Genomics and Applied Biology 2024, Vol.15, No.4, 212-222
http://bioscipublisher.com/index.php/gab

Review and Progress **Open Access**

Genome-Wide Association Studies in Fabaceae: Progress and Prospects

Qishan Chen
Modern Agricultural Research Center, Cuixi Academy of Biotechnology, Zhuji, 311800, Zhejiang, China
Corresponding author: qishan.chen@cuixi.org

Genomics and Applied Biology, 2024, Vol.15, No.4 doi: 10.5376/gab.2024.15.0023
Received: 01 Jul., 2024
Accepted: 11 Aug., 2024
Published: 26 Aug., 2024

Preferred citation for this article:
Chen Q.S., 2024, Genome-wide association studies in fabaceae: progress and prospects, Genomics and Applied Biology, 15(4): 212-222 (doi: 10.5376/gab.2024.15.0023)

Abstract Recent advancements in GWAS have significantly enhanced our understanding of the genetic architecture of complex traits in Fabaceae. Key discoveries include the identification of numerous genomic variants linked to agronomic traits, such as yield, stress tolerance, and biochemical properties. The development of novel methodologies, such as mixed model frameworks and haplotype-based fine-mapping, has improved the accuracy and resolution of GWAS, reducing false positives and increasing the power to detect rare variants. Additionally, the integration of next-generation sequencing technologies has facilitated the rapid identification of candidate genes and their functional validation. The findings from GWAS in Fabaceae have profound implications for plant breeding and genetic engineering. By uncovering the genetic basis of complex traits, these studies provide valuable insights that can be leveraged to enhance crop performance and resilience. Future research should focus on optimizing GWAS models, exploring epistatic interactions, and utilizing genomic data to advance our understanding of biological processes and improve crop breeding strategies.

Keywords Genome-wide association studies; Fabaceae; Genetic variants; Complex traits; Crop improvement; Next-generation sequencing; Plant breeding

1 Introduction

The Fabaceae family, also known as Leguminosae, is the third largest family of flowering plants, comprising approximately 19 500 species across 770 genera (Xu and Deng, 2017; Jishtu and Goraya, 2020). This family is of immense agricultural importance, second only to the Poaceae family (grasses), due to its significant contribution to global food security and agricultural sustainability (Zhu et al., 2005; Smýkal et al., 2015). Fabaceae includes a wide variety of economically important crops such as soybeans, chickpeas, lentils, and peas, which are crucial sources of protein for human diets and animal feed (Smýkal et al., 2015; Jishtu and Goraya, 2020). Additionally, many Fabaceae species possess the unique ability to fix atmospheric nitrogen through symbiotic relationships with rhizobial bacteria, enhancing soil fertility and reducing the need for synthetic fertilizers (Zhao et al., 2021).

Genome-Wide Association Studies (GWAS) have emerged as a powerful tool in plant genetics and breeding, enabling the identification of genetic variants associated with important agronomic traits (Koenen et al., 2019; Shen et al., 2024). In the context of Fabaceae, GWAS can facilitate the discovery of genes linked to traits such as disease resistance, drought tolerance, and yield improvement, thereby accelerating the development of superior crop varieties (Bruneau et al., 2013; Koenen et al., 2019). The integration of GWAS with other genomic approaches, such as phylogenomics and comparative genomics, can provide deeper insights into the genetic architecture and evolutionary history of legume crops, further informing breeding strategies (Zhu et al., 2005; Koenen et al., 2019).

This study aims to provide a comprehensive overview of the progress and prospects of Genome-Wide Association Studies in the Fabaceae family. It will cover the current state of GWAS research in major legume crops, highlight key findings and methodological advancements, and discuss the potential applications of GWAS in crop improvement programs.

2 Background on GWAS in Plant Genomics

2.1 Fundamental concepts of GWAS

Genome-wide association studies (GWAS) have become a cornerstone in the field of genomics, particularly for identifying the genetic basis of phenotypic variation. The fundamental concept of GWAS involves scanning the genome for genetic variants, typically single nucleotide polymorphisms (SNPs), that are associated with specific traits. This approach leverages the natural genetic diversity within a population to identify loci that contribute to phenotypic differences (Burghardt et al., 2017; Cortes et al., 2021). The mixed model framework has been particularly influential in reducing false positives, thereby increasing the reliability of GWAS findings (Cortes et al., 2021).

2.2 Historical development of GWAS in Plants

The application of GWAS in plant genomics has evolved significantly over the past decade. Initially, the focus was on human genetics, but the methodology quickly found applications in plant science due to advances in genotyping and sequencing technologies (Visscher et al., 2017; Liu and Yan, 2018). Early plant GWAS studies were limited by the availability of genomic resources and computational tools. However, the development of high-throughput sequencing and improved statistical methods has enabled more comprehensive and accurate association studies in plants. Over 1000 GWAS have been conducted in various crop species, revealing substantial genotype-phenotype associations and providing insights into the genetic architecture of complex traits (Liu and Yan, 2018).

2.3 Technical approaches and methodologies in GWAS

Several technical approaches and methodologies have been developed to enhance the efficiency and accuracy of GWAS. The use of high-density genotyping arrays and next-generation sequencing technologies has been pivotal in capturing a wide range of genetic variants (Yang et al., 2015; Liu and Yan, 2018). Statistical methods such as the mixed model framework, which accounts for population structure and relatedness, have been crucial in reducing false positives (Burghardt et al., 2017; Cortes et al., 2021). Additionally, novel methods like extreme-phenotype GWAS (XP-GWAS) have been introduced to identify trait-associated variants by sequencing pools of individuals with extreme phenotypes, thereby reducing the need for extensive genotyping (Yang et al., 2015). These advancements have facilitated the detection of genomic variants associated with both traditional agronomic traits and molecular phenotypes (Cortes et al., 2021).

