

Genomic Insights into the Domestication of Major Legume Crops

Xiaoxi Zhou ¹, Shengyu Chen ²✉

¹ Institute of Life Sciences, Jiyang College, Zhejiang A&F 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.2 doi: [10.5376/lgg.2024.15.0007](https://doi.org/10.5376/lgg.2024.15.0007)

Received: 05 Mar., 2024

Accepted: 06 Apr., 2024

Published: 16 Apr., 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, Genomic insights into the domestication of major legume crops, *Legume Genomics and Genetics*, 15(2): 49-61 (doi: [10.5376/lgg.2024.15.0007](https://doi.org/10.5376/lgg.2024.15.0007))

Abstract The domestication of major legume crops has been a pivotal process in the development of agriculture, providing essential nutrients and improving soil fertility through biological nitrogen fixation. This study synthesizes current genomic insights into the domestication and improvement of legume crops, highlighting the significant advancements made possible by modern genomic technologies. The study discusses the co-evolutionary process of domestication, the genetic bottlenecks encountered, and the role of wild relatives as reservoirs of genetic diversity for crop improvement. It also explores the impact of climate change on legume cultivation and the potential of genomic approaches to enhance stress tolerance and disease resistance. Furthermore, the study examines the contributions of genomic tools in understanding the molecular basis of agronomically important traits and the development of superior legume varieties through sequence-based breeding. By integrating genomic data with phenotyping and agronomic practices, this study provides a comprehensive perspective on the future directions for legume crop improvement, aiming to increase yield, quality, and resilience in the face of environmental challenges.

Keywords Legume domestication; Genomic technologies; Crop improvement; Stress tolerance; Disease resistance

1 Introduction

Legume crops, belonging to the family Fabaceae, are a cornerstone of global agriculture and human nutrition. They are a diverse group of plants that include beans, peas, lentils, chickpeas, and soybeans, among others. Legumes are highly valued for their ability to fix atmospheric nitrogen through symbiotic relationships with Rhizobium bacteria, which enhances soil fertility and reduces the need for synthetic fertilizers (Smýkal et al., 2020; Kagale and Close, 2021). This unique trait makes them crucial for sustainable agricultural practices and environmental health. Additionally, legumes are rich in proteins, vitamins, and minerals, making them an essential component of the human diet and a key source of animal feed (Smýkal et al., 2020; Kagale and Close, 2021).

The domestication of legumes dates back to the Neolithic era, approximately 12 000 years ago, when early agricultural societies began to cultivate these crops alongside cereals (Bohra et al., 2022). This co-evolutionary process has led to significant morphological and physiological changes in legume species, distinguishing them from their wild ancestors (Bohra et al., 2022). Despite their early domestication, the genetic and genomic understanding of legume domestication has lagged behind that of cereals (Bohra et al., 2022). Recent advances in genomic technologies, such as next-generation sequencing and high-throughput genotyping, have begun to shed light on the complex genetic changes associated with legume domestication and improvement (Kagale and Close, 2021; Zhang et al., 2022). These technologies have enabled researchers to identify key domestication traits and understand the molecular mechanisms underlying these traits (Kagale and Close, 2021; Zhang et al., 2022).

The study is to provide a comprehensive overview of the genomic insights into the domestication of major legume crops. By synthesizing recent research findings, this study aims to highlight the progress made in understanding the genetic basis of legume domestication and the potential applications of this knowledge in modern breeding programs. This study underscores the importance of integrating genomic approaches with traditional breeding methods to enhance the genetic gains in legume crops, thereby improving their yield, nutritional quality, and stress resilience. Furthermore, this study discusses the challenges and future directions in legume genomics research,

emphasizing the need for continued investment in genomic resources and technologies to fully realize the potential of legume crops in sustainable agriculture and global food security.

2 Historical Context of Legume Domestication

2.1 Early domestication events

The domestication of grain legumes occurred in parallel with cereals in various regions around the world, forming the economic foundation of early agricultural societies (Zhang et al., 2022). Human plant domestication began approximately 12 000 years ago in the Fertile Crescent and subsequently spread to other regions, including China, Mesoamerica, the Andes, Near Oceania, Sub-Saharan Africa, and eastern North America. The Indus Valley civilizations played a significant role in the domestication of several grain legumes, such as pigeon pea, black gram, green gram, lablab bean, moth bean, and horse gram. These early domestication events were characterized by the selection of desirable traits such as taste, yield, storage, and ease of cultivation (Ambika et al., 2022).

2.2 Archaeological evidence of legume use

Archaeological records provide substantial evidence of legume use and domestication. For instance, lentils (*Lens culinaris* Medik. subsp. *culinaris*) were first cultivated in Southwest Asia (SWA) around 8 000 to 10 000 years ago (Liber et al., 2021). However, the exact number of independent domestication events and the specific regions within SWA where these events occurred remain unclear (Liber et al., 2021). Neolithic archaeological records indicate that several legume crops were first domesticated by early civilizations in the Indian subcontinent (Ambika et al., 2022). These records highlight the importance of legumes in early agricultural societies and their role in the development of human civilization.

2.3 Geographic distribution of domesticated legumes

The geographic distribution of domesticated legumes is extensive, reflecting their adaptability and importance in various agricultural systems. Legumes were domesticated in multiple regions, including the Fertile Crescent, China, Mesoamerica, the Andes, Near Oceania, Sub-Saharan Africa, and eastern North America (Ambika et al., 2022). The spread of domesticated legumes was facilitated by human migration and trade, leading to the establishment of these crops in diverse environments. For example, the domestication and subsequent spread of lentils resulted in the emergence of two domesticated gene pools in Southwest Asia, with admixed varieties found throughout their range (Figure 1) (Liber et al., 2021). This widespread distribution underscores the significance of legumes in global agriculture and their contribution to food security and agrobiodiversity.

Liber et al. (2021) highlights the significant geographic and genetic differentiation among *Lens culinaris* populations. The spatial interpolation maps depict distinct distribution patterns, revealing how historical and environmental factors have shaped the genetic structure of lentil populations across various regions. The differentiation among the groups suggests varying degrees of adaptation and domestication processes. This study underscores the importance of understanding regional genetic diversity for effective breeding strategies and conservation efforts. By identifying specific genetic clusters, the research provides a framework for improving lentil varieties, enhancing their resilience, and maintaining genetic resources critical for sustainable agriculture.

