Review and Progress Open Access

Accelerating Yam Breeding Cycles with Genomic Selection

Wenhui Yu [1], Quantong Cheng [2]
1 Quzhou Academy of Agricultural and Forestry Sciences, Quzhou, 324000, Zhejiang, China
2 Kaihua County Agricultural Technology Extension Center, Kaihua, 324300, Zhejiang, China
Corresponding author: 565326696@qq.com

Genomics and Applied Biology, 2024, Vol.15, No.6 doi: 10.5376/gab.2024.15.0033
Received: 22 Oct., 2024
Accepted: 28 Nov., 2024
Published: 13 Dec., 2024

Preferred citation for this article:
Yu W.H., and Cheng Q.T., 2024, Accelerating yam breeding cycles with genomic selection, Genomics and Applied Biology, 15(6): 319-331 (doi: 10.5376/gab.2024.15.0033)

Abstract This study evaluates the potential of genomic selection (GS) in yam breeding, particularly its advantages in shortening breeding cycles and enhancing the selection efficiency of key agronomic traits. Traditional yam breeding is constrained by long generation times and complex genetic structures. By utilizing genomic data, GS can significantly accelerate generation turnover and increase genetic gain. This study discusses the role of integrating GS with speed breeding in yam breeding, as well as the advantages of GS in managing the complexity of genotype-by-environment (GxE) interactions. In addition, the development and optimization of genomic resources and their potential applications in yam breeding are explored. The application of GS not only improves breeding efficiency but also has a positive impact on the quality and productivity of yam varieties, providing new approaches to enhance global food security.

Keywords Yam breeding; Genomic selection (GS); Speed breeding; Genetic gain; Genotype-by-environment interaction (GxE)

1 Introduction

Yam (*Dioscorea* spp.) is an essential staple food crop for millions of people, particularly in tropical and subtropical regions. It serves as a major source of carbohydrates and plays a vital role in the food security and economic stability of many developing countries. However, yam breeding faces significant challenges that hinder the development of improved cultivars. These challenges include the long growth cycle of yams, the genetic complexity of the crop, and the susceptibility to various biotic and abiotic stresses. Breeding programs, therefore, require innovative strategies to overcome these obstacles and meet the growing demands for more resilient, higher-yielding varieties (Crossa et al., 2017; Cappetta et al., 2020).

One of the main limitations in yam breeding is the extended duration of the breeding cycle. Conventional breeding methods often require multiple years to achieve meaningful genetic gains due to the long vegetative and reproductive phases of the crop (Lin et al., 2016). This slow pace delays the release of new cultivars that can respond to emerging challenges such as climate change, pests, and diseases. Thus, accelerating the breeding cycle is essential to enhance productivity and sustainability in yam cultivation (Asfaw et al., 2020).

Genomic Selection (GS) has emerged as a promising approach that can transform breeding programs by substantially reducing the time needed for cultivar development. Unlike traditional marker-assisted selection, which relies on identifying specific genetic markers associated with desirable traits, GS leverages genome-wide marker data to predict the breeding values of individuals. This method enables breeders to make more informed and faster selection decisions, thereby shortening the breeding cycle and enhancing genetic gain per unit of time. The integration of GS in yam breeding holds the potential to revolutionize the process, making it more efficient and responsive to the needs of both farmers and consumers. This study explores the challenges faced in yam breeding, the necessity of accelerating breeding cycles, and the transformative role that GS could play in addressing these challenges to drive the development of superior yam varieties.

2 Application of Genomic Selection in Yam Breeding

Genomic selection (GS) has emerged as a revolutionary tool for accelerating breeding cycles in crops, including yam. By predicting the genetic potential of candidate plants using genome-wide markers, GS enables breeders to make selection decisions early in the breeding cycle, thus reducing the time required for field evaluation and increasing genetic gain. In yam breeding, where generation times are long, GS presents a promising solution to enhance breeding efficiency and respond to the growing demand for resilient, high-yield varieties (An et al., 2024).

2.1 Basic concepts and technical background of genomic selection

Genomic Selection (GS) utilizes high-density genetic marker data distributed across the entire genome to predict the breeding value of individuals, which differs significantly from traditional marker-assisted selection (MAS). MAS focuses primarily on identifying and using a limited number of key quantitative trait loci (QTLs), which only explain part of the genetic variation for complex traits. In contrast, GS captures the effects of thousands of small-effect loci using genome-wide markers, providing a more comprehensive reflection of the genetic basis of complex traits. This approach greatly enhances the accuracy and effectiveness of predictions, especially for polygenic traits controlled by multiple small-effect genes, which are common in crops such as yams. By leveraging more extensive genetic information, GS accelerates the selection process, improves the efficiency and success rate of breeding programs, and facilitates the rapid development of new varieties with superior characteristics (Crossa et al., 2017).