2.4 Challenges and limitations in plant GWAS

Despite the progress, several challenges and limitations persist in plant GWAS. One major issue is the complexity of plant genomes, which often include large amounts of repetitive DNA and polyploidy, complicating the identification of causal variants (Liu and Yan, 2018). Population structure and relatedness within plant populations can also lead to spurious associations, necessitating the use of sophisticated statistical models to control for these factors (Barsh et al., 2012; Burghardt et al., 2017). Additionally, the phenomenon of synthetic associations, where non-causal variants appear to be associated with traits due to linkage disequilibrium, can lead to false conclusions (Liu and Yan, 2018) Finally, the validation of GWAS findings remains a significant hurdle, as it requires extensive functional studies to confirm the biological relevance of identified variants (Barsh et al., 2012).

3 GWAS in Fabaceae: Progress and Key Findings

3.1 Overview of fabaceae and genetic diversity

The Fabaceae family, also known as the legume family, is one of the largest and most diverse plant families. This family is economically and ecologically significant, particularly due to its ability to fix atmospheric nitrogen through symbiotic relationships with rhizobial bacteria. The genetic diversity within Fabaceae is vast, with numerous polyploidization events contributing to its evolutionary history and adaptation to various environments (Zhao et al., 2021).

3.2 Major Crops studied (e.g., soybean, pea, common bean, lentil)

Several major crops within the Fabaceae family have been extensively studied using genome-wide association studies (GWAS). These include:

Soybean (*Glycine max*): Soybean is a crucial crop for its high protein and oil content. Numerous GWAS have been conducted to identify loci associated with agronomic traits, disease resistance, and seed quality (Shook et al., 2021; Kim et al., 2022; Yoosefzadeh-Najafabadi et al., 2023).

Pea (*Pisum sativum*): Field pea is another important legume crop, with GWAS identifying loci related to agronomic traits, seed morphology, and seed quality (Gali et al., 2019).

Common Bean (*Phaseolus vulgaris*): Common bean is widely cultivated for its nutritional value. GWAS in common bean have focused on yield and yield-contributing traits, leveraging the genetic diversity present in germplasm collections (Mir et al., 2021).

Lentil (*Lens culinaris*): Although not explicitly covered in the provided data, lentils are another significant legume crop studied for various agronomic traits through GWAS.

3.3 Identification of trait-associated loci in fabaceae

GWAS have been instrumental in identifying loci associated with various traits in Fabaceae crops. For instance, in soybean, significant associations have been found for traits such as days to flowering, seed coat color, and node number (Kim et al., 2022). Meta-GWAS approaches have further enhanced the detection of quantitative trait loci (QTL) by combining data from multiple studies, leading to the identification of hundreds of loci associated with traits like seed yield, plant height, and disease resistance (Shook et al., 2021). In field pea, GWAS have identified SNP markers associated with traits such as plant height, lodging resistance, and seed protein concentration (Gali et al., 2019).

3.4 Case study: GWAS for drought tolerance in soybean

Drought tolerance is a critical trait for soybean, especially in the context of climate change. A study focusing on root traits in soybean landraces identified 112 significant SNP loci associated with seven root traits, which are crucial for drought tolerance (Kim et al., 2023). This study highlights the importance of understanding the genetic basis of root traits to develop drought-resistant soybean varieties.

3.5 Success stories: breeding applications from GWAS in fabaceae

The application of GWAS findings in breeding programs has led to significant advancements in Fabaceae crop improvement. For example, the integration of GWAS and genotyping-by-sequencing (GBS) in soybean has facilitated the identification of mutation hotspots and the acceleration of genome evolution, which can be harnessed for breeding purposes (Kim et al., 2022). Additionally, the use of machine learning algorithms in GWAS has improved the detection of durable QTL associated with soybean seed quality traits, providing new tools for genomic-based breeding approaches (Yoosefzadeh-Najafabadi et al., 2023) (Figure 1). In common bean, the identification of stable and major marker-trait associations has paved the way for molecular breeding programs aimed at enhancing yield (Mir et al., 2021).

4 Advances in GWAS Technologies and Methods

4.1 High-throughput genotyping and sequencing technologies

High-throughput genotyping and sequencing technologies have revolutionized genome-wide association studies (GWAS) by enabling the rapid and cost-effective generation of vast amounts of genomic data. These technologies facilitate the identification of genetic variants associated with complex traits by providing comprehensive genomic coverage. For instance, the integration of high-throughput phenotyping with GWAS has been shown to enhance the quality of trait data, thereby improving the accuracy of genetic association studies (Xiao et al., 2021). Additionally, the development of advanced sequencing platforms has allowed for the detailed examination of genetic variants across diverse populations, furthering our understanding of the genetic basis of complex traits (Cortes et al., 2021).