3 Genomic Techniques and Tools

3.1 Sequencing technologies

The advent of next-generation sequencing (NGS) technologies has revolutionized the field of legume genomics. NGS allows for high-throughput sequencing, which has significantly reduced the time and cost associated with genome sequencing. This technology has enabled the sequencing and re-sequencing of numerous legume species, providing comprehensive insights into their genetic makeup (Varshney et al., 2018; Afzal et al., 2020). The availability of genome sequence assemblies and the development of high-density genetic maps have facilitated the identification of markers associated with various agronomic traits (Varshney et al., 2018). Additionally, RNA sequencing (RNA-Seq) has been instrumental in characterizing differentially expressed genes, thereby contributing to functional genomics studies (Afzal et al., 2020).

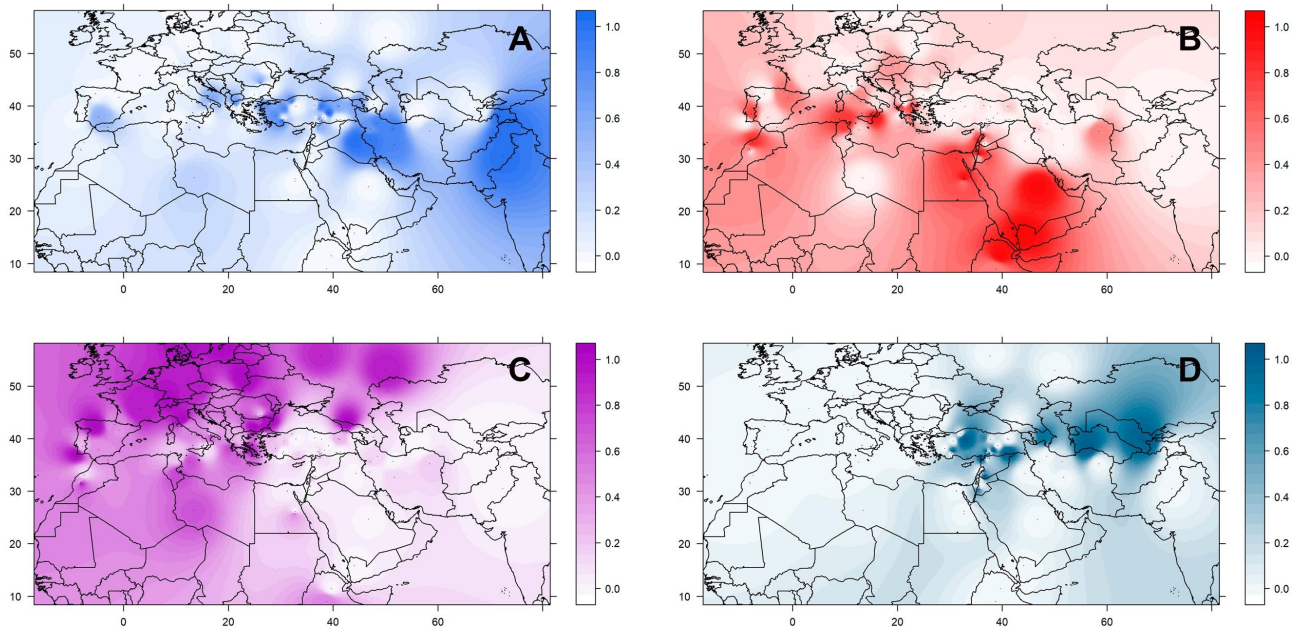


Figure 1 Geographical distribution of the four groups (A-D) of *L. culinaris* identified in the STRUCTURE model $K = 4$ (Adopted from Liber et al., 2021)

Image caption: Spatial interpolation is based on the Q-matrix of proportional memberships of individual accessions to each of the four groups (Adopted from Liber et al., 2021)

3.2 Genome-wide association studies (GWAS)

Genome-wide association studies (GWAS) have emerged as a powerful tool for dissecting complex traits in legumes. GWAS involves scanning the genome for single-nucleotide polymorphisms (SNPs) that are associated with specific phenotypic traits. This approach has been widely adopted in agricultural genetics to understand the genetic basis of traits such as protein content and disease resistance (Liu and Yan, 2018; Susmitha et al., 2023). The integration of machine learning methods has further enhanced the efficiency of GWAS by addressing challenges related to dimensionality reduction (Susmitha et al., 2023). The markers identified through GWAS can be utilized for genomic selection, thereby accelerating the breeding of superior legume lines (Susmitha et al., 2023).

3.3 Comparative genomics

Comparative genomics involves comparing the genomes of different species to identify conserved and unique genetic elements. This approach has been particularly useful in translating knowledge from model legumes to less-studied legume crops. Databases like LegumeIP V3 provide genomic and transcriptomic data for multiple legume species, facilitating the discovery of key genes and pathways that regulate important agronomic traits (Dai et al., 2020). Comparative genomics has also been instrumental in understanding the molecular underpinnings of domestication and crop evolution in grain legumes (Zhang et al., 2022; Bohra et al., 2022).

3.4 Transcriptomics and epigenomics

Transcriptomics, particularly RNA-Seq, has provided valuable insights into the gene expression profiles of legumes under various conditions. This technique has been used to study the host transcriptional changes during interactions between legumes and pathogens, identifying key players in both resistant and susceptible interactions (Kankanala et al., 2019). Epigenomics, which involves the study of heritable changes in gene expression that do not involve changes to the underlying DNA sequence, is another emerging field. Although less explored in legumes, epigenomic studies have the potential to uncover regulatory mechanisms that contribute to trait variation and adaptation (Bohra et al., 2022).

In summary, the integration of advanced genomic techniques and tools such as NGS, GWAS, comparative genomics, and transcriptomics has significantly advanced our understanding of legume genomics. These technologies not only facilitate the identification of key genetic elements but also accelerate the breeding of improved legume varieties, thereby contributing to sustainable agricultural practices.