The technical foundation of GS involves the use of genomic estimated breeding values (GEBVs), calculated by fitting statistical models that incorporate marker data. Training populations are genotyped and phenotyped to establish prediction models, which can then be applied to predict GEBVs of untested individuals based on their genotypes alone. Common statistical approaches include best linear unbiased prediction (BLUP) and various machine learning methods like Bayesian regression and random forests, which help account for complex genetic architectures and genotype-by-environment interactions (Heslot et al., 2012).

2.2 Potential of genomic selection in yam breeding

The application of GS in yam breeding has the potential to address several challenges that have traditionally hindered progress in this crop. Yam, particularly the *Dioscorea* species, is vital for food security in regions like West Africa but suffers from a long breeding cycle due to its reproductive characteristics and the need for extensive field trials. By allowing breeders to select superior genotypes early, GS can drastically shorten breeding cycles, leading to faster varietal release and more efficient breeding programs (Asfaw et al., 2020).

In addition to cycle acceleration, GS can enhance genetic gain by improving the accuracy of selection for complex traits like yield, disease resistance, and quality, which are crucial for yam improvement. Since these traits are influenced by multiple genetic loci, GS provides a more holistic approach than traditional phenotypic selection methods. Furthermore, GS allows for simultaneous improvement of multiple traits, reducing the trade-offs often encountered in conventional breeding (Lin et al., 2016).

2.3 Current status of genomic selection breeding worldwide and in domestic programs

Globally, the adoption of GS has increased significantly across crops like maize, wheat, and perennial ryegrass, where it has demonstrated substantial genetic gains and reduced breeding cycle times. For instance, in maize, GS has enabled faster accumulation of favorable alleles for yield and disease resistance, establishing GS as an integral part of modern breeding programs (Xu et al., 2019; Wang, 2024). Similarly, GS has shown promise in high-value crops like tomato, where rapid selection for complex traits has accelerated the breeding process (Cappetta et al., 2020).

In yam, GS is still an emerging tool but has garnered attention in regions where yam is a staple crop, particularly in West African countries. Collaborative breeding programs led by institutions such as the International Institute of Tropical Agriculture (IITA) are pioneering the integration of GS with conventional methods to enhance the genetic diversity and resilience of yam varieties. These programs focus on training populations representative of local varieties and environmental conditions, ensuring that genomic predictions are robust and applicable to target regions. As the cost of genotyping continues to decrease, it is expected that GS will become a core component of yam breeding programs worldwide, contributing to improved food security and agricultural sustainability in yam-growing regions (Saski et al., 2015).

3 Challenges and Opportunities in Yam Breeding

3.1 Major challenges in yam breeding: genetic diversity and complexity

Yam breeding, particularly for *Dioscorea* species, faces significant challenges due to the complexity of their genomes, which are often polyploid. Polyploid genomes, containing multiple sets of chromosomes, introduce layers of genetic intricacy that complicate breeding efforts. These complex genomes contain a high degree of genetic diversity, which, while essential for adaptation and resilience, presents hurdles in selectively breeding for desirable agronomic traits. Traits such as yield, drought tolerance, and disease resistance are controlled by multiple genes, a phenomenon known as polygenic control, which makes it challenging to pinpoint specific genes responsible for these characteristics (Saski et al., 2015). This polygenic nature, combined with high genetic variation, slows down the breeding process as it is difficult to isolate and stabilize desired traits.

Another critical challenge in yam breeding is the extensive genetic diversity among yam cultivars, which, although advantageous for the species' adaptability to different environments, complicates selection processes. This diversity means that breeders must carefully consider genotype-by-environment interactions to ensure that the selected traits perform consistently across different growing conditions. Such interactions can influence the expression of key traits, making it challenging to achieve stable trait performance in a single variety that can adapt to a wide range of environments (Agre et al., 2021). Consequently, breeders must conduct extensive multi-environment trials to identify cultivars that can reliably express traits like high yield and resistance to pests and diseases under varied environmental pressures.

3.2 Limitations of traditional breeding methods

Traditional yam breeding methods are constrained by long generation times, as yam plants require extended growth periods to develop observable phenotypic traits. This leads to prolonged breeding cycles, often spanning several years, significantly delaying the development of new varieties. Additionally, traditional selection methods rely heavily on phenotypic assessments, which can be inconsistent across different environments, impacting the accuracy and stability of trait evaluation. Such variability poses challenges in selection, as traits that perform well in one environment may not exhibit the same favorable expression in another, limiting the effectiveness and adaptability of new varieties (Asfaw et al., 2020).