4.2 Multi-trait and multi-locus GWAS approaches

Traditional GWAS methods often focus on single-marker associations, which can limit the detection of significant loci due to stringent multiple testing corrections. Multi-trait and multi-locus GWAS approaches have been developed to address these limitations by considering multiple traits and loci simultaneously. These methods increase the statistical power to detect associations and reduce false positives. For example, multi-locus GWAS methodologies such as mrMLM, ISIS EMBLASSO, and FASTmrMLM have been proposed to improve the detection of significant loci by leveraging the genetic architecture of complex traits (Zhang et al., 2019). Moreover, joint-GWAS approaches, which combine data from multiple studies, have been shown to enhance the power of GWAS by increasing sample sizes and capturing more genetic variation (Müller et al., 2018).

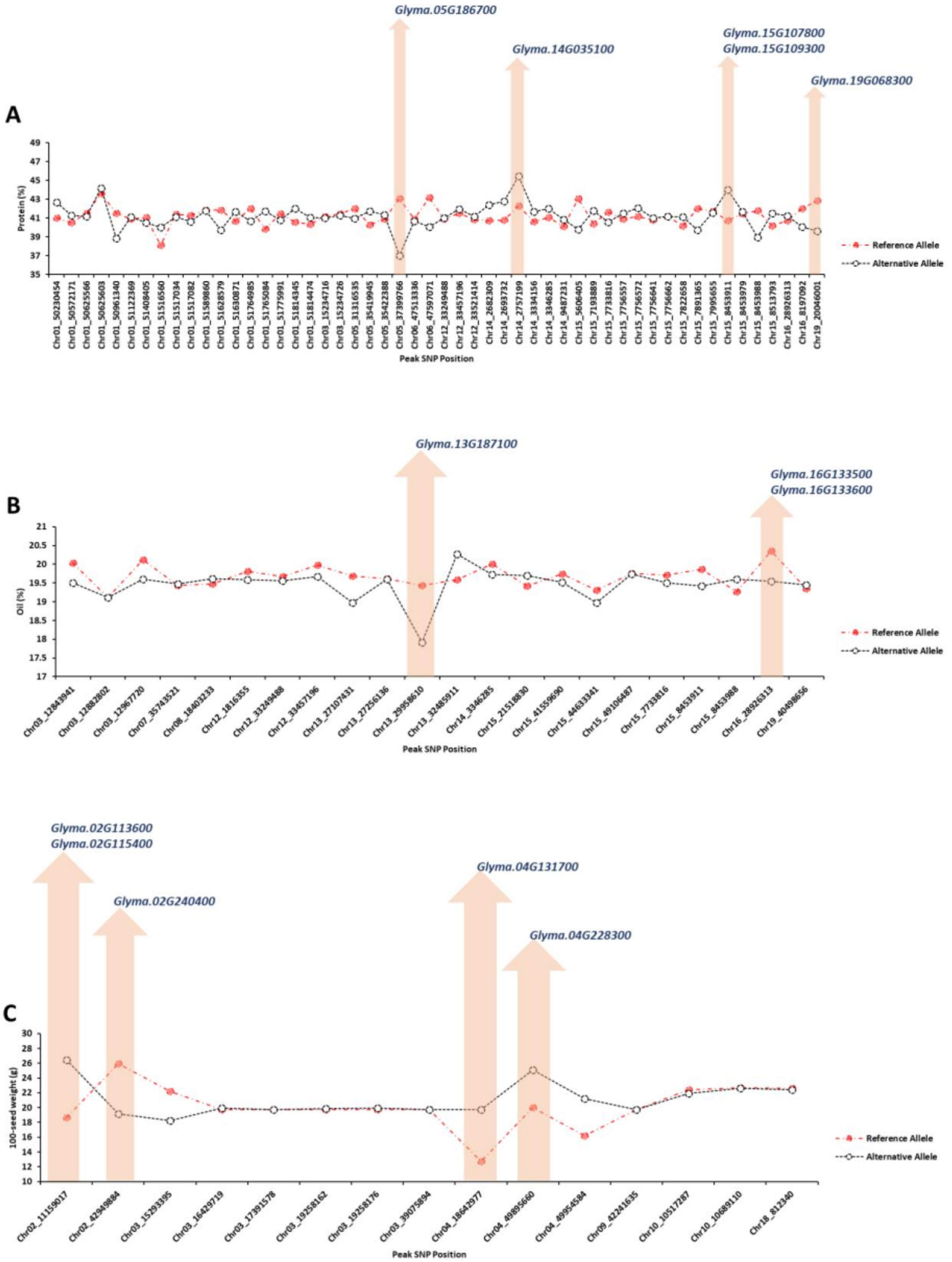

Figure 1 GWAS results of the top significant SNPs associated with soybean seed protein (A), oil (B), and 100-seed weight (C) across different environments. The y-axis represents the value of the trait of interest, and the x-axis represents the genomic position of each SNP on the soybean genome (Adopted from Yoosefzadeh-Najafabadi et al., 2023)

4.3 Integration of GWAS with other omics data (e.g., transcriptomics, metabolomics)

The integration of GWAS with other omics data, such as transcriptomics and metabolomics, provides a more comprehensive understanding of the genetic basis of complex traits. This multi-omics approach allows researchers to link genetic variants to gene expression and metabolic pathways, thereby uncovering the molecular mechanisms underlying trait variation. For instance, multi-omics strategies for transcriptome-wide association studies (TWAS) have been developed to incorporate distal-SNPs and regulatory biomarkers, improving the power to detect gene-trait associations (Bhattacharya et al., 2020). Additionally, integrative analyses combining genomics, transcriptomics, and metabolomics have been used to identify candidate genes for traits such as metabolic and vegetative growth variation in canola (Knoch et al., 2023).

4.4 Case study: multi-omics integration for disease resistance in common bean

A case study on the integration of multi-omics data for disease resistance in common bean illustrates the potential of this approach. By combining GWAS with transcriptomic and metabolomic data, researchers can identify candidate genes and pathways involved in disease resistance. This integrative approach enables the discovery of genetic variants that may not be detectable through GWAS alone. For example, the integration of high-throughput phenotyping and multi-omics data has been shown to enhance the identification of genetic loci associated with complex traits, providing valuable insights for crop breeding and genetic engineering (Xiao et al., 2021; Knoch et al., 2023). This case study highlights the importance of multi-omics integration in advancing our understanding of the genetic basis of disease resistance and other complex traits in plants.

5 Prospects and Future Directions in Fabaceae GWAS

5.1 Emerging trends in GWAS for fabaceae

Genome-wide association studies (GWAS) have become a cornerstone in plant genetics, enabling the dissection of complex traits through the identification of genetic variants associated with phenotypic traits. Recent advancements in genomic technologies, such as next-generation sequencing and high-throughput genotyping, have significantly enhanced the power and resolution of GWAS. These technologies have facilitated the detection of both common and rare variants, contributing to a more comprehensive understanding of the genetic architecture of traits in Fabaceae (Liu and Yan, 2018; Cortes et al., 2021). Additionally, the integration of machine learning methods for dimensionality reduction and the development of sophisticated statistical models have further improved the accuracy and efficiency of GWAS (Susmitha et al., 2023).

5.2 Potential for GWAS in neglected and underutilized fabaceae species

While major legume crops like soybean and chickpea have benefited from extensive genomic research, many underutilized and minor legumes remain largely unexplored. These species, including adzuki bean, cluster bean, and winged bean, hold significant potential for enhancing food security and agricultural sustainability, particularly in marginal environments. Recent progress in genomics, such as the completion of genome assemblies and the development of genomic resources, has opened new avenues for GWAS in these neglected species. By leveraging these advancements, researchers can identify key genetic variants that contribute to desirable traits, thereby facilitating the development of improved cultivars (Jha et al., 2022).

5.3 Role of GWAS in climate-resilient crop development

Climate change poses a significant threat to global agriculture, necessitating the development of climate-resilient crops. GWAS plays a crucial role in this endeavor by identifying genetic variants associated with traits that enhance resilience to abiotic stresses such as drought, heat, and salinity. For instance, GWAS has been instrumental in uncovering loci linked to drought tolerance and heat resistance in various crops. By applying these insights to Fabaceae, researchers can develop legume varieties that are better equipped to withstand changing climatic conditions, thereby ensuring stable yields and food security (Susmitha et al., 2023; Uprety, 2023).

5.4 Case study: GWAS for nutritional traits in soybean

Soybeans are a vital source of protein and essential nutrients, making them an important crop for addressing nutritional deficiencies. Recent GWAS efforts have focused on identifying genetic variants associated with nutritional traits in soybeans, such as protein content, amino acid composition, and micronutrient levels. These studies have revealed several loci that significantly influence these traits, providing valuable markers for breeding programs aimed at enhancing the nutritional quality of soybeans. The integration of these markers into genomic selection frameworks can accelerate the development of nutritionally superior soybean varieties (Susmitha et al., 2023) (Figure 2).

5.5 Collaborative initiatives and global research efforts

The success of GWAS in Fabaceae is greatly enhanced by collaborative initiatives and global research efforts. Platforms like the GWAS Atlas provide curated resources of genome-wide variant-trait associations, facilitating data sharing and collaboration among researchers worldwide (Tian et al., 2019). Additionally, international consortia and breeding programs are increasingly adopting joint-GWAS approaches, which combine data from multiple studies to increase statistical power and validate findings across diverse populations (Müller et al., 2018). These collaborative efforts are essential for advancing our understanding of the genetic basis of complex traits and for translating GWAS findings into practical breeding applications.

6 Challenges and Considerations for Future GWAS Studies

6.1 Data quality and reproducibility issues

Data quality and reproducibility are critical challenges in genome-wide association studies (GWAS). Ensuring high-quality data involves rigorous quality control measures to filter out errors and artifacts that can lead to false associations. For instance, the tutorial by Marees et al. (2018) emphasizes the importance of quality control and statistical analysis in GWAS, highlighting the need for dedicated genetics software to manage these tasks effectively. Additionally, reproducibility of results is a significant concern, as demonstrated by the MultiGWAS tool, which facilitates the replication of GWAS by testing different parameters and models to validate results (Garreta et al., 2020). This approach helps in identifying false associations and ensuring that findings are robust and reliable.

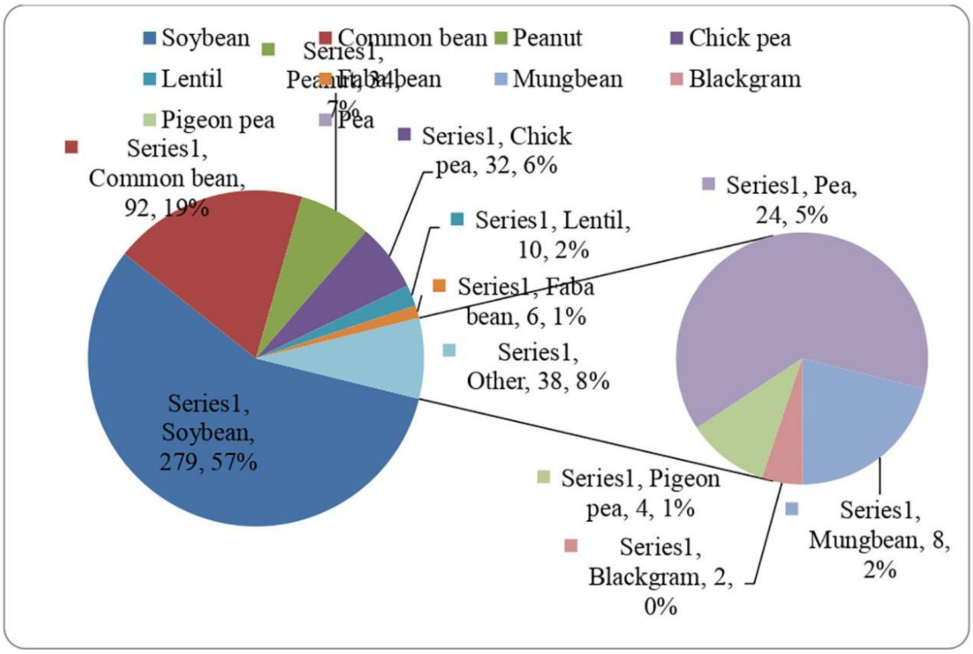

Figure 2 Number of GWAS studies conducted in different leguminous crops from timeline 2012 to 2023 (Adopted from Susmitha et al., 2023)

6.2 Statistical challenges and advances in GWAS analysis

Statistical challenges in GWAS include managing the vast amount of data and controlling for population structure to avoid spurious associations. Advances in statistical methods have been crucial in addressing these issues. For example, the development of mixed model frameworks has significantly reduced false positives compared to naïve methods (Cortes et al., 2021). Moreover, the use of joint-GWAS approaches, as seen in the study on Eucalyptus, has increased the power to detect significant associations by combining data from multiple populations (Müller et al., 2018). These advancements are essential for improving the accuracy and reliability of GWAS findings.

6.3 Ethical considerations and intellectual property rights

Ethical considerations in GWAS encompass issues related to data privacy, informed consent, and the equitable representation of diverse populations. The importance of including ancestrally diverse populations in GWAS is highlighted by Peterson et al. (2019), which discusses the scientific and ethical imperatives of broadening the genetic research base. Additionally, intellectual property rights pose challenges in terms of data sharing and the commercialization of genetic findings. The NHGRI-EBI GWAS Catalog provides a high-quality curated collection of published GWAS, promoting transparency and accessibility while addressing some of these ethical concerns (Buniello et al., 2018).

6.4 Case study: overcoming population structure in GWAS of fabaceae crops

Population structure can confound GWAS results, leading to false associations. Overcoming this challenge requires sophisticated statistical methods and careful study design. The study on Eucalyptus breeding populations serves as a relevant case study, where joint-GWAS approaches were used to combine data from different populations, thereby increasing the power to detect true associations (Müller et al., 2018). This method can be applied to Fabaceae crops to address population structure issues and improve the accuracy of GWAS findings. By leveraging joint-GWAS and other advanced statistical techniques, researchers can better understand the genetic architecture of Fabaceae crops and enhance breeding programs.

7 Applications of GWAS in Breeding Programs

7.1 Incorporating GWAS findings into marker-assisted selection

Genome-wide association studies (GWAS) have become a cornerstone in modern plant breeding, particularly for marker-assisted selection (MAS). By identifying single nucleotide polymorphisms (SNPs) associated with desirable traits, GWAS enables breeders to select plants with favorable genetic profiles more efficiently. For instance, in soybean, SNP markers associated with yield, maturity, plant height, and seed weight have been identified, facilitating their use in MAS to improve these traits (Ravelombola et al., 2021). Similarly, in field pea, SNP markers linked to agronomic traits such as plant height, lodging resistance, and grain yield have been identified, which can be utilized in MAS to accelerate cultivar improvement (Gali et al., 2019).

7.2 Strategies for accelerating genetic gain through GWAS

To maximize genetic gain, breeders can integrate GWAS findings with genomic selection (GS). This approach leverages the predictive power of SNP markers to estimate the breeding values of individuals, thereby accelerating the selection process. In soybean, the combination of GWAS and GS has shown high accuracy in selecting for yield and other agronomic traits, demonstrating the potential for rapid genetic improvement (Ravelombola et al., 2021). Additionally, the use of advanced statistical models such as ridge regression best linear unbiased prediction (rrBLUP) and Bayesian methods can further enhance the accuracy of GS, as seen in legume breeding programs (Susmitha et al., 2023).

7.3 Case study: GWAS-driven breeding for yield improvement in pea

A notable example of GWAS application in breeding is the improvement of yield in field pea (*Pisum sativum* L.). A comprehensive GWAS involving 135 pea accessions from diverse breeding programs identified several SNP markers associated with key agronomic traits, including grain yield. These markers were consistent across multiple trials and locations, underscoring their reliability for use in breeding programs (Gali et al., 2019) (Figure 3). By incorporating these markers into MAS, breeders can more effectively select high-yielding pea varieties, thereby enhancing overall productivity.

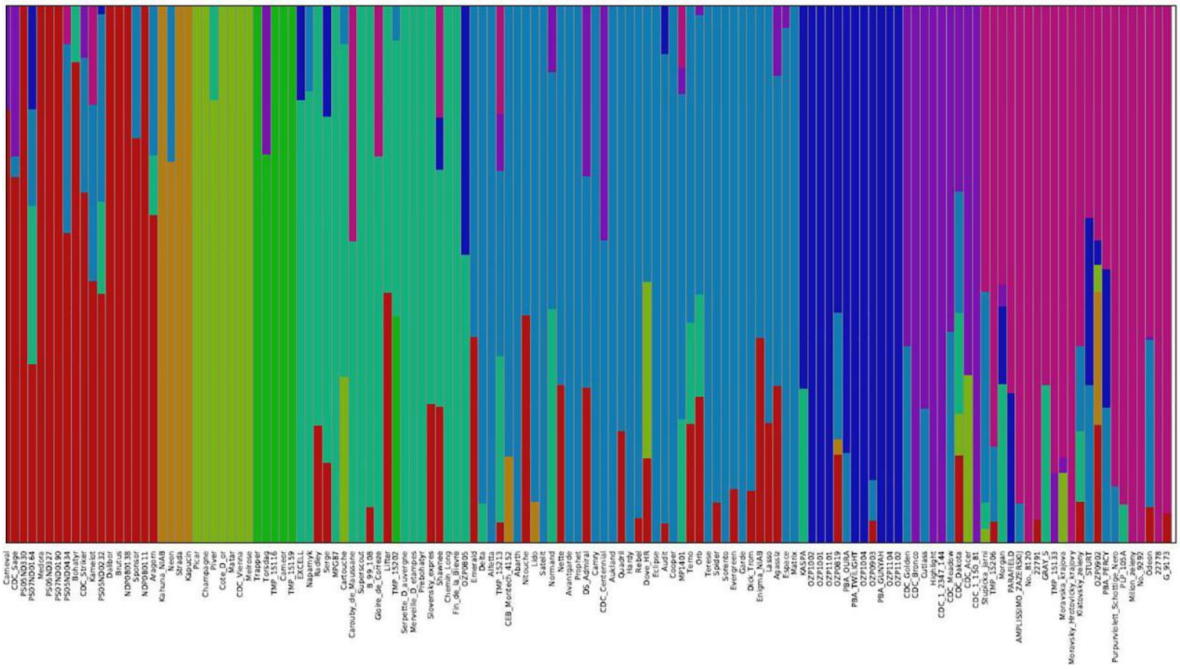

Figure 3 Population structure of 135 pea accessions based on K = 9. In the panel, each accession is indicated as a vertical bar partitioned into colored segments where the respective length of these segments represents the proportion of the individual's genome in a given group (Adopted from Gali et al., 2019)

7.4 Future prospects in the integration of GWAS with modern breeding technologies

The future of GWAS in breeding programs lies in its integration with cutting-edge technologies such as high-throughput genotyping, phenomics, and bioinformatics tools. The development of meta-GWAS approaches, which combine data from multiple years and environments, offers a robust method for identifying marker-trait associations in unbalanced datasets, as demonstrated in wheat breeding programs (Battenfield et al., 2018). Furthermore, the use of machine learning techniques for dimensionality reduction and the application of advanced statistical models can significantly improve the efficiency and accuracy of GWAS, paving the way for more precise and rapid genetic improvements in crops (Susmitha et al., 2023).

8 Concluding Remarks

Genome-wide association studies (GWAS) have significantly advanced our understanding of the genetic architecture of complex traits in Fabaceae. The application of GWAS in various Fabaceae species, such as common bean (*Phaseolus vulgaris* L.) and field pea (*Pisum sativum* L.), has led to the identification of numerous loci associated with important agronomic traits. For instance, in common bean, significant SNP markers associated with phenology, biomass, yield components, and seed yield traits were identified, providing insights into the genetic basis of these traits. Similarly, in field pea, GWAS identified SNP markers linked to agronomic traits such as days to flowering, plant height, and seed yield, as well as seed quality traits like protein and starch concentrations. These studies highlight the power of GWAS in uncovering the genetic determinants of key traits in Fabaceae crops.

The findings from GWAS in Fabaceae have several implications for future research and breeding programs. Firstly, the identification of trait-associated loci provides valuable markers for marker-assisted selection (MAS), which can accelerate the breeding of improved varieties with desirable traits. For example, the significant SNPs identified in common bean and field pea can be used to develop molecular markers for traits such as yield and seed quality, facilitating the selection of superior genotypes in breeding programs. Additionally, the integration of GWAS with other genomic tools, such as genomic selection and gene editing, holds promise for further enhancing the efficiency and precision of crop improvement efforts. Future research should also focus on validating the identified loci and understanding their functional roles in trait expression, which will provide deeper insights into the molecular mechanisms underlying complex traits in Fabaceae.

GWAS has emerged as a powerful tool for dissecting the genetic basis of complex traits in Fabaceae, contributing to significant advancements in crop improvement. The ability to identify and utilize genetic variants associated with important agronomic traits has the potential to revolutionize breeding programs, leading to the development of high-yielding, resilient, and nutritionally enhanced Fabaceae crops. However, challenges such as population structure, linkage disequilibrium, and the need for large sample sizes must be addressed to fully harness the potential of GWAS. Continued advancements in genomic technologies and statistical methods will undoubtedly enhance the resolution and accuracy of GWAS, paving the way for more effective and sustainable crop improvement strategies in Fabaceae.

Acknowledgments

The author appreciates the feedback from two anonymous peer reviewers on the manuscript of this study, whose careful evaluation and constructive suggestions have contributed to the improvement of the manuscript.

Conflict of Interest Disclosure

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

References

Barsh G., Copenhaver G., Gibson G., and Williams S., 2012, Guidelines for genome-wide association studies, PLoS Genetics, 8(7): e1002812.
https://doi.org/10.1371/journal.pgen.1002812

Battenfield S., Sheridan J., Silva L., Miclaus K., Dreisigacker S., Wolfinger R., Peña R., Singh R., Jackson E., Fritz A., Guzmán C., and Poland J., 2018, Breeding-assisted genomics: Applying meta-GWAS for milling and baking quality in CIMMYT wheat breeding program, PLoS ONE, 13(11): e0204757.
https://doi.org/10.1371/journal.pone.0204757

Bhattacharya A., Li Y., and Love M., 2020, MOSTWAS: multi-omic strategies for transcriptome-wide association studies, PLoS Genetics, 17(3): e1009398.
https://doi.org/10.1371/journal.pgen.1009398

Bruneau A., Doyle J., Herendeen P., Hughes C., Kenicer G., Lewis G., Mackinder B., Pennington R., Sanderson M., Wojciechowski M., Boatwright S., Brown G., Cardoso D., Crisp M., Egan A., Fortunato R., Hawkins J., Kajita T., Klitgaard B., Koenen E., Lavin M., Luckow M., Marazzi B., McMahon M., Miller J., Murphy D., Ohashi H., Queiroz L., Rico L., Särkinen T., Schrire B., Simon M., Souza É., Steele K., Torke B., Wieringa J., and Wyk B., 2013, Legume phylogeny and classification in the 21st century: Progress, prospects and lessons for other species-rich clades, Taxon, 62: 217-248.
https://doi.org/10.12705/622.8

Buniello A., MacArthur J., Cerezo M., Harris L., Hayhurst J., Malangone C., McMahon A., Morales J., Mountjoy E., Sollis E., Suveges D., Vrousgou O., Whetzel P., Amode R., Guillen J., Riat H., Trevanion S., Hall P., Junkins H., Flicek P., Burdett T., Hindorff L., Cunningham F., and Parkinson H., 2018, The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019, Nucleic Acids Research, 47: D1005-D1012.
https://doi.org/10.1093/nar/gky1120

Burghardt L., Young N., and Tiffin P., 2017, A guide to genome-wide association mapping in plants, Current Protocols in Plant Biology, 2(1): 22-38.
https://doi.org/10.1002/cppb.20041

Cortes L., Zhang Z., and Yu J., 2021, Status and prospects of genome-wide association studies in plants, The Plant Genome, 14(1): e20077.
https://doi.org/10.1002/tpg2.20077

Gali K., Sackville A., Tafesse E., Lachagari V., Mcphee K., Hybl M., Mikić A., Smýkal P., McGee R., Burstin J., Domoney C., Ellis T., Tar'an B., and Warkentin T., 2019, Genome-wide association mapping for agronomic and seed quality traits of field pea (*Pisum sativum* L.), Frontiers in Plant Science, 10: 1538.
https://doi.org/10.3389/fpls.2019.01538

Garreta L., Cerón-Souza I., Palacio M., and Reyes-Herrera P., 2020, MultiGWAS: an integrative tool for genome wide association studies in tetraploid organisms, Ecology and Evolution, 11: 7411-7426.
https://doi.org/10.1002/ece3.7572

Jha U., Nayyar H., Parida S., Bakır 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

Jishtu V., and Goraya G., 2020, Leguminosae (nom. alt. Fabaceae)-its diversity, use and role in environmental conservation in the Harsh environs of the cold deserts of north-west India, In: Hasanuzzaman, M., Araújo S., Gill S. (eds), The Plant Family Fabaceae, Springer, Singapore, pp.261-285.
https://doi.org/10.1007/978-981-15-4752-2_10

Kim D., Lyu J., Kim J., Seo J., Choi H., Jo Y., Kim S., Eom S., Ahn J., Bae C., and Kwon S., 2022, Identification of loci governing agronomic traits and mutation hotspots via a GBS-based genome-wide association study in a soybean mutant diversity pool, International Journal of Molecular Sciences, 23(18): 10441.
https://doi.org/10.3390/ijms231810441

Kim S., Tayade R., Kang B., Hahn B., Ha B., and Kim Y., 2023, Genome-wide association studies of seven root traits in soybean (*Glycine max* L.) landraces, International Journal of Molecular Sciences, 24(1): 873.
https://doi.org/10.3390/ijms24010873

Knoch D., Meyer R., Heuermann M., Riewe D., Peleke F., Szymański J., Abbadi A., Snowdon R., and Altmann T., 2023, Integrated multi-omics analyses and genome-wide association studies reveal prime candidate genes of metabolic and vegetative growth variation in canola, The Plant Journal : for Cell and Molecular Biology, 117(3): 713-728.
https://doi.org/10.1111/tpj.16524

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

Liu H., and Yan J., 2018, Crop genome-wide association study: a harvest of biological relevance, The Plant Journal, 97: 8-18.
https://doi.org/10.1111/tpj.14139

Marees A., Kluiver H., Stringer S., Vorspan F., Curis E., Marie-Claire C., and Derks E., 2018, A tutorial on conducting genome-wide association studies: Quality control and statistical analysis, International Journal of Methods in Psychiatric Research, 27(2): e1608.
https://doi.org/10.1002/mpr.1608

Mir R., Choudhary N., Bawa V., Jan S., Singh B., Bhat M., Paliwal R., Kumar A., Chitikineni A., Thudi M., and Varshney R., 2021, Allelic diversity, structural analysis, and genome-wide association study (GWAS) for yield and related traits using unexplored common bean (*Phaseolus vulgaris* L.) germplasm from western Himalayas, Frontiers in Genetics, 11: 609603.
https://doi.org/10.3389/fgene.2020.609603

Müller B., Filho J., Lima B., Garcia C., Missiaggia A., Aguiar A., Takahashi E., Kirst M., Gezan S., Silva-Junior O., Neves L., and Grattapaglia D., 2018, Independent and joint-GWAS for growth traits in Eucalyptus by assembling genome-wide data for 3373 individuals across four breeding populations, The New Phytologist, 221(2): 818-833.
https://doi.org/10.1111/nph.15449

Peterson R., Kuchenbaecker K., Walters R., Chen C., Popejoy A., Periyasamy S., Lam M., Iyegbe C., Strawbridge R., Brick L., Carey C., Martin A., Meyers J., Su J., Chen J., Edwards A., Kalungi A., Koen N., Majara L., Schwarz E., Smoller J., Stahl E., Sullivan P., Vassos E., Mowry B., Prieto M., Cuellar-Barboza A., Bigdeli T., Edenberg H., Huang H., and Duncan L., 2019, Genome-wide association studies in ancestrally diverse populations: opportunities, methods, pitfalls, and recommendations, Cell, 179: 589-603.
https://doi.org/10.1016/j.cell.2019.08.051

Ravelombola W., Qin J., Shi A., Song Q., Yuan J., Wang F., Chen P., Yan L., Feng Y., Zhao T., Meng Y., Guan K., Yang C., and Zhang M., 2021, Genome-wide association study and genomic selection for yield and related traits in soybean, PLoS ONE, 16(8): e0255761.
https://doi.org/10.1371/journal.pone.0255761

Shen W.L., Huang Y.P., Chen R.C., and Lin H.M., 2024, Impact of whole genome duplication events on the diversification of legumes, Legume Genomics and Genetics, 15(1): 37-45.
https://doi.org/10.5376/lgg.2024.15.0005

Shook J., Zhang J., Jones S., Singh A., Diers B., and Singh A., 2021, Meta-GWAS for quantitative trait loci identification in soybean, G3: Genes|Genomes|Genetics, 11(7): jkab117.
https://doi.org/10.1093/g3journal/jkab117

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

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

Tian D., Wang P., Tang B., Teng X., Li C., Liu X., Zou D., Song S., and Zhang Z., 2019, GWAS Atlas: a curated resource of genome-wide variant-trait associations in plants and animals, Nucleic Acids Research, 48: D927-D932.
https://doi.org/10.1093/nar/gkz828

Uprety S., 2023, A prospect for genome wide association studies in crops, International Journal of Applied Sciences and Biotechnology,
https://doi.org/10.3126/ijasbt.v11i2.56114

Visscher P., Wray N., Zhang Q., Sklar P., McCarthy M., Brown M., and Yang J., 2017, 10 years of GWAS discovery: biology, function, and translation, American journal of human genetics, 101(1): 5-22.
https://doi.org/10.1016/j.ajhg.2017.06.005

Xiao Q., Bai X., Zhang C., and He Y., 2021, Advanced high-throughput plant phenotyping techniques for genome-wide association studies: A review, Journal of Advanced Research, 35: 215-230.
https://doi.org/10.1016/j.jare.2021.05.002

Xu Z., and Deng M., 2017, Fabaceae or Leguminosae, In: Identification and Control of Common Weeds: Volume 2. Springer, Dordrecht, pp.547-615.
https://doi.org/10.1007/978-94-024-1157-7_43

Yang J., Jiang H., Yeh C., Yu J., Jeddeloh J., Nettleton D., and Schnable P., 2015, Extreme-phenotype genome-wide association study (XP-GWAS): a method for identifying trait-associated variants by sequencing pools of individuals selected from a diversity panel, The Plant Journal : for Cell and Molecular Biology, 84(3): 587-596.
https://doi.org/10.1111/tpj.13029

Yoosefzadeh-Najafabadi M., Torabi S., Tulpan D., Rajcan I., and Eskandari M., 2023, Application of SVR-mediated GWAS for identification of durable genetic regions associated with soybean seed quality traits, Plants, 12(14): 2659.
https://doi.org/10.3390/plants12142659

Zhang Y., Jia Z., and Dunwell J., 2019, Editorial: the applications of new multi-locus GWAS methodologies in the genetic dissection of complex traits, Frontiers in Plant Science, 10: 100.
https://doi.org/10.3389/fpls.2019.00100

Zhao Y., Zhang R., Jiang K., Qi J., Hu Y., Guo J., Zhu R., Zhang T., Egan A., Yi T., Huang C., and Ma H., 2021, Nuclear phylotranscriptomics/phylogenomics support numerous polyploidization events and hypotheses for the evolution of rhizobial nitrogen-fixing symbiosis in Fabaceae, Molecular Plant, 14(5): 748-773.
https://doi.org/10.1016/j.molp.2021.02.006

Zhu H., Choi H., Cook D., and Shoemaker R., 2005, Bridging model and crop legumes through comparative genomics, Plant Physiology, 137: 1189-1196.
https://doi.org/10.1104/pp.104.058891