4 Key Genomic Findings in Legume Domestication

4.1 Soybean (*Glycine max*)

Soybean domestication has led to significant genetic bottlenecks, which are evident in the reduced genetic diversity of cultivated varieties compared to their wild relatives. The resequencing of 481 diverse soybean accessions, including both wild (*Glycine soja*) and cultivated (*Glycine max*) varieties, identified 7.8 million SNPs and revealed numerous highly conserved regions among cultivated varieties, suggesting strong selection during domestication (Figure 2) (Valliyodan et al., 2021). This selection has resulted in the loss of certain alleles and the fixation of others that are beneficial for traits such as seed size, oil content, and stress resistance (Petereit et al., 2022).

Genomic studies have identified specific regions under selection in soybean. For instance, the development of soybean pangenomes has highlighted significant structural variations between individuals, providing insights into what has been selected for or lost during domestication and breeding (Petereit et al., 2022). Additionally, the draft genome sequence of *Glycine latifolia*, a wild relative of soybean, has provided valuable comparative data, revealing genes related to defense responses and stress tolerance that are overrepresented in wild species but may have been lost or altered in cultivated soybeans (Liu et al., 2018).

The phylogenetic analysis of 481 diverse soybean accessions using 25 496 SNPs by Valliyodan et al. (2021) provides a comprehensive understanding of the genetic variation and evolutionary relationships between cultivated and wild soybean species. The tree, rooted between *G. max* and *G. soja*, illustrates the genetic diversity within and between these groups. The inclusion of accessions from various countries (China, Korea, Japan, the United States, and Russia) highlights the global genetic diversity of soybeans and the impact of geographical origin on genetic variation. This analysis underscores the significance of wild relatives as genetic resources and the importance of maintaining genetic diversity for crop improvement. The detailed phylogenetic relationships revealed by this study can guide future breeding programs aimed at improving stress resistance, yield, and other agronomic traits in soybeans.

4.2 Common bean (*Phaseolus vulgaris*)

Common bean exhibits considerable genetic diversity and complex population structure, influenced by its domestication history and adaptation to diverse agroecological settings. A study involving 315 accessions of the USDA common bean core collection identified significant genetic variation and population structure, with 4 654 SNPs used to evaluate resistance to soybean cyst nematode (SCN) (Shi et al., 2021). This genetic diversity is crucial for breeding programs aimed at improving resistance to biotic and abiotic stresses.

The domestication of common bean has led to a suite of traits known as the domestication syndrome, which includes changes in seed size, pod dehiscence, and growth habit. Genomic prediction models have been developed to select new lines based on genomic data, accounting for genotype-by-environment interactions (G×E) to improve traits such as yield and stress resistance (Keller et al., 2020). Furthermore, genomic studies have identified regions associated with resistance to common bacterial blight (CBB), with significant peaks near the SAP6 molecular marker, providing targets for marker-assisted breeding (Simons et al., 2021). The integration of genomic resources and marker-assisted methods is essential for developing common bean varieties that are resilient to future environmental challenges (Assefa et al., 2019).

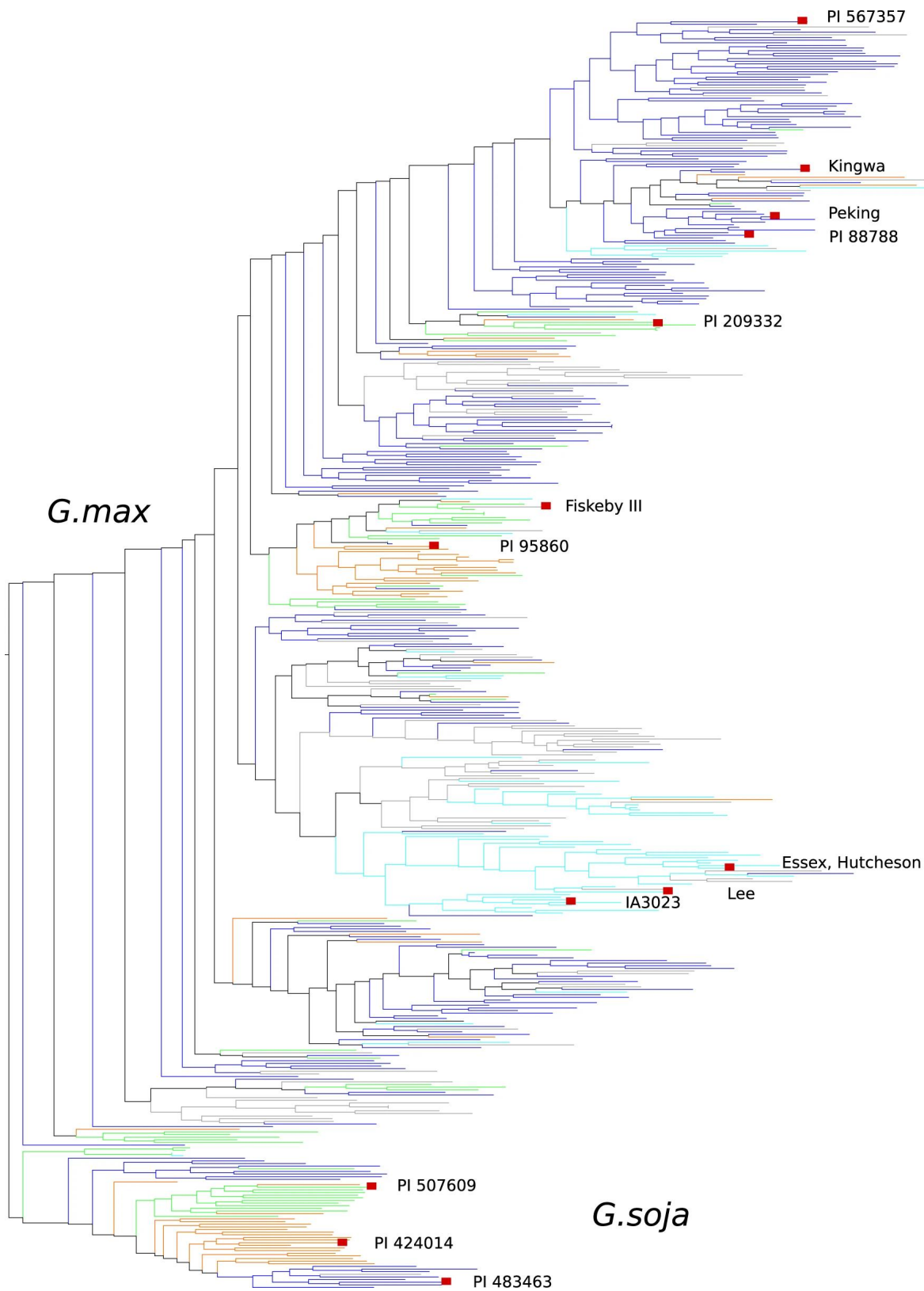


Figure 2 Genetic variation among 481 diverse soybean accessions, inferred from genomic re-sequencing (Adopted from Valliyodan et al., 2021)

Image caption: Phylogenetic tree of the 481 re-sequenced accessions. The tree is rooted between *G. max* and *G. soja* accessions. Colors indicate countries of origin: blue, China; orange, Korea; green, Japan; cyan, United States; gray, all others (predominantly from Russia). Cultivars of interest are highlighted on the tree (Adopted from Valliyodan et al., 2021)

4.3 Chickpea (*Cicer arietinum*)

Chickpea (*Cicer arietinum*) is one of the most widely cultivated legumes globally, known for its nutritional benefits and adaptability to various climatic conditions. The domestication of chickpea has led to a significant reduction in genetic diversity, necessitating the exploration of wild species for crop improvement. Studies have shown that crossing cultivated chickpea with wild species such as *C. reticulatum* and *C. echinospermum* can help identify genetic traits beneficial for domestication and breeding (Shin et al., 2019). The use of genomic technologies, including genomics-assisted breeding and high-throughput phenotyping, has facilitated the development of stress-resistant and high-yielding chickpea varieties (Koul et al., 2022). Additionally, the Translational Chickpea Genomics Consortium (TCGC) has been instrumental in deploying modern genomics approaches to enhance chickpea production and productivity (Palakurthi et al., 2021).

Research has identified several genomic regions associated with key traits in chickpea. For instance, a dense genetic map and QTL analysis have revealed nine main-effect QTLs and 955 epistatic QTLs linked to resistance against the pod borer, *Helicoverpa armigera*, with significant phenotypic variance (Barmukh et al., 2020). Another study identified QTLs controlling vigour and related traits, with robust markers developed for marker-assisted selection (Nguyen et al., 2021). Furthermore, the MutMap approach has enabled the rapid identification of candidate genes and development of markers associated with early flowering and enhanced seed size, crucial for mitigating yield losses under drought stress (Manchikatla et al., 2021). These findings underscore the importance of genomic regions in improving chickpea traits and enhancing its adaptability to various environmental conditions.

4.4 Lentil (*Lens culinaris*)

Lentil (*Lens culinaris*) has a rich domestication history, first cultivated in Southwest Asia around 8 000–10 000 years ago. Genotyping-by-sequencing (GBS) of wild and domesticated lentil accessions has provided insights into its evolutionary history and domestication process. The study identified four groups within the *Lens* genus and confirmed gene flow between cultivated varieties and their progenitor, *L. culinaris* subsp. *orientalis*, albeit at low levels (Liber et al., 2021). The findings suggest a protracted domestication process with two domesticated gene pools emerging in Southwest Asia. This genetic diversity is crucial for understanding the distribution of genetic traits and targeting genes in breeding programs.

The genomic regions under selection in lentil have been explored through various studies. The collection and conservation of lentil germplasm have highlighted the importance of core collections for germplasm users and the necessity of sharing omic and non-omic characterization data (Piergiovanni et al., 2022). Additionally, the identification of alleles involved in domestication and adaptation to climatic variables has provided valuable insights into the genetic basis of lentil traits. These genomic regions are essential for improving lentil varieties and enhancing their adaptability to different environments.

5 Comparative Analysis of Domestication Genomics Across Legumes

5.1 Common domestication traits

Domestication of legumes, much like other crops, has led to significant morpho-physiological changes that distinguish domesticated varieties from their wild ancestors. Common traits selected during domestication include reduced seed dormancy, decreased pod shattering, increased organ size, and earlier flowering and maturity (Yundaeng et al., 2019; Bohra et al., 2022; Zhang et al., 2022). These traits have been crucial for adapting legumes to human agricultural practices and improving their yield and reliability as food sources. For instance, in moth bean (*Vigna aconitifolia*), key domestication traits such as seed dormancy and pod shattering are controlled by large-effect QTLs, indicating that a few major genetic changes have had significant impacts on domestication (Yundaeng et al., 2019). Similarly, in narrow-leafed lupin (*Lupinus angustifolius*), genes associated with vernalization, cold-responsiveness, and phosphatidylinositol signaling pathways have been linked to early phenology, while genes controlling photosynthesis and abiotic stress tolerance are associated with high yield (Figure 3) (Plewiński et al., 2020).

Plewiński et al. (2020) investigates the phenotypic variation and genetic heritability of key agronomic traits in 126 accessions of *Lupinus angustifolius* across different domestication types, over two locations and years. The findings reveal significant variability due to environmental factors, with high heritability for most traits except yield. Cultivars generally exhibited superior traits, such as early flowering and higher yield, compared to wild or primitive lines. The data underscore the influence of environmental conditions, such as temperature, on trait expression and the importance of heritability in breeding programs. This research provides valuable insights for the genetic improvement of *L. angustifolius*, emphasizing the need for tailored breeding strategies that consider both genetic and environmental interactions.

5.2 Divergent evolutionary pathways

Despite the common traits selected during domestication, different legume species have followed divergent evolutionary pathways influenced by their unique environmental conditions and human selection pressures. For example, the pea (*Pisum sativum*) genome shows intense gene dynamics and significant genomic rearrangements compared to other legumes, likely due to the expansion of repetitive elements (Kreplak et al., 2019). This divergence is also evident in the genetic linkage maps and QTL analyses of various legumes, which reveal species-specific adaptations. In moth bean, comparative genome analysis shows high synteny with other *Vigna* species, yet distinct domestication-related QTLs have been identified that are unique to moth bean (Yundaeng et al., 2019). Similarly, rice bean (*Vigna umbellata*) exhibits unique genetic traits related to flowering potential, habit, and palatability, which are crucial for its efficient domestication and nutritional value (Kaul et al., 2022).

5.3 Insights from cross-species genomic comparisons

Cross-species genomic comparisons have provided valuable insights into the domestication and evolution of legumes. Phylogenomic studies have resolved deep divergences in the legume phylogeny, revealing that all six subfamilies of legumes originated nearly simultaneously, challenging the traditional view of some subfamilies as ‘basal’ or ‘early-diverging’ (Koenen et al., 2019). This simultaneous origin suggests that rapid successive speciation events have shaped the diversity of legume traits. Additionally, the use of whole genome sequencing and omics-level data has expanded our understanding of the genetic changes accompanying domestication. For instance, genomic tools have enabled the identification of key players in the interactions between legumes and pathogens, providing insights into the molecular basis of disease resistance (Kankanala et al., 2019). Furthermore, meta-QTL analyses in common bean (*Phaseolus vulgaris*) have identified stable genomic regions associated with yield-related traits, which are valuable for breeding programs (Arriagada et al., 2022).

In summary, the comparative analysis of domestication genomics across legumes highlights both common traits and divergent evolutionary pathways. Cross-species genomic comparisons have enriched our understanding of the genetic basis of domestication and provided practical insights for modern breeding strategies aimed at improving legume crops.

6 Implications for Breeding and Crop Improvement

6.1 Utilizing genomic data for breeding programs

The integration of genomic data into breeding programs has revolutionized the way legume crops are improved. Modern genomic tools, such as next-generation sequencing, have facilitated the development of high-density genetic maps and the identification of markers associated with key agronomic traits (Varshney et al., 2018). These advancements enable marker-assisted selection and backcrossing, which have been successfully used to develop superior legume lines (Varshney et al., 2018). Additionally, genomic prediction models, which utilize both recent and historical phenotypic and genotypic data, have shown promise in selecting new lines without the need for extensive phenotyping (Keller et al., 2020). This approach is particularly beneficial for under-utilized legume crops like common beans, where genotype-by-environment interactions are critical for developing stress-resistant varieties (Keller et al., 2020).

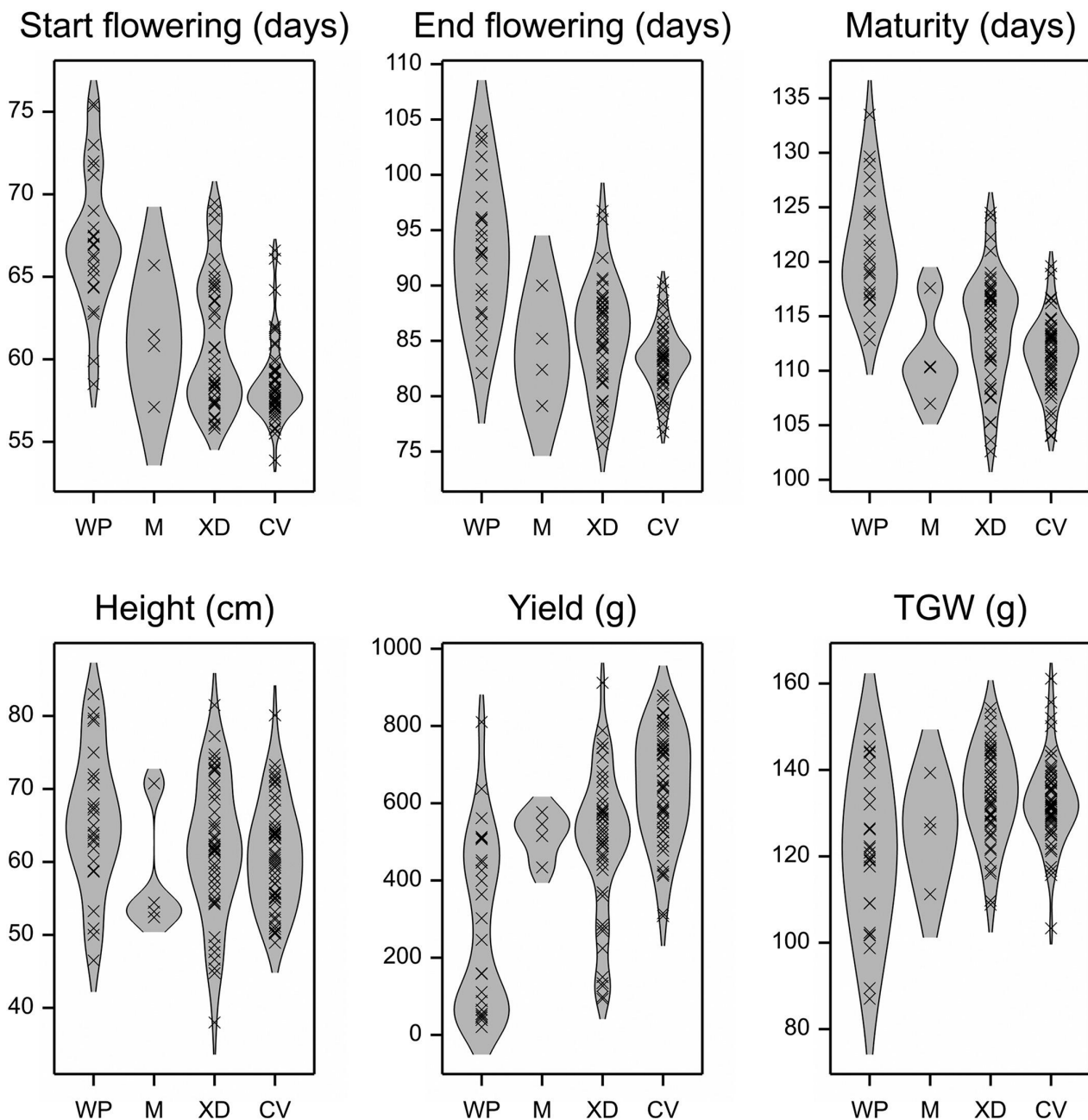


Figure 2 Distributions of general mean values for observed traits (over years, locations) in four groups according to domestication type of analysed lines (Adopted from Plewiński et al., 2020)

Image caption: Traits include start flowering (days from sowing to start flowering of 1/3 plants in a plot), end flowering (days from sowing to end flowering), maturity (number of days measured when majority of pods in the plot was ready to harvest), yield (total seed yield per plot), height (plant height at maturity), TGW (thousand grain weight). Experiments were performed at Plant Breeding Smolice Ltd. in Przebędowo (52°35'N 17°01'E) and Poznan Plant Breeding Ltd. in Wiatrowo (52°45'N 17°08'E) during growing seasons of 2015 and 2016. Domestication types are as follows, XD: Cross derivative or breeding line; CV: Cultivar; M: Mutant; WP: Wild or primitive (Adopted from Plewiński et al., 2020)

6.2 Potential for improving yield and stress resistance

Genomic approaches have significantly contributed to understanding the molecular mechanisms underlying stress tolerance in legumes. Proteomic analyses have identified stress-responsive proteins involved in photosynthesis, carbohydrate metabolism, and signal transduction, which are crucial for developing stress-tolerant varieties (Jan et al., 2022). Furthermore, genomic studies have elucidated the host transcriptional changes during interactions between legumes and pathogens, providing insights into disease resistance mechanisms (Kankanala et al., 2019).

The integration of genomic data with high-throughput phenomics and simulation modeling can enhance genetic gains in legume breeding programs, leading to improved varieties that perform well under environmental stresses such as drought and low soil fertility (Varshney et al., 2018). This is essential for increasing productivity and ensuring food security in the face of climate change (Mousavi-Derazmahalleh et al., 2018; Thudi et al., 2020).

6.3 Future directions in legume crop genomics

Future research in legume crop genomics should focus on several key areas to further enhance breeding programs. First, there is a need for continuous population improvement after each breeding cycle to accelerate genetic gains (Varshney et al., 2018). This can be achieved through sequence-based breeding approaches, which include parental selection, enhancing genetic diversity, and forward breeding for early generation selection (Varshney et al., 2018). Additionally, the adoption of speed breeding technologies, which allow for multiple generations per year, can significantly shorten breeding cycles and increase genetic gains (Varshney et al., 2018).

Moreover, the exploration of wild relatives of legumes as reservoirs of novel genetic variation can provide new alleles for crop improvement (Bohra et al., 2022). Whole genome sequencing and omics-level data can expand our capacity to monitor genetic changes and identify breeding targets for developing modern cultivars with improved yield, quality, and stress tolerance (Bohra et al., 2022). Finally, integrating genomic breeding approaches with machine learning and artificial intelligence can enhance the precision and efficiency of breeding programs, ultimately leading to the development of climate-resilient and high-nutrition legume varieties (Thudi et al., 2020).

By leveraging these genomic insights and technologies, breeding programs can make significant strides in improving legume crops, ensuring their productivity and resilience in the face of evolving agricultural challenges.

7 Challenges and Future Perspectives

7.1 Technological and analytical challenges

The rapid advancement in genomic technologies has significantly enhanced our understanding of legume domestication and crop improvement. However, several technological and analytical challenges persist. One major challenge is the high cost and complexity associated with high-quality and precise phenotyping, which is crucial for accurate genomic selection and breeding (Varshney et al., 2018; Afzal et al., 2020). Additionally, the integration of large-scale genomic data with phenotypic data requires sophisticated bioinformatics tools and expertise, which can be a limiting factor for many research programs (Varshney et al., 2018; Afzal et al., 2020). The development of comprehensive genomic resources for underutilized legume species also remains limited, hindering their potential for genetic improvement (Jha et al., 2022). Furthermore, the continuous evolution of pathogens necessitates ongoing research to identify and introgress new disease resistance genes into legume cultivars (Kankanala et al., 2019).

7.2 Ethical and regulatory considerations

The application of advanced genomic technologies, such as gene editing, in legume breeding raises several ethical and regulatory concerns. The use of CRISPR/Cas9 and other gene-editing tools has the potential to revolutionize legume crop improvement by introducing desirable traits with high precision (Baloğlu et al., 2022). However, the regulatory frameworks governing the use of these technologies vary significantly across different regions. For instance, the European Union has stringent regulations on genetically modified organisms (GMOs), which can impede the adoption of gene-edited crops (Baloğlu et al., 2022). In contrast, the United States has a more permissive regulatory environment, which may facilitate the development and commercialization of gene-edited legumes (Baloğlu et al., 2022). Ethical considerations also include the potential impact on biodiversity and the need to ensure that the benefits of these technologies are accessible to smallholder farmers and developing countries (Varshney et al., 2018; Jha et al., 2022).

7.3 Prospects for future research

Future research in legume genomics should focus on several key areas to address the existing challenges and harness the full potential of genomic technologies. First, there is a need to develop cost-effective and

high-throughput phenotyping methods to complement genomic selection and accelerate breeding programs (Varshney et al., 2018; Afzal et al., 2020). Second, expanding genomic resources for underutilized legume species will be crucial for enhancing their genetic diversity and resilience to environmental stresses (Jha et al., 2022). Third, integrating multi-omics approaches, including genomics, transcriptomics, proteomics, and metabolomics, can provide a holistic understanding of the molecular mechanisms underlying important agronomic traits (Afzal et al., 2020; Jan et al., 2022). Additionally, the adoption of speed breeding techniques can significantly reduce the breeding cycle time and accelerate the development of improved legume varieties (Varshney et al., 2018; Jha et al., 2022). Finally, addressing the ethical and regulatory challenges associated with gene editing will be essential for the widespread adoption of these technologies in legume breeding (Baloğlu et al., 2022). By overcoming these challenges and leveraging advanced genomic tools, future research can contribute to the sustainable improvement of legume crops, ensuring food security and agricultural resilience in the face of climate change and other global challenges.

8 Concluding Remarks

The domestication of major legume crops has been a complex process influenced by both natural and human selection. Recent genomic studies have provided significant insights into this process. For instance, whole genome sequencing and omics-level data have expanded our understanding of the genetic changes accompanying domestication, revealing the importance of wild relatives as reservoirs of genetic variation for crop improvement. Genomic approaches have also highlighted the centers of origin and crucial domestication traits of grain legumes, emphasizing the impact of domestication on both above-ground and below-ground traits. Additionally, the identification of disease resistance genes and their interactions with pathogens has been revolutionized by genomic tools, providing valuable insights for breeding programs. The integration of modern genomics, phenotyping, and systems modeling has been proposed to accelerate genetic gains in legumes, particularly under stress conditions. Furthermore, the sequencing of legume genomes, such as the reference genome for pea, has shed light on legume genome evolution and the molecular basis of agronomically important traits.

Genomic research has proven to be indispensable in understanding and advancing legume domestication. The advent of next-generation sequencing (NGS) and high-throughput genotyping methods has opened new avenues for legume research, enabling the identification of structural variations and functional genomics. These technologies have facilitated the development of high-quality genome assemblies and pangenomes, which are crucial for studying variations between individuals and populations. Genomic tools have also been instrumental in identifying stress-responsive proteins and understanding the molecular mechanisms of stress tolerance in legumes, which is vital for improving crop resilience under changing environmental conditions. Moreover, the exploration of wild legume species for domestication has been supported by genomic studies, which provide key preliminary information for breeding and improvement programs.

The integration of genomic research into legume domestication has significantly advanced our understanding of the genetic and molecular basis of domestication traits. This knowledge is crucial for developing improved legume varieties that are more resilient, productive, and nutritionally valuable. Future research should continue to leverage genomic technologies to explore the untapped genetic diversity in wild relatives and to address the challenges posed by climate change and evolving pathogens. By doing so, we can ensure the sustainable production of legume crops, which are essential for global food security and agricultural biodiversity.

Acknowledgments

The authors extend 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

- Afzal M., Alghamdi S.S., Migdadi H.H., Khan M.A., Mirza S.B., and El-Harty E., 2020. Legume genomics and transcriptomics: from classic breeding to modern technologies, Saudi Journal of Biological Sciences, 27(1): 543-555.
<https://doi.org/10.1016/j.sjbs.2019.11.018>
- Ambika Aski M.S., Gayacharan Hamwiah A., Talukdar A., Gupta S., Sharma B., Joshi R., Upadhyaya H., 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>
- Arriagada O., Arévalo B., Cabeza R., Carrasco B., and Schwember A., 2022, Meta-QTL analysis for yield components in common bean (*Phaseolus vulgaris* L.), Plants, 12(1): 117.
<https://doi.org/10.3390/plants12010117>
- Assef 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>
- Baloğlu M., Altunoğlu Y., Baloglu P., Yıldız A., Türkölmez N., and Çiftçi Y., 2022, Gene-editing technologies and applications in legumes: progress, evolution, and future prospects, Frontiers in Genetics, 13: 859437.
<https://doi.org/10.3389/fgene.2022.859437>
- Barmukh R., Roorkiwal M., Jaba J., Chitikineni A., Mishra S., Sagurthi S., Munghate R., Sharma H., and Varshney R., 2020, Development of a dense genetic map and QTL analysis for pod borer *Helicoverpa armigera* (Hübner) resistance component traits in chickpea (*Cicer arietinum* L.), The Plant Genome, 14(1): e20071.
<https://doi.org/10.1002/tpg2.20071>
- 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(11): 1554-1572.
<https://doi.org/10.1093/pcp/pcac086>
- Dai X., Zhuang Z., Boschiero C., Dong Y., and Zhao P., 2020, LegumeIP V3: from models to crops—an integrative gene discovery platform for translational genomics in legumes, Nucleic Acids Research, 49(D1): D1472-D1479.
<https://doi.org/10.1093/nar/gkaa976>
- Jan N., Rather A., John R., Chaturvedi P., Ghatak A., Weckwerth W., Zargar S., Mir R., Khan M., and Mir R., 2022, Proteomics for abiotic stresses in legumes: present status and future directions, Critical Reviews in Biotechnology, 43(2): 171-190.
<https://doi.org/10.1080/07388551.2021.2025033>
- Jha U., Nayyar H., Parida S., Bakir M., Wettberg E., and Siddique K., 2022, Progress of genomics-driven approaches for sustaining underutilized legume crops in the post-genomic era, Frontiers in Genetics, 13: 831656.
<https://doi.org/10.3389/fgene.2022.831656>
- Kagale S., and Close T., 2021, Legumes: embracing the genome era, Legume Science, 3(3): 1-4.
<https://doi.org/10.1002/leg3.113>
- Kankanala P., Nandety R., and Mysore K., 2019, Genomics of plant disease resistance in legumes, Frontiers in Plant Science, 10: 1345.
<https://doi.org/10.3389/fpls.2019.01345>
- Kaul T., Easwaran M., Thangaraj A., Meyyazhagan A., Nehra M., Raman N., Verma R., Sony S., Abdel K., Bharti J.G., Badapanda C., and Balasubramanian B., 2022, De novo genome assembly of rice bean (*Vigna umbellata*)-A nominated nutritionally rich future crop reveals novel insights into flowering potential, habit, and palatability centric - traits for efficient domestication, Frontiers in Plant Science, 13: 739654.
<https://doi.org/10.3389/fpls.2022.739654>
- Keller B., Ariza-Suarez D., Hoz J., Aparicio J., Portilla-Benavides A., Buendia H., Mayor V., Studer B., and Raatz B., 2020, Genomic prediction of agronomic traits in common bean (*Phaseolus vulgaris* L.) under environmental stress, Frontiers in Plant Science, 11: 1001.
<https://doi.org/10.3389/fpls.2020.01001>
- 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(3): 1355-1369.
<https://doi.org/10.1111/nph.16290>
- Koul B., Sharma K., Sehgal V., Yadav D., Mishra M., and Bharadwaj C., 2022, Chickpea (*Cicer arietinum* L.) biology and biotechnology: from domestication to biofortification and biopharming, Plants, 11(21): 2926.
<https://doi.org/10.3390/plants11212926>
- Kreplak J., Madoui M., Čápal 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., Belser C., Berrabah W., Toegelová H., Milec Z., Vrána J., Lee H., Kougbéadjó 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., Lichtenzveig 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(9): 1411-1422.
<https://doi.org/10.1038/s41588-019-0480-1>
- 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>
- Liu H., and Yan J., 2018, Crop genome-wide association study: a harvest of biological relevance, *The Plant Journal*, 97(1): 8-18.
<https://doi.org/10.1111/tpj.14139>
- Liu Q., Chang S., Hartman G., and Domier L., 2018, Assembly and annotation of a draft genome sequence for *Glycine latifolia*, a perennial wild relative of soybean, *The Plant Journal*, 95(1): 71-85.
<https://doi.org/10.1111/tpj.13931>
- Manchikatlal P., Kalavikatte D., Mallikarjuna B., Palakurthi R., Khan A., Jha U., Bajaj P., Singam P., Chitikineni A., Varshney R., and Thudi M., 2021, MutMap approach enables rapid identification of candidate genes and development of markers associated with early flowering and enhanced seed size in chickpea (*Cicer arietinum* L.), *Frontiers in Plant Science*, 12: 688694.
<https://doi.org/10.3389/fpls.2021.688694>
- Mousavi-Derazmahalleh M., Bayer P., Hane J., Valliyodan B., Nguyen H., Nelson M., Erskine W., Varshney R., Papa R., and Edwards D., 2018, Adapting legume crops to climate change using genomic approaches, *Plant, Cell and Environment*, 42(1): 6-19.
<https://doi.org/10.1111/pce.13203>
- Nguyen D., Hayes J., Atieno J., Li Y., Baumann U., Pattison A., Bramley H., Hobson K., Roorkiwal M., Varshney R., Colmer T., and Sutton T., 2021, The genetics of vigour-related traits in chickpea (*Cicer arietinum* L.): insights from genomic data, *Theoretical and Applied Genetics*, 135(1): 107-124.
<https://doi.org/10.1007/s00122-021-03954-4>
- Palakurthi R., Jayalakshmi V., Kumar Y., Kulwal P., Yasin M., Kute N., Laxuman C., Yeri S., Vemula A., Rathore A., Samineni S., Soren K., Mondal B., Dixit G., Bharadwaj C., Chaturvedi S., Gaur P., Roorkiwal M., Thudi M., Singh N., and Varshney R., 2021, Translational chickpea genomics consortium to accelerate genetic gains in chickpea (*Cicer arietinum* L.), *Plants*, 10(12): 2583.
<https://doi.org/10.3390/plants10122583>
- Petereit J., Marsh J., Bayer P., Danilevich M., Thomas W., Batley J., and Edwards D., 2022, Genetic and genomic resources for soybean breeding research, *Plants*, 11(9): 1181.
<https://doi.org/10.3390/plants11091181>
- Piergiorganni A., 2022, Ex situ conservation of plant genetic resources: an overview of chickpea (*Cicer arietinum* L.) and lentil (*Lens culinaris* Medik.) worldwide collections, *Diversity*, 14(11): 941.
<https://doi.org/10.3390/d14110941>
- Plewiński P., Ćwiek-Kuczyńska H., Rudy E., Bielski W., Rychel-Bielska S., Stawiński S., Barzyk P., Krajewski P., Naganowska B., Wolko B., and Książkiewicz M., 2020, Innovative transcriptome-based genotyping highlights environmentally responsive genes for phenology, growth and yield in a non-model grain legume, *Plant, Cell and Environment*, 43(11): 2680-2698.
<https://doi.org/10.1111/pce.13880>
- Shi A., Gepts P., Song Q., Xiong H., Michaels T., and Chen S., 2021, Genome-wide association study and genomic prediction for soybean cyst nematode resistance in USDA common bean (*Phaseolus vulgaris*) core collection, *Frontiers in Plant Science*, 12: 624156.
<https://doi.org/10.3389/fpls.2021.624156>
- Shin M., Bulyntsev S., Chang P., Korbu L., Carrasquilla-Garcia N., Vishnyakova M., Samsonova M., Cook D., and Nuzhdin S., 2019, Multi-trait analysis of domestication genes in *Cicer arietinum* - *Cicer reticulatum* hybrids with a multidimensional approach: modeling wide crosses for crop improvement, *Plant Science*, 285: 122-131.
<https://doi.org/10.1016/J.PLANTSCI.2019.04.018>
- Simons K., Oladzad A., Lamma P., McClean P., Osorno J., and Pasche J., 2021, Using breeding populations with a dual purpose: cultivar development and gene mapping—a case study using resistance to common bacterial blight in dry bean (*Phaseolus vulgaris* L.), *Frontiers in Plant Science*, 12: 621097.
<https://doi.org/10.3389/fpls.2021.621097>
- Smykal P., Wettberg E., and McPhee K., 2020, Legume genetics and biology: from Mendel's pea to legume genomics, *International Journal of Molecular Sciences*, 21(9): 3336.
<https://doi.org/10.3390/ijms21093336>
- Susmitha P., Kumar P., Yadav P., Sahoo S., Kaur G., Pandey M., Singh V., Tseng T., and Gangurde S., 2023, Genome-wide association study as a powerful tool for dissecting competitive traits in legumes, *Frontiers in Plant Science*, 14: 1123631.
<https://doi.org/10.3389/fpls.2023.1123631>
- Thudi M., Palakurthi R., Schnable J., Chitikineni 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>
- Valliyodan B., Brown A., Wang J., Patil G., Liu Y., Otyama P., Nelson R., Vuong T., Song Q., Musket T., Wagner R., Marri P., Reddy S., Sessions A., Wu X., Grant D., Bayer P., Roorkiwal M., Varshney R., Liu X., Edwards D., Xu D., Joshi T., Cannon S., and Nguyen H., 2021, Genetic variation among 481 diverse soybean accessions, inferred from genomic re-sequencing, *Scientific Data*, 8(1): 50.
<https://doi.org/10.1038/s41597-021-00834-w>
- Varshney R., Pandey M., Bohra A., Singh V., Thudi M., and Saxena R., 2018, Toward the sequence-based breeding in legumes in the post-genome sequencing

era, Theoretische Und Angewandte Genetik, 132(3): 797-816.

<https://doi.org/10.1007/s00122-018-3252-x>

Varshney R., Thudi M., Pandey M., Tardieu F., Ojiewo C., Vadez V., Whitbread A., Siddique K., Nguyen H., Carberry P., and Bergvinson D., 2018, Accelerating genetic gains in legumes for the development of prosperous smallholder agriculture: integrating genomics, phenotyping, systems modelling and agronomy, Journal of Experimental Botany, 69(13): 3293-3312.

<https://doi.org/10.1093/jxb/ery088>

Yundaeng C., Somta P., Amkul K., Kongjaimun A., Kaga A., and Tomooka N., 2019, Construction of genetic linkage map and genome dissection of domestication-related traits of moth bean (*Vigna aconitifolia*), a legume crop of arid areas, Molecular Genetics and Genomics, 294: 621-635.

<https://doi.org/10.1007/s00438-019-01536-0>

Zhang H., Mascher M., Abbo S., and Jayakodi M., 2022, Advancing grain legumes domestication and evolution studies with genomics, Plant and Cell Physiology, 63(11): 1540-1553.

<https://doi.org/10.1093/pcp/pcac062>



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.