Traditional breeding approaches also depend on labor-intensive field trials and complex crossbreeding processes, further slowing down the breeding timeline. Field trials require multiple observations and measurements over extended periods to gather sufficient data, consuming significant time and resources. This dependency on lengthy and labor-intensive methods hampers the ability of breeders to respond quickly to emerging agricultural needs, such as shifting disease pressures, climate change, and evolving market demands. Yam’s extended growth cycles and environmental requirements make it difficult to efficiently select individuals with desired traits, resulting in breeding programs that struggle to keep pace with modern agricultural challenges and limiting the speed and adaptability of yam crop improvement.

3.3 Feasibility of genomic selection as a solution

Genomic selection (GS) offers a transformative approach to yam breeding by allowing early selection based on genomic estimated breeding values (GEBVs). This approach can significantly reduce the breeding cycle and improve the accuracy of selecting for complex traits, such as yield, disease resistance, and stress tolerance, that are difficult to assess through traditional phenotypic selection alone. By utilizing genome-wide molecular markers, GS enables breeders to make informed, data-driven selection decisions even before plants reach maturity, eliminating the need for extended growth periods to observe full phenotypic traits. This early selection capability not only speeds up the breeding process but also reduces the costs and labor associated with extensive field trials, allowing breeders to focus on plants with the highest genetic potential from the start.

This passage discusses the application of Genomic Selection (GS) in yam breeding, particularly in predicting key agronomic traits such as yield and disease resistance. Studies have shown that, by using dense genome-wide markers, GS models can more reliably predict breeding values, overcoming limitations of traditional methods, which often require time-consuming phenotypic assessments that are influenced by environmental factors (Norman et al., 2022). Furthermore, advancements in genomic resources for yam, such as high-quality reference genomes and comprehensive marker datasets, have made the integration of GS into breeding projects increasingly feasible and cost-effective. The reduction in genotyping costs, combined with improved computational tools, enables breeders to incorporate GS into yam breeding strategies, enhancing the overall efficiency, responsiveness, and adaptability of breeding programs to meet evolving agricultural demands (Crossa et al., 2017).

The adoption of GS in yam breeding also holds the potential to address long-standing challenges associated with genotype-by-environment interactions by allowing breeders to assess genetic potential across diverse environments without requiring extensive multi-location trials. This shift towards genomics-driven breeding represents a viable solution for accelerating yam crop improvement and achieving stable, high-performance varieties capable of meeting future agricultural needs more effectively.

4 Development and Utilization of Genomic Resources

4.1 Methods for establishing genomic resources for yam

The establishment of genomic resources for yam has greatly benefited from advancements in next-generation sequencing (NGS) technologies, which have provided critical genetic data to support breeding and improvement efforts. Techniques such as genotyping-by-sequencing (GBS), expressed sequence tag (EST) sequencing, and whole-genome sequencing (WGS) are among the primary NGS approaches used to generate detailed genomic information for yam. Genotyping-by-sequencing, for example, has proven effective in producing high-density single nucleotide polymorphism (SNP) markers across yam genomes. These SNPs are invaluable for accurately mapping complex traits, including yield, disease resistance, and environmental adaptability, which are essential for developing improved yam varieties. By identifying specific genomic regions associated with these traits, breeders can make informed selections to accelerate yam breeding cycles and enhance genetic gains (Saski et al., 2015).

In addition to SNP markers, microsatellite markers, or simple sequence repeats (SSRs), have been developed for yam and offer further benefits for breeding programs. SSRs are highly polymorphic, which allows them to reveal substantial genetic diversity within and across yam species. This diversity is crucial for yam breeding, as it provides the genetic variation necessary for adaptability and resilience under changing environmental conditions. Moreover, SSR markers are known for their cross-species transferability, enabling their application across various yam species. This cross-applicability is especially advantageous for breeding programs, as it allows breeders to efficiently analyze genetic diversity and track desirable traits across different breeding populations, making marker-assisted selection (MAS) both feasible and effective (Diouf et al., 2023).

These genomic resources not only streamline traditional breeding methods but also facilitate advanced breeding techniques like genome-wide association studies (GWAS) and genomic selection (GS). With access to comprehensive genomic data, breeders can better understand the genetic underpinnings of important traits and develop varieties tailored to specific agricultural challenges. As the cost of NGS continues to decrease, the expansion of these genomic resources will further enhance yam breeding programs, helping to create high-yielding, resilient yam varieties that can meet the demands of future food security and sustainability.

4.2 Applications of whole-genome profiling and marker development

Whole-genome profiling has become a cornerstone in modern yam breeding, enabling efficient marker discovery, detailed genetic diversity analysis, and quantitative trait locus (QTL) mapping for complex, multi-gene traits. Single nucleotide polymorphisms (SNPs) and simple sequence repeat (SSR) markers, generated through whole-genome sequencing (WGS) and genotyping-by-sequencing (GBS), allow for high-resolution trait mapping. This detailed genetic information is instrumental for trait association studies, supporting breeders in identifying the genetic basis of desirable traits such as disease resistance, yield, and environmental adaptability. By using these markers within genomic selection models, breeders can assess and predict the genetic potential of yam plants early in the breeding cycle, accelerating the development of improved varieties (Tamiru et al., 2017).

The creation of comprehensive marker databases, such as the Yam Microsatellite Markers Database (Y2MD), has further streamlined the breeding process. Y2MD provides breeders with rapid access to an extensive collection of SSR markers, promoting data sharing and collaboration within the yam breeding community. This shared resource allows for more efficient screening and selection across breeding programs, fostering the development of resilient yam cultivars suited to diverse environmental conditions and agricultural challenges (Diouf et al., 2023). As these resources continue to grow, they play a pivotal role in enhancing the precision and speed of yam breeding efforts globally.

4.3 Role of genomic resources in yam breeding

Genomic resources have transformed yam breeding by offering critical tools for identifying genetic diversity and advancing the development of resilient yam cultivars. These resources, which include comprehensive databases of genetic markers, allow breeders to employ both marker-assisted selection (MAS) and genomic selection (GS) with high precision. By integrating MAS and GS, breeders can significantly streamline the selection process, allowing for faster breeding cycles and more targeted trait selection. This approach is particularly valuable for enhancing traits like adaptability, yield, and disease resistance, which are essential for ensuring sustainable and productive yam varieties under varying environmental conditions.

The availability of genotyping technologies, such as genotyping-by-sequencing (GBS), simple sequence repeats (SSRs), and single nucleotide polymorphisms (SNPs), has greatly supported trait-specific breeding efforts. With these tools, breeders can make data-driven decisions, especially for complex, polygenic traits like yield and disease resistance, which are typically challenging to select using traditional methods alone. By enabling precise trait mapping and early prediction of genetic potential, these genomic resources ensure that yam breeding programs are better equipped to meet agricultural demands, promoting food security and sustainability in yam-producing regions (Mulualem et al., 2018).

5 Development and Optimization of Genomic Prediction Models

The development and optimization of genomic prediction models are fundamental for the successful application of genomic selection in yam breeding. These models enable breeders to predict the genetic potential of breeding candidates based on their genomic profiles, allowing for earlier and more precise selection. Genomic prediction models can be tuned and improved to better predict complex traits, which are common in yam, including yield, disease resistance, and adaptability. Each model has its own strengths and limitations, and careful selection of the right model for specific traits and conditions is crucial to maximize genetic gains and reduce breeding cycle time.

5.1 Commonly used genomic prediction models and their principles

Genomic prediction models are essential tools in plant and animal breeding, allowing the estimation of breeding values based on genome-wide markers. These models vary significantly in their assumptions and mathematical approaches. One of the most widely used models is the best linear unbiased prediction (BLUP) approach, particularly genomic BLUP (GBLUP). GBLUP estimates breeding values by incorporating a genomic relationship matrix derived from SNP (single nucleotide polymorphism) data. This approach assumes that genetic effects are additive and normally distributed. GBLUP is efficient for high-dimensional data as it treats all markers as having small, equal effects, which is useful for predicting polygenic traits (Ge et al., 2020).

Bayesian models are another common class of genomic prediction models, including methods like Bayes A, Bayes B, and Bayes Cπ. These models differ from GBLUP in that they allow some markers to have larger effects than others, which can be advantageous for traits controlled by a few major genes. Bayes A assumes that all markers have a similar prior variance, while Bayes B and Bayes Cπ use a mixture prior where some markers are assumed to have no effect. This flexibility enables Bayesian models to capture genetic architectures where only a subset of markers contributes significantly to trait variability (Crossa et al., 2017).

In addition to these traditional models, nonparametric and machine learning-based approaches are increasingly applied in genomic prediction. Support vector machines (SVM) and deep learning models, like convolutional neural networks (CNN) and multilayer perceptrons (MLP), are examples. These methods do not assume an additive genetic architecture and can model non-additive genetic effects, which are common in complex traits. Nonparametric models are especially useful for capturing complex interactions between markers and environmental variables, although they require large datasets and extensive computational power (Zhao et al., 2020). Altogether, the choice of model depends on the trait architecture, computational resources, and specific breeding goals.

5.2 Comparative effectiveness of different models in yam genomic selection

Different genomic prediction models demonstrate varying levels of effectiveness depending on the trait being studied, the genetic architecture, and the available data. In yam breeding, where traits like yield, disease resistance, and tuber quality are highly complex, the choice of an effective prediction model is critical. Studies comparing the performance of models, such as GBLUP and Bayesian methods, suggest that Bayesian models often outperform GBLUP for traits controlled by a few major loci due to their flexibility in handling marker-specific effects. For example, in traits where non-additive genetic effects are substantial, Bayesian models like Bayes B and Bayes Cπ have shown better predictive accuracy because they allow some markers to have larger effects than others. From the study by Asfaw et al. (2020), it can be seen that the genetic variance composition differs among traits. For example, the dominance component is higher for tuber yield and tuber number, while the additive component is lower for fresh tuber yield. This information helps in understanding the genetic control patterns of different traits (Asfaw et al., 2020) (Figure 1).

However, GBLUP remains widely used and is considered robust for a broad range of traits, particularly polygenic ones where many genes with small effects collectively contribute to the trait. GBLUP's simplicity and computational efficiency make it a practical choice, especially for breeders working with limited resources. Machine learning models, such as random forests, support vector machines, and deep learning approaches, have shown promising results in other crops for predicting traits with complex architectures. In recent research, deep learning models like convolutional neural networks (CNN) and multilayer perceptrons (MLP) have outperformed traditional models for some complex traits in crops such as wheat (Sandhu et al., 2021). These models could offer similar benefits in yam, where complex traits like yield and disease resistance involve intricate gene-environment interactions.

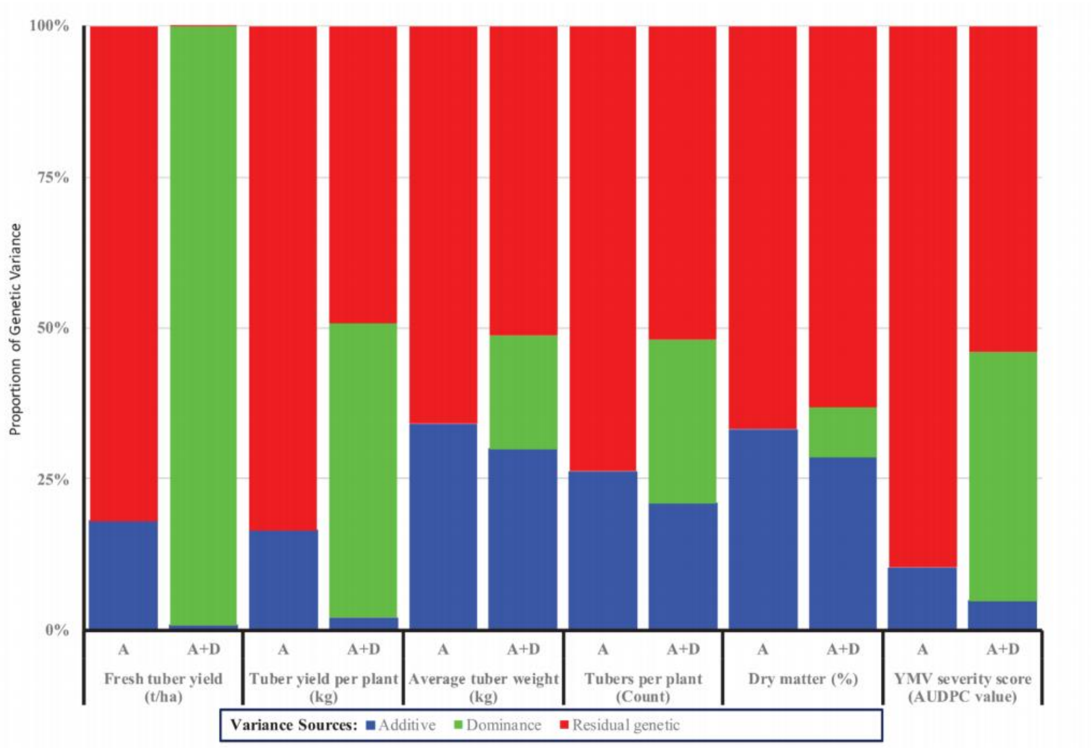

Figure 1 Genetic variance distribution of six traits in white Guinea yam (Adopted from Asfaw et al., 2020)

In conclusion, each model has its strengths and weaknesses, and no single model is universally optimal. