accelerated domestication and diversification of food legumes, *Frontiers in Genetics*, 13: 932430.
<https://doi.org/10.3389/fgene.2022.932430>
- Araújo S., Beebe S., Crespi M., Delbreil B., Gonzalez E., Gruber V., Lejeune-Hénaut I., Link W., Monteros M., Prats E., Rao I., Vadez V., and Patto M., 2015, Abiotic stress responses in legumes: strategies used to cope with environmental challenges, *Critical Reviews in Plant Sciences*, 34: 237-280.
<https://doi.org/10.1080/07352689.2014.898450>
- Assefa T., Mahama A., Brown A., Cannon E., Rubyogo J., Rao I., Rao I., Blair M., and Cannon S., 2019, A review of breeding objectives, genomic resources, and marker-assisted methods in common bean (*Phaseolus vulgaris* L.), *Molecular Breeding*, 39: 1-23.
<https://doi.org/10.1007/s11032-018-0920-0>
- Bitocchi E., Bellucci E., Giardini A., Rau D., Rodriguez M., Biagetti E., Santilocchi R., Zeuli P., Gioia T., Logozzo G., Attene G., Nanni L., and Papa R., 2013, Molecular analysis of the parallel domestication of the common bean (*Phaseolus vulgaris*) in Mesoamerica and the Andes, *The New Phytologist*, 197(1): 300-313.
<https://doi.org/10.1111/j.1469-8137.2012.04377.x>
- Bohra A., Tiwari A., Kaur P., Ganie S., Raza A., Roorkiwal M., Mir R., Fernie A., Smýkal P., and Varshney R., 2022, The key to the future lies in the past: insights from grain legume domestication and improvement should inform future breeding strategies, *Plant and Cell Physiology*, 63: 1554-1572.
<https://doi.org/10.1093/pcp/pcac086>
- Burle M., Fonseca J., Kami J., and Gepts P., 2010, Microsatellite diversity and genetic structure among common bean (*Phaseolus vulgaris* L.) landraces in Brazil, a secondary center of diversity, *Theoretical and Applied Genetics*, 121: 801-813.
<https://doi.org/10.1007/s00122-010-1350-5>
- Choi H., Mun J., Kim D., Zhu H., Baek J., Mudge J., Roe B., Ellis N., Doyle J., Kiss G., Young N., and Cook D., 2004, Estimating genome conservation between crop and model legume species, *Proceedings of the National Academy of Sciences of the United States of America*, 101(43): 15289-15294.
<https://doi.org/10.1073/PNAS.0402251101>
- Coyne C., Kumar S., Wettberg E., Marques E., Berger J., Redden R., Ellis T., Brus J., Zablatzká L., and Smýkal P., 2020, Potential and limits of exploitation of crop wild relatives for pea, lentil, and chickpea improvement, *Legume Science*, 2(2): e36.
<https://doi.org/10.1002/leg3.36>
- Cramer S., Fior S., Zoller S., and Widmer A., 2021, A target capture approach for phylogenomic analyses at multiple evolutionary timescales in rosewoods (*Dalbergia* spp.) and the legume family (Fabaceae), *Molecular Ecology Resources*, 22: 3087-3105.
<https://doi.org/10.1111/1755-0998.13666>
- Delêtre M., Soengas B., Vidaurre P., Meneses R., Vásquez O., Balbín I., Santayana M., Heider B., and Sørensen M., 2017, Ecotypic differentiation under farmers' selection: molecular insights into the domestication of *Pachyrhizus* Rich. ex DC. (Fabaceae) in the Peruvian Andes, *Evolutionary Applications*, 10: 498-513.
<https://doi.org/10.1111/eva.12472>
- Duc G., Agrama H., Bao S., Berger J., Bourion V., Ron A., Gowda C., Mikić A., Millot D., Singh K., Tullu A., Vandenberg A., Patto M., Warkentin T., and Zong X., 2015, Breeding annual grain legumes for sustainable agriculture: new methods to approach complex traits and target new cultivar ideotypes, *Critical Reviews in Plant Sciences*, 34: 381-411.
<https://doi.org/10.1080/07352689.2014.898469>
- Egan A., and Vatanparast M., 2019, Advances in legume research in the genomics era, *Australian Systematic Botany*, 32: 459-483.
<https://doi.org/10.1071/SB19019>
- Fuller D., 2007, Contrasting patterns in crop domestication and domestication rates: recent archaeobotanical insights from the old world, *Annals of Botany*, 100: 903-924.
<https://doi.org/10.1093/aob/mcm048>
- Gioia T., Logozzo G., Marzario S., Zeuli P., and Gepts P., 2019, Evolution of SSR diversity from wild types to U.S. advanced cultivars in the Andean and Mesoamerican domestications of common bean (*Phaseolus vulgaris*), *PLoS One*, 14(1): e0211342.
<https://doi.org/10.1371/journal.pone.0211342>
- Gore P., Gupta V., Singh R., Tripathi K., Kumar R., Kumari G., Madhavan L., Dikshit H., Venkateswaran K., Pandey A., Singh N., Bhat K., Nair R., and Pratap A., 2022, Insights into the genetic diversity of an underutilized Indian legume, *Vigna stipulacea* (Lam.) Kuntz., using morphological traits and microsatellite markers, *PLoS One*, 17(1): e0262634.
<https://doi.org/10.1371/journal.pone.0262634>
- Gujaria-Verma N., Vail S., Carrasquilla-García N., Penmetsa R., Cook D., Farmer A., Vandenberg A., Bett K., Gepts P., and Akhunov E., 2014, Genetic mapping of legume orthologs reveals high conservation of synteny between lentil species and the sequenced genomes of *Medicago* and chickpea, *Frontiers in Plant Science*, 5: 676.
<https://doi.org/10.3389/fpls.2014.00676>
- Jansen R., Wojciechowski M., Sanniyasi E., Lee S., and Daniell H., 2008, Complete plastid genome sequence of the chickpea (*Cicer arietinum*) and the phylogenetic distribution of *rps12* and *clpP* intron losses among legumes (Leguminosae), *Molecular Phylogenetics and Evolution*, 48(3): 1204-1217.
<https://doi.org/10.1016/j.ympev.2008.06.013>

- Koenen E., Ojeda D., Bakker F., Wieringa J., Kidner C., Hardy O., Pennington R., Herendeen P., Bruneau A., and Hughes C., 2020, The origin of the legumes is a complex paleopolyploid phylogenomic tangle closely associated with the cretaceous-paleogene (K-Pg) mass extinction event, *Systematic Biology*, 70: 508-526.
<https://doi.org/10.1093/sysbio/syaa041>
- 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>
- Kreplak J., Madoui M., Cápál P., Novák P., Labadie K., Aubert G., Bayer P., Gali K., Syme R., Main D., Klein A., Bérard A., Vrbová I., Fournier C., d'Agata L., Belsler C., Berrabah W., Toegelová H., Milec Z., Vrána J., Lee H., Kougbeadjo A., Térézol M., Huneau C., Turo C., Mohellibi N., Neumann P., Falque M., Gallardo K., McGee R., Tar'an B., Bendahmane A., Aury J., Batley J., Paslier M., Ellis N., Warkentin T., Coyne C., Salse J., Edwards D., Lichtenzweig J., Macas J., Doležel J., Wincker P., and Burstin J., 2019, A reference genome for pea provides insight into legume genome evolution, *Nature Genetics*, 51: 1411-1422.
<https://doi.org/10.1038/s41588-019-0480-1>
- Kumar J., Srivastava E., Singh M., and Pratap A., 2014, Genomics in studying the legume genome evolution, In: Gupta S., Nadarajan, N., and Gupta D. (eds), *Legumes in the omic era*, Springer, New York, USA, pp.287-300.
https://doi.org/10.1007/978-1-4614-8370-0_14
- Liber M., Duarte I., Maia A., and Oliveira H., 2021, The history of lentil (*Lens culinaris* subsp. *culinaris*) domestication and spread as revealed by genotyping-by-sequencing of wild and landrace accessions, *Frontiers in Plant Science*, 12: 628439.
<https://doi.org/10.3389/fpls.2021.628439>
- Lioi L., Zuluaga D., Pavan S., and Sonnante G., 2019, Genotyping-by-sequencing reveals molecular genetic diversity in Italian common bean landraces, *Diversity*, 11(9): 154.
<https://doi.org/10.3390/d11090154>
- Mamidi S., Rossi M., Annam D., Moghaddam S., Lee R., Papa R., and McClean P., 2011, Investigation of the domestication of common bean (*Phaseolus vulgaris*) using multilocus sequence data, *Functional Plant Biology*, 38(12): 953-967.
<https://doi.org/10.1071/FP11124>
- Nanni L., Bitocchi E., Bellucci E., Rossi M., Rau D., Attene G., Gepts P., and Papa R., 2011, Nucleotide diversity of a genomic sequence similar to *SHATTERPROOF (PvSHP1)* in domesticated and wild common bean (*Phaseolus vulgaris* L.), *Theoretical and Applied Genetics*, 123: 1341-1357.
<https://doi.org/10.1007/s00122-011-1671-z>
- O'Rourke J., Bolon Y., Bucciarelli B., and Vance C., 2014, Legume genomics: understanding biology through DNA and RNA sequencing, *Annals of Botany*, 113(7): 1107-1120.
<https://doi.org/10.1093/aob/mcu072>
- Pavan S., Delvento C., Nazzicari N., Ferrari B., D'Agostino N., Taranto F., Lotti C., Ricciardi L., and Annicchiarico P., 2022, Merging genotyping-by-sequencing data from two ex situ collections provides insights on the pea evolutionary history, *Horticulture Research*, 9: uhab062.
<https://doi.org/10.1093/hr/uhab062>
- Pfeil B., Schlueter J., Shoemaker R., and Doyle J., 2005, Placing paleopolyploidy in relation to taxon divergence: a phylogenetic analysis in legumes using 39 gene families, *Systematic Biology*, 54(3): 441-454.
<https://doi.org/10.1080/10635150590945359>
- Pratap A., Das A., Kumar S., and Gupta S., 2021, Current perspectives on introgression breeding in food legumes, *Frontiers in Plant Science*, 11: 589189.
<https://doi.org/10.3389/fpls.2020.589189>
- Prudent M., Dequiedt S., Sorin C., Girodet S., Nowak V., Duc G., Salon C., and Maron P., 2019, The diversity of soil microbial communities matters when legumes face drought, *Plant, Cell and Environment*, 43(4): 1023-1035.
<https://doi.org/10.1111/pce.13712>
- Queiroz L., Pastore J., Cardoso D., Snak C., Lima A., Gagnon E., Vatanparast M., Holland A., and Egan A., 2015, A multilocus phylogenetic analysis reveals the monophyly of a circumscribed papilionoid legume tribe Diocleae with well-supported generic relationships, *Molecular Phylogenetics and Evolution*, 90: 1-19.
<https://doi.org/10.1016/j.ympev.2015.04.016>
- Raatz B., Mukankusi C., Lobaton J., Male A., Chisale V., Amsalu B., Fourie D., Mukamuhirwa F., Muimui K., Mutari B., Nchimbi-Msolla S., Nkalubo S., Tumsa K., Chirwa R., Maredia M., and He C., 2019, Analyses of African common bean (*Phaseolus vulgaris* L.) germplasm using a SNP fingerprinting platform: diversity, quality control and molecular breeding, *Genetic Resources and Crop Evolution*, 66: 707-722.
<https://doi.org/10.1007/s10722-019-00746-0>
- Raina A., Khan S., Wani M.R., Laskar R.A., and Mushtaq W., 2019, Chickpea (*Cicer arietinum* L.) cytogenetics, genetic diversity and breeding, *Advances in Plant Breeding Strategies: Legumes*, 7: 53-112.
<https://doi.org/10.37118/ijdr.26892.06.2023>
- Ren L., Huang W., and Cannon S., 2019, Reconstruction of ancestral genome reveals chromosome evolution history for selected legume species, *The New Phytologist*, 223(4): 2090-2103.
<https://doi.org/10.1111/nph.15770>

- Rendón-Anaya M., Montero-Vargas J., Saburido-Álvarez S., Vlasova A., Capella-Gutiérrez S., Ordaz-Ortiz J., Aguilar O., Vianello-Brondani R., Santalla M., Delaye L., Gabaldón T., Gepts P., Winkler R., Guigó R., Delgado-Salinas A., and Herrera-Estrella A., 2017, Genomic history of the origin and domestication of common bean unveils its closest sister species, *Genome Biology*, 18: 1-17.
<https://doi.org/10.1186/s13059-017-1190-6>
- Rivera A., Plans M., Sabaté J., Casañas F., Casals J., Rull A., and Simó J., 2018, The Spanish core collection of common beans (*Phaseolus vulgaris* L.): an important source of variability for breeding chemical composition, *Frontiers in Plant Science*, 9: 1642.
<https://doi.org/10.3389/fpls.2018.01642>
- Scarano D., Rubio F., Ruiz J., Rao R., and Corrado G., 2014, Morphological and genetic diversity among and within common bean (*Phaseolus vulgaris* L.) landraces from the Campania region (Southern Italy), *Scientia Horticulturae*, 180: 72-78.
<https://doi.org/10.1016/J.SCIENTA.2014.10.013>
- Schmutz J., McClean P., Mamidi S., Wu G., Cannon S., Grimwood J., Jenkins J., Shu S., Song Q., Chavarro C., Torres-Torres M., Geffroy V., Geffroy V., Moghaddam S., Gao D., Abernathy B., Barry K., Blair M., Brick M., Chovatia M., Gepts P., Goodstein D., Gonzales M., Hellsten U., Hyten D., Hyten D., Jia G., Kelly J., Kudrna D., Lee R., Richard M., Miklas P., Osorno J., Rodrigues J., Rodrigues J., Thareau V., Urrea C., Wang M., Yu Y., Zhang M., Wing R., Cregan P., Rokhsar D., and Jackson S., 2014, A reference genome for common bean and genome-wide analysis of dual domestications, *Nature Genetics*, 46: 707-713.
<https://doi.org/10.1038/ng.3008>
- Sharma S., Upadhyaya H., Varshney R., and Gowda C., 2013, Pre-breeding for diversification of primary gene pool and genetic enhancement of grain legumes, *Frontiers in Plant Science*, 4: 309.
<https://doi.org/10.3389/fpls.2013.00309>
- Smýkal P., Coyne C., Ambrose M., Maxted N., Schaefer H., Blair M., Berger J., Greene S., Nelson M., Besharat N., Vymyslický T., Toker C., Saxena R., Roorkiwal M., Pandey M., Hu J., Li Y., Wang L., Guo Y., Qiu L., Redden R., and Varshney R., 2015, Legume crops phylogeny and genetic diversity for science and breeding, *Critical Reviews in Plant Sciences*, 34: 104-143.
<https://doi.org/10.1080/07352689.2014.897904>
- Thudi M., Palakurthi R., Schnable J., Chitkineni A., Dreisigacker S., Mace E., Srivastava R., Satyavathi C., Odeny D., Tiwari V., Lam H., Hong Y., Singh V., Li G., Xu Y., Chen X., Kaila S., Nguyen H., Sivasankar S., Jackson S., Close T., Shubo W., and Varshney R., 2020, Genomic resources in plant breeding for sustainable agriculture, *Journal of Plant Physiology*, 257: 153351.
<https://doi.org/10.1016/j.jplph.2020.153351>
- Varshney R., 2016, Exciting journey of 10 years from genomes to fields and markets: Some success stories of genomics-assisted breeding in chickpea, pigeonpea and groundnut, *Plant Science*, 242: 98-107.
<https://doi.org/10.1016/j.plantsci.2015.09.009>
- Varshney R., Varshney R., Song C., Saxena R., Azam S., Yu S., Sharpe A., Cannon S., Baek J., Rosen B., Tar'an B., Millán T., Zhang, X., Ramsay L., Iwata A., Wang Y., Nelson W., Farmer A., Gaur P., Soderlund C., Penmetsa R., Xu C., Bharti A., He W., Winter P., Zhao S., Hane J., Carrasquilla-García N., Condie J., Upadhyaya H., Luo M., Thudi M., Gowda C., Singh N., Lichtenzweig J., Gali K., Rubio J., Nadarajan N., Doležel J., Bansal K., Xu X., Edwards D., Zhang G., Kahl G., Gil J., Singh K., Singh K., Datta S., Jackson S., Wang J., Wang J., and Cook D., 2013, Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement, *Nature Biotechnology*, 31: 240-246.
<https://doi.org/10.1038/nbt.2491>
- Wojciechowski M., Lavin M., and Sanderson M., 2004, A phylogeny of legumes (Leguminosae) based on analysis of the plastid *matK* gene resolves many well-supported subclades within the family, *American Journal of Botany*, 91(11): 1846-1862.
<https://doi.org/10.3732/ajb.91.11.1846>
- Yung W., Huang C., Li M., and Lam H., 2022, Changes in epigenetic features in legumes under abiotic stresses, *The Plant Genome*, 16(4): e20237.
<https://doi.org/10.1002/tpg2.20237>



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.