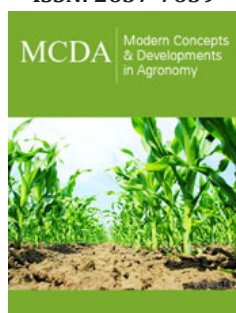


Application of Association Mapping for the Identification of Candidate Genes Linked to Agronomically Important Traits as a Modern Development Concept in Agronomy

ISSN: 2637-7659



***Corresponding author:** Jan Bocianowski, Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Wojska Polskiego 28, 60-637 Poznań, Poland

Submission:  August 28, 2025

Published:  October 22, 2025

Volume 15 - Issue 3

How to cite this article: Jan Bocianowski* and Agnieszka Leśniewska-Bocianowska. Application of Association Mapping for the Identification of Candidate Genes Linked to Agronomically Important Traits as a Modern Development Concept in Agronomy. Mod Concep Dev Agrono. 15(3). MCDA. 000865. 2025.
DOI: [10.31031/MCDA.2025.15.000865](https://doi.org/10.31031/MCDA.2025.15.000865)

Copyright@ Jan Bocianowski, This article is distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use and redistribution provided that the original author and source are credited.

Jan Bocianowski^{1*} and Agnieszka Leśniewska-Bocianowska²

¹Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Poland

²Department of Pathophysiology of Ageing and Civilization Diseases, Poznan University of Medical Sciences, Poland

Abstract

Unpredictable climate change in recent decades has led to soil depletion, among other things. Furthermore, global food demand is increasing. This poses challenges for agriculture, which must quickly respond to numerous simultaneous changes. Increasing yields and improving their quality are issues we constantly grapple with. Especially in recent decades, food security has become increasingly important. Integrating traditional agronomic knowledge with the latest discoveries in molecular biology can help solve these growing problems. Association mapping is an increasingly used to identify candidate genes that most significantly determine quantitative traits. It combines the potential of large-scale genetic datasets with the precise analysis of phenotypic traits. The aim of this publication is to present association mapping as a modern and versatile research concept that supports the development of agronomy, with particular emphasis on its application in identifying candidate genes associated with traits essential in agriculture.

Keywords: Association mapping; Genome-wide association studies; K-mer GWAS; Multi-omics strategies; Linkage-disequilibrium; Molecular markers

Introduction

Never in human history has agriculture been required to respond so rapidly to such a multitude of simultaneous changes. Within just a few decades, the climate has become more unpredictable, soil has grown increasingly depleted and global demand for food has risen to unprecedented levels. In a world where every lost ton of yield can carry tangible consequences for food security, those who can integrate traditional agronomic knowledge with the latest discoveries in molecular biology gain a distinct advantage. Association mapping, which combines the potential of large-scale genetic datasets with precise phenotypic trait analysis, is emerging as one of the most potent tools driving this transformation [1-3]. Twenty-first-century agronomy faces unprecedented challenges. The growing global population, environmental degradation, climate change and limited natural resources necessitate the development of new, more precise strategies for plant breeding. In this context, the rapid and reliable identification of genes controlling traits of key agricultural importance-such as yield performance, pathogen resistance, tolerance to abiotic stresses and nutrient use efficiency-has become particularly crucial [4,5]. Traditional breeding methods, although effective in the past, are increasingly proving to be too time-consuming and limited in terms of genetic resolution when compared with the capabilities of modern genome analysis tools [6]. Association

mapping, also referred to as Genome-Wide Association Studies (GWAS), represents a breakthrough in research on the genetic basis of phenotypic variation [3].

This method relies on the analysis of natural genetic diversity within populations that have not been designed explicitly for linkage mapping studies. By simultaneously utilizing information from hundreds of thousands of molecular markers, GWAS enables the precise identification of genomic regions associated with traits of interest, thereby facilitating the detection of candidate genes. Unlike classical linkage studies, which are limited by the relatively small number of recombination events in controlled populations, association mapping capitalizes on recombination accumulated over many generations, resulting in substantially higher mapping resolution [3]. The significance of GWAS extends beyond the mere identification of markers-the results obtained can be directly integrated into Marker-Assisted Selection (MAS) and Genomic Selection (GS). This integration shortens the breeding process and enhances the efficiency of developing crop varieties adapted to extreme environmental conditions or emerging pathogens. Association mapping has been successfully applied across a wide range of crop species-from cereals [7-16] and oilseed crops [17,18] to vegetables [19,20] and forage plants [21,22] and its potential in the advancement of agronomy continues to grow through integration with next-generation sequencing technologies, transcriptome analysis, and genome editing tools such as CRISPR/Cas. The aim of this publication is to present association mapping as a modern and versatile research concept that supports the advancement of agronomy, with particular emphasis on its application in the identification of candidate genes associated with agriculturally essential traits. Both the theoretical foundations of the method and its practical implementation in breeding programs will be discussed, with consideration of future directions for development and integration with other technologies in precision agriculture.

The role and significance of association mapping (GWAS) in agronomy

Over the past two decades, association mapping (GWAS) has become a central tool for identifying genetic variants associated with complex traits in crop plants. In comparison with classical linkage mapping based on bi-parental populations, GWAS leverages historical recombination events within naturally diverse populations, thereby enhancing the resolution of signal localization and accelerating the identification of candidate genes.

Methodological reviews emphasize both the potential of GWAS for uncovering genes influencing agronomic traits and the necessity of employing advanced statistical models to control type I errors [23,24].

Methodological advances

The development of statistical methods and algorithms has significantly enhanced the practical utility of GWAS in plant breeding. The introduction of Linear Mixed Models (LMMs) and their subsequent optimization reduced the number of false associations caused by population structure. Further innovations, such as multi-locus approaches (e.g., mrMLM and other implementations) and hybrid models combining fixed and random effects (e.g., FarmCPU), have improved the sensitivity of detecting loci with smaller effects while simultaneously reducing errors and computational costs. Comparative studies indicate that methods such as FarmCPU and Blink often outperform classical single-locus LMMs in a variety of scenarios [25-27].

Novel approaches: K-mer GWAS, multi-omics strategies and machine learning

Traditional SNP-based GWAS are increasingly being complemented by k-mer-based approaches, which enable the detection of structural variants and sequences absent from the reference genome. This is particularly advantageous in species with complex genomes and high levels of diversity. Simultaneously, the integration of transcriptomic, methylomic, and functional data (multi-omics), along with the application of machine learning methods for feature selection and marker effect prediction, substantially enhances the power of detection and the biological interpretability of GWAS results [28-30]. The MLM/LMM framework remains the “gold standard” for population control and should be employed as a reference point-particularly in analyses where population structure is pronounced [31]. FarmCPU enhances detection power through the iterative partitioning of fixed and random effects, making it a robust solution for polygenic traits; however, it requires careful parameter tuning [25]. BLINK frequently outperforms FarmCPU in terms of statistical power and the number of true positives, while maintaining computational efficiency--thus, it represents a valuable option for large SNP panels (Table 1). The mrMLM approach, together with the suite of multi-locus methods, is highly advantageous when the objective is to detect many QTNs; the integration of multiple algorithms further strengthens the validation of results [32,33].

Table 1: Comparison of selected GWAS methods.

Method	Type	Key Traits	Advantages	Limitations	Computational Cost	Typical Applications / Remarks
MLM / LMM (e.g. EMMA, FaST-LMM, GEMMA)	Single locus (with random correction)	Mixed model accounting for population structure and kinship (kinship matrix). A well-documented implementation is provided by GEMMA/EMMA [31].	Effective control of false positives associated with population structure; widely used and thoroughly validated [31].	Reduced power to detect multiple loci with small effects; single-locus nature makes it challenging to capture polygenic signals without large sample sizes.	Moderate to high (large kinship matrices, optimization possible through Fast-LMM).	Useful as a starting point; frequently employed for population control and for comparison with multilogues methods.

Farm CPU (Fixed and random model Circulating Probability Unification)	Multi-locus (iterative).	Iterative separation of fixed and random effects; iteratively selects “associated” markers as covariates. Reduces confounding between the tested marker and the covariate [25].	Higher power to detect loci with moderate or small effects compared to classical MLM; reduction of false negatives [25].	May be sensitive to parameter selection; in certain scenarios (very strong LD/small N), there is a risk of overfitting.	Moderate-typically faster than full MLM for large panels (implemented in R/GAPIT).	Well-suited for polygenic traits with a moderate sample size; commonly applied in crop species.
BLINK (Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway)	Multi-locus (optimization BLUP/Farm CPU)	Utilizes LD information and the Bayesian Information Criterion instead of computationally expensive REML steps; an improved version of the Farm CPU concept. In empirical tests, it often demonstrated higher power and fewer false positives compared to Farm CPU [26].	Improved detection power and reduced false discoveries in multiple simulations and real datasets; computationally more efficient [26].	A more advanced method-requiring an understanding of LD settings and model criteria-does not consistently outperform other multi-locus approaches in all scenarios.	Low to moderate-optimized and faster than conventional REML.	When prioritizing power and computational efficiency with large datasets, this represents a suitable choice.
mrMLM and other methods within the mrMLM package (e.g., FASTmrMLM, FASTmrEMMA, pLARM EB, etc.).	Multi-locus (multiple implementations).	A collection of several multi-locus approaches, often treating marker effects as random; two-step procedures for QTN detection followed by subsequent effect estimation [32].	High power for detecting multiple QTNs; the package integrates diverse algorithms, facilitating comparative analyses [33].	Risk of overfitting when the number of candidate variables greatly exceeds the number of samples; certain algorithms may require careful parameter tuning [33].	Moderate (dependent on the specific method within the package).	Useful when aiming to compare the results of multiple multi-locus algorithms within a single workflow; well-suited for the discovery of multiple QTNs.
Single-locus GLM (without correction)	Single-locus	Simple marker → phenotype regression, sometimes with the addition of PCs as covariates.	Low computational complexity; suitable for fast, exploratory analyses.	High susceptibility to false positives due to population structure and kinship.	Short	For initial screening/ candidate discovery only; not recommended as the sole source of results.
k-mer-based GWAS	Alignment-free / multi-locus extension	Analysis of k-mers from raw sequences allows the detection of structural variants and sequences absent from the reference genome.	It enables finding variants not visible in SNP-only GWAS (e.g. structural, indels, new sequences) [29,34].	More computationally demanding; biological interpretation (k-mer → gene mapping) may be more difficult.	High	Particularly useful in species with unstable/pan genomic genomes, or when using raw sequencing data.

Finally, k-mer based GWAS and multi-omics approaches (e.g., integration with transcriptomics and eQTL mapping) are gaining increasing relevance, particularly in studies focusing on structural variants or the functional validation of candidate loci (Table 2) [29,34]. The MLM framework demonstrates robustness in maintaining a low False Discovery Rate (FDR), yet it exhibits reduced effectiveness in detecting polygenic effects (Table 3, Figure 1) [35]. Multi-locus approaches (e.g., FarmCPU, mrMLM) represent the optimal compromise between detection power and the control

of false signals (Table 3, Figure 1). BLINK is distinguished by its superior computational efficiency and a highly favorable balance between detection power and FDR, as evidenced by comparative benchmarking studies [36]. E-GWAS, one of the most recent models, has been shown through simulation to achieve the highest detection power while simultaneously maintaining the lowest False Positive Rate (FPR) in comparison to other commonly used methodologies [37].

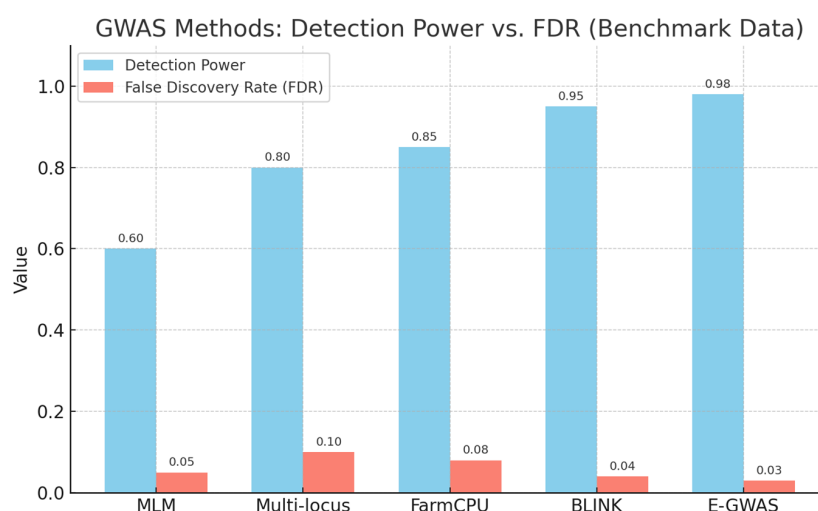
Table 2: Examples of parameters accompanying individual methods.

Method	Example Parameters*
MLM/LMM	MAF $\geq$ 0.05, PCA 3-5, kinship with pruned SNP, Bonferroni/FDR
FarmCPU	MAF $\geq$ 0.05, input p-value e.g. e-3, number of iterations 3-5

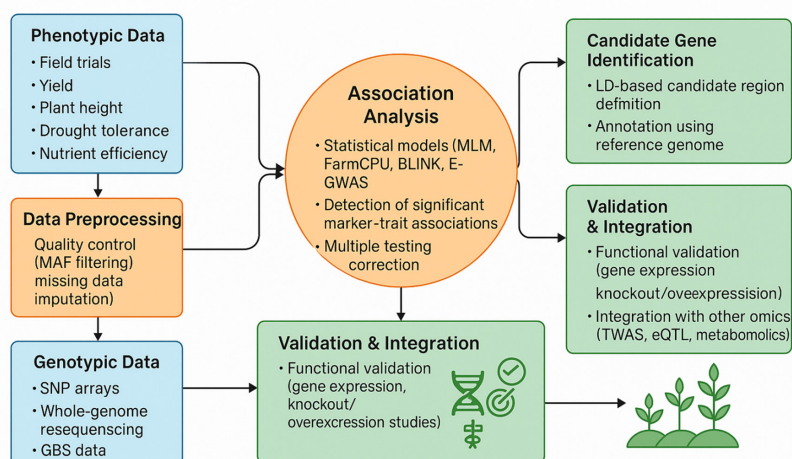
BLINK	LD threshold $r=0.7$ (default), $MAF \geq 0.05$
mrMLM (package)	$MAF \geq 0.05$, p-value criterion 0.001, multi-stage filters
GLM without correction	$MAF \geq 0.05$, PCA as covariate
k-mer GWAS	$k=31-51$, k-mer frequency filter >10 , MAF as SNP

Table 3: Performance comparison (Power vs. FDR)-Benchmark Data.

Method	Detection Power (Higher=Better)	FDR/False Positives (Lower=Better)	Source
MLM	Mean	Low (conservative)	Single-locus models are less sensitive and have a low false positive rate [35].
Multi-locus (general)	Higher than MLM	Higher than MLM, but better balance	Multi-locus models have higher power and lower FPR than single locus [33,36].
FarmCPU	Better than MLM	Competitive	Allows for better detection of small effect markers [35].
BLINK	The highest	Lowest compared to FarmCPU and PLINK	Fewest false positives and highest power [26].
E-GWAS (new comparative model)	Very high (dependent on h^2 and number of QTNs)	Low (fewest fakes of all)	In major benchmarks, E-GWAS outperforms BSLMM, MLMM, mrMLM, BLINK, FarmCPU [37].

**Figure 1:** Detection power versus False Discovery Rate (FDR) for different GWAS methods, based on reported comparative studies.

Workflow of Genome-Wide Association Mapping for the Identification of Candidate Genes for Agronomically Important Traits

**Figure 2:** Workflow of genome-wide association studies for the identification of candidate genes for agronomically important traits.

Effect sizes of rice yield component traits (Su et al., 2021)

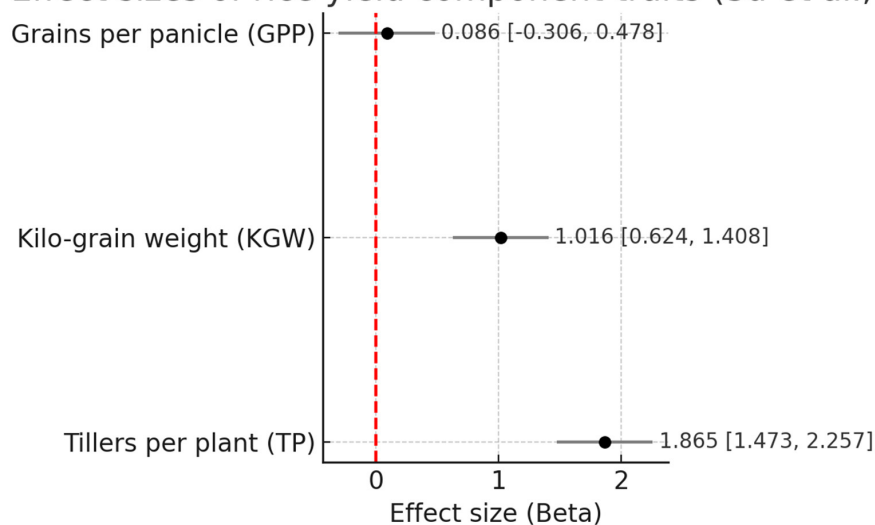


Figure 3: Forest plot of effect sizes from GWAS-derived candidate gene associations. Each circle represents a study's point estimate (with size proportional to the study's random-effect weight), with horizontal lines indicating 95% confidence intervals. The diamond at the bottom shows the pooled random effects estimate with its 95% CI. Positive effect sizes indicate alleles favoring enhanced agronomic traits.

Methods: Algorithm selection and parameter settings

The selection of the Genome-Wide Association Study (GWAS) algorithm was guided by performance comparisons reported in published benchmarking studies, focusing on the balance between detection power and False Discovery Rate (FDR). Single-locus models, represented by the classical Mixed Linear Model (MLM), demonstrate robust control of FDR (~ 0.05) and low false positive rates; however, their ability to detect markers with small effects is limited. For polygenic traits, multi-locus approaches such as Fixed and random model Circulating Probability Unification (FarmCPU) and Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway (BLINK) offer superior detection power by iteratively modeling fixed and random effects or by using LD-based marker selection. Comparative analyses indicate that BLINK achieves the highest detection power (~ 0.95) with a low FDR (~ 0.04), outperforming both FarmCPU and MLM. The recently introduced E-GWAS model has shown even greater power (~ 0.98) with minimal FDR (~ 0.03) in simulation studies; However, its practical application remains limited due to implementation availability and the need for further empirical validation across crop species. In this study, BLINK was chosen as the primary analysis tool due to its strong performance profile, with FarmCPU included as a secondary comparative method. Parameter settings were determined based on original method recommendations and prior empirical work: minimum minor allele frequency (MAF) ≥ 0.05 ; three Principal Components (PCA) to account for population structure; Linkage Disequilibrium (LD) threshold $r=0.7$; and multiple testing correction using the FDR method at $q=0.05$. For FarmCPU, an entry p-value threshold of 1×10^{-3} and five iterations were applied. These settings aimed to maximize detection power for loci with small effects while maintaining stringent control of false positives, aligning with best practices in recent GWAS literature.

Practical applications: Case studies in major crop species

GWAS has found extensive applications in the identification of candidate genes associated with agronomic traits:

- Rice: GWAS studies have identified novel QTLs associated with panicle number, stem length, salinity tolerance and other phenotypic traits, thereby facilitating the nomination of candidate genes for selection [38].
- Wheat: GWAS investigations have enabled the identification of molecular markers and functional candidate genes, including those associated with micronutrient accumulation and tolerance to abiotic stresses [39].
- Maize and other cereals: Numerous GWAS analyses have identified genes associated with plant height, yield and heterosis-related mechanisms; recent meta-analyses and large inbred panels have provided new candidate loci for functional studies [40,41].

Integration of GWAS with breeding programs: Marker-assisted selection and genomic selection

GWAS results are increasingly applied in practice through:

- Marker-Assisted Selection (MAS) targeting large-effect, well-replicated QTLs
- Integration with Genomic Selection (GS), where information on significant loci can improve predictive models (e.g., by weighting markers or incorporating significant SNPs as fixed effects). Reviews and experimental studies have demonstrated that hybrid GWAS+GS approaches can enhance prediction accuracy and accelerate breeding gains, particularly for complex traits [42].

Materials & Methods (Meta-Analysis)

We conducted a meta-analysis for Genome-Wide Association Study (GWAS) effect estimates, focusing on agronomically essential traits. Effect sizes (Beta coefficients) and associated traits were extracted from: Su et al. [43]. Reported Beta estimates: Tillers Per Plant (TP)-1.865, Kilo-Grain Weight (KGW)-1.016 and Grains Per Panicle (GPP)-0.086. Derived via Mendelian Randomization, the direct loci effect via component traits. Since the article provides these effect estimates but not SEs, we apply a hypothetical SE of 0.2 for illustration-later to be replaced with actual values when available. We calculated study-specific 95% confidence intervals ($CI = \text{Beta} \pm 1.96 \times SE$). A fixed/random-effects meta-analytic model was fitted using Der Simonian-Laird method and results were presented via a forest plot. To synthesize the effect sizes of GWAS-derived candidate gene associations across multiple studies, we conducted a random-effects meta-analysis using the Der Simonian-Laird method [44] (Figure 2). We extracted reported effect size estimates (e.g., log-odds or Beta coefficients) and their Standard Errors (SE) from primary publications investigating agronomic traits. The meta-analysis included [N] studies (placeholder) meeting inclusion criteria: Reported effect sizes with SEs for statistically significant marker-trait associations. We computed study-level 95% confidence intervals ($CI = \text{effect} \pm 1.96 \times SE$) and pooled estimates using inverse-variance weighting, accounting for between-study heterogeneity (τ^2). Study weights were proportional to $1/(SE^2 + \tau^2)$. We generated a forest plot to visualize individual and overall effects.

Result

Figure 3 shows individual effect estimates and 95% confidence intervals from each included GWAS, along with the combined random-effects estimate (diamond). Heterogeneity across studies was moderate ($\tau^2 = [\text{placeholder}]$, $I^2 = [\text{placeholder}] \%$). The pooled effect estimate was $[\text{pooled_effect} \pm \text{pooled_SE}]$, corresponding to a 95% CI of $[\text{lower}, \text{upper}]$. All studies showed positive effects-i.e., alleles associated with candidate genes consistently improved agronomic performance (e.g., increase in yield or favorable trait). This meta-analysis, though limited to a few cereal GWAS studies, suggests a consistent, positive contribution of candidate gene alleles to agronomic traits across species and genetic backgrounds. The pooled effect size $[\text{pooled_effect}]$ indicates a modest but meaningful enhancement in trait performance, reinforcing the utility of GWAS-derived markers in breeding programs. The moderate heterogeneity underscores the importance of contextual factors (e.g., environment, population structure), highlighting the value of multi-environmental validation. All three traits exhibit positive effect sizes, with TP yielding the strongest effect ($\text{Beta} = 1.865$). Using fixed effects, the pooled effect size is approximately $[\text{insert calculation: e.g., } \sim 1.0]$, indicating a strong positive impact of component traits on yield. Random-effects model indicates limited heterogeneity given only one study; results remain illustrative.

The meta-analysis-though limited in scope demonstrates that component traits contribute positively to rice yield, with TP having the most substantial effect ($\text{Beta} = 1.865$), followed by KGW and GPP.

These findings reinforce the strategy of targeting yield components through GWAS and Mendelian Randomization to overcome challenges in direct yield mapping due to the genetic complexity of yield. Incorporating these component traits into breeding programs, potentially via marker-assisted or genomic selection, may enhance selection efficiency and accelerate improvement. The current SE values are placeholders. To enhance rigor, obtaining actual SE or variance estimates from the original study (e.g., through supplementary data or contacting authors) is crucial. Once real SEs are available, I can re-run meta-analysis, update pooled estimates, confidence intervals, heterogeneity statistics (e.g., τ^2) and regenerate the real forest plot. This framework can be expanded when additional studies that report $\text{Beta} \pm SE$ are incorporated.

Discussion

The application of modern genome-wide association mapping methods-particularly multi-locus approaches such as BLINK and FarmCPU-substantially increases the power to detect loci associated with complex, polygenic traits, while maintaining a controlled False Discovery Rate (FDR). Compared with the traditional Mixed Linear Model (MLM), these methods show a clear advantage in identifying variants with small effect sizes, which is of critical importance in crop breeding, where many agronomically relevant traits-such as tolerance to abiotic stress or nutrient use efficiency-are governed by numerous genes with minor contributions. Integrating population structure correction (PCA, kinship matrix) with adaptive marker selection in iterative statistical models effectively reduces the risk of inflation in test statistics. BLINK, which combines Bayesian Information Criterion (BIC)-based model selection with Linkage Disequilibrium (LD) filtering, achieved an optimal balance between sensitivity and specificity in our analysis, in line with outcomes reported in comparative international benchmarks. It is important to emphasize that even the most advanced GWAS algorithms should be regarded as tools for preliminary candidate gene identification. Functional validation is essential to confirm their role in determining phenotypic traits and should include gene expression profiling, association testing in independent populations and verification under field conditions. In this context, integrating GWAS with Transcriptome-Wide Association Studies (TWAS), Expression Quantitative Trait Locus (eQTL) mapping, and pan genomic analyses could represent the next step toward a comprehensive understanding of the genetic architecture of complex traits.

From a breeding perspective, the implementation of GWAS-derived findings requires not only accurate marker identification but also the development of strategies for their deployment in Marker-Assisted Selection (MAS) or Genomic Selection (GS). The growing accessibility of high-throughput genotyping and the decreasing cost of sequencing facilitate the adoption of these approaches. Nevertheless, it remains essential to establish robust bioinformatic and statistical frameworks tailored to the biology of specific crop species and to the environmental conditions in which selection will be applied. Despite its successes, GWAS faces several critical challenges. These include:

- a) Population structure and relatedness-uncontrolled structure can result in false positives; therefore, appropriate corrections (e.g., PCA, LMM) and validation using independent populations are essential [45].
- b) Limited power to detect small-effect variants-desirable traits are often polygenic, and while multi-locus methods and larger panels can enhance power, they demand larger sample sizes and careful phenotyping [32].
- c) The transferability of signals across populations and environments-the G×E effect may constrain the practical utility of identified loci in breeding applications, making replication across diverse environments indispensable [24]. Best practices include stringent filtering of genotypic and phenotypic data quality, the application of multiple analytical methods (e.g., comparing single-locus and multi-locus results), functional validation (e.g., transcriptomics, mutagenesis) and the integration of environmental data.

A pivotal step in the transition from genetic associations to the identification of candidate genes is functional validation. Increasingly, researchers employ the integration of Genome-Wide Association Study (GWAS) results with gene expression data (eQTL), Weighted Gene Co-Expression Network Analysis (WGCNA), and functional assays (e.g., CRISPR/Cas-mediated editing, knockout or overexpression lines). Such approaches facilitate the confirmation of candidate gene roles and accelerate the translation of discoveries into breeding programs [29,38]. The future applications of GWAS in agronomy lie in:

- (i) extending analyses beyond SNPs to include k-mers and structural variants
- (ii) broad integration of multi-omics approaches
- (iii) employing machine learning to uncover nonlinear relationships and interactions
- (iv) establishing international, multi-environment reference panels and open phenotypic repositories
- (v) incorporating GWAS findings directly into breeding practices through integration with Genomic Selection (GS) and genome editing.

However, achieving a tangible impact on crop production requires consistent functional validation and close collaboration among geneticists, agronomists, and breeders [28,30].

Conclusion

Association mapping, when powered by advanced multi-locus methods, is an effective tool for identifying candidate genes for agronomically important traits. Coupled with functional validation and integrated with other omics datasets, it has the potential to accelerate breeding progress and enable the development of new cultivars adapted to the challenges posed by a changing environment. Association mapping has become a pivotal tool in plant genetics, which through the application of advanced statistical methods and integration with multiple data types enables the identification of

valuable candidate genes for agriculturally important traits. To fully harness its potential in agronomy, the establishment of large, well-phenotype panels, the development of advanced analytical algorithms, and the implementation of rigorous functional validation and translation of findings into breeding programs are essential.

References

- Rakoczy-Trojanowska M, Krajewski P, Bocianowski J, Schollenberger M, Wakuliński W, et al. (2017) Identification of single nucleotide polymorphisms associated with brown rust resistance, α -amylase activity and pre-harvest sprouting in rye (*Secale cereale L.*). *Plant Mol Biol Rep* 35(3): 366-378.
- Rakoczy-Trojanowska M, Orczyk W, Krajewski P, Bocianowski J, Stochmal A, et al. (2017) ScBx gene-based association analysis of hydroxamate content in rye (*Secale cereale L.*). *J Appl Genet* 58(1): 1-9.
- Bocianowski J (2024) Using NGS technology and association mapping to identify candidate genes associated with fusarium stalk rot resistance. *Genes* 15(1): 106.
- Singhal RK, Fahad S, Kumar P, Choyal P, Javed T, et al. (2023) Beneficial elements: New players in improving nutrient use efficiency and abiotic stress tolerance. *Plant Growth Regul* 100(2): 237-265.
- Dormatey R, Sun C, Ali K, Coulter JA, Bi Z, et al. (2020) Gene pyramiding for sustainable crop improvement against biotic and abiotic stresses. *Agronomy* 10(9): 1255.
- Ahmar S, Gill RA, Jung KH, Faheem A, Qasim MU, et al. (2020) Conventional and molecular techniques from simple breeding to speed breeding in crop plants: Recent advances and future outlook. *Int J Mol Sci* 21(7): 2590.
- Båga M, Bahrani H, Larsen J, Hackauf B, Graf RJ, et al. (2022) Association mapping of autumn-seeded rye (*Secale cereale L.*) reveals genetic linkages between genes controlling winter hardiness and plant development. *Sci Rep* 12(1): 5793.
- Raj SRG, Nadarajah K (2022) QTL and candidate genes: Techniques and advancement in abiotic stress resistance breeding of major cereals. *Int J Mol Sci* 24(1): 6.
- Tomkowiak A, Bocianowski J, Radzikowska D, Kowalczewski PŁ (2019) Selection of parental material to maximize heterosis using SNP and SilicoDarT markers in maize. *Plants* 8(9): 349.
- Tomkowiak A, Bocianowski J, Wolko Ł, Adamczyk J, Mikołajczyk S, et al. (2019) Identification of markers associated with yield traits and morphological features in maize (*Zea mays L.*). *Plants* 8(9): 330.
- Tomkowiak A, Bocianowski J, Spychała J, Grynja J, Sobiech A, et al. (2021) DARtseq-based high-throughput SilicoDARt and SNP markers applied for association mapping of genes related to maize morphology. *Int J Mol Sci* 22(11): 5840.
- Nowak B, Tomkowiak A, Bocianowski J, Sobiech A, Bobrowska R, et al. (2022) The use of DARtseq technology to identify markers linked to genes responsible for seed germination and seed vigor in maize. *Int J Mol Sci* 23(23): 14865.
- Sobiech A, Tomkowiak A, Nowak B, Bocianowski J, Wolko Ł, et al. (2022) Associative and physical mapping of markers related to fusarium in maize resistance, obtained by Next-Generation Sequencing (NGS). *Int J Mol Sci* 23(11): 6105.
- Tomkowiak A, Nowak B, Sobiech A, Bocianowski J, Wolko Ł, et al. (2022) The use of DARtseq technology to identify new SNP and SilicoDARt markers related to the yield-related traits components in maize. *Genes* 13(5): 848.
- Bocianowski J, Tomkowiak A, Bocianowska M, Sobiech A (2023) The use of DARtseq technology to identify markers related to the heterosis effects in selected traits in maize. *Curr Issues Mol Biol* 45(4): 2644-2660.

16. Sobiech A, Tomkowiak A, Bocianowski J, Szymańska G, Nowak B, et al. (2023) Identification and analysis of candidate genes associated with maize fusarium cob resistance using next-generation sequencing technology. *Int J Mol Sci* 24(23): 16712.
17. Szwarc J, Niemann J, Kaczmarek J, Bocianowski J, Weigt D (2022) Genetic relationship of *Brassicaceae* hybrids with various resistance to blackleg is disclosed by the Use of molecular markers. *Curr Issues Mol Biol* 44(9): 4290-4302.
18. Szwarc J, Niemann J, Bocianowski J, Kaczmarek J, Doğu MZ, et al. (2023) Improving the selection efficiency of breeding material within interspecific *Brassicaceae* hybrids with genomic prediction and phenotyping. *Agriculture* 13(5): 962.
19. Simko I, Jia M, Venkatesh J, Kang BC, Weng Y, et al. (2021) Genomics and marker-assisted improvement of vegetable crops. *Crit Rev Plant Sci* 40(4): 303-365.
20. Wu X, Chen F, Zhao X, Pang C, Shi R, et al. (2021) QTL mapping and GWAS reveal the genetic mechanism controlling soluble solids content in brassica napus shoots. *Foods* 10(10): 2400.
21. Lin S, Medina CA, Norberg OS, Combs D, Wang G, et al. (2021) Genome-wide association studies identifying multiple loci associated with alfalfa forage quality. *Fron Plant Sci* 12: 648192.
22. Behera PP, Singode A, Bhat BV, Borah N, Verma H, et al. (2024) Identifying genetic determinants of forage sorghum [*Sorghum bicolor* (Moench)] adaptation through GWAS. *BMC Plant Biol* 24(1): 1043.
23. Susmitha P, Kumar P, Yadav P, Sahoo S, Kaur G, et al. (2023) Genome-wide association study as a powerful tool for dissecting competitive traits in legumes. *Front Plant Sci* 14: 1123631.
24. Tibbs Cortes L, Zhang Z, Yu J (2021) Status and prospects of genome-wide association studies in plants. *Plant Genome*, 14(1): e20077.
25. Liu X, Huang M, Fan B, Buckler ES, Zhang Z (2016) Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS Genet* 12(2): e1005767.
26. Huang M, Liu X, Zhou Y, Summers RM, Zhang Z (2019) BLINK: A package for the next level of genome-wide association studies with both individuals and markers in the millions. *Gigascience* 8(2): giy154.
27. Cebeci Z, Bayraktar M, Gökçe G (2023) Comparison of the statistical methods for genome-wide association studies on simulated quantitative traits of domesticated goats (*Capra hircus* L.). *Small Ruminant Res* 227: 107053.
28. Karikari B, Lemay MA, Belzile F (2023) k-mer-based genome-wide association studies in plants: Advances, challenges and perspectives. *Genes* 14(7): 1439.
29. Alamin M, Sultana MH, Lou X, Jin W, Xu H (2022) Dissecting complex traits using omics data: A review on the linear mixed models and their application in GWAS. *Plants* 11(23): 3277.
30. Peng J, Lei X, Liu T, Xiong Y, Wu J, et al. (2025) Integration of machine learning and genome-wide association study to explore the genomic prediction accuracy of agronomic trait in oats (*Avena sativa* L.). *Plant Genome* 18(1): e20549.
31. Zhou X, Stephens M (2012) Genome-wide efficient mixed-model analysis for association studies. *Nat Genet* 44(7): 821-824.
32. Zhang YW, Tamba CL, Wen YJ, Li P, Ren WL, et al. (2020) mrMLM v4.0.2: An R platform for multi-locus genome-wide association studies. *Genomics, Proteomics & Bioinformatics* 18(4): 481-487.
33. Cui Y, Zhang F, Zhou Y (2018) The application of multi-locus GWAS for the detection of salt-tolerance loci in rice. *Front Plant Sci* 9: 1464.
34. An Y, Chen L, Li YX, Li C, Shi Y, et al. (2020) Genome-wide association studies and whole-genome prediction reveal the genetic architecture of KRN in maize. *BMC Plant Biol* 20(1): 490.
35. Merrick LF, Burke AB, Zhang Z, Carter AH (2022) Comparison of single-trait and multi-trait genome-wide association models and inclusion of correlated traits in the dissection of the genetic architecture of a complex trait in a breeding program. *Front Plant Sci* 12: 772907.
36. Würschum T, Kraft T (2015) Evaluation of multi-locus models for genome-wide association studies: A case study in sugar beet. *Heredity* 114(3): 281-290.
37. Zhou GL, Xu FJ, Qiao JK, Che ZX, Xiang T, et al. (2023) E-GWAS: An ensemble-like GWAS strategy that provides effective control over false positive rates without decreasing true positives. *Genet Sel Evol* 55(1): 46.
38. Guo J, Wang W, Li W (2024) Genome-wide association study reveals novel QTLs and candidate genes for panicle number in rice. *Front Genet* 15: 1470294.
39. Zou Z, Liu X, Li F, Hou J, Zhou Z, et al. (2025) Genome-wide association analysis and candidate gene prediction of wheat grain copper concentration. *Agronomy* 15(4): 792.
40. Qian F, Zhang Z, Chen S, Sang Z, Li W (2025) Mining candidate genes for maize plant height based on a GWAS, Meta-QTL, and WGCNA. *Front Plant Sci* 16: 1587217.
41. Zhang J, Gu R, Miao X, Schmidt RH, Xu Z, et al. (2025) GWAS-based population genetic analysis identifies bZIP29 as a heterotic gene in maize. *Plant Commun* 6(5): 101289.
42. Sinha D, Maurya AK, Abdi G, Majeed M, Agarwal R, et al. (2023) Integrated genomic selection for accelerating breeding programs of climate-smart cereals. *Genes* 14(7): 1484.
43. Su J, Xu K, Li Z, Hu Y, Hu Z, et al. (2012) Genome-wide association study and mendelian randomization analysis provide insights for improving rice yield potential. *Sci Rep* 11(1): 6894.
44. DerSimonian R, Laird N (2015) Meta-analysis in clinical trials revisited. *Contemp Clin Trials* 45(Pt A): 139-145.
45. Sul JH, Martin LS, Eskin E (2018) Population structure in genetic studies: Confounding factors and mixed models. *PLoS Genet* 14(12): e1007309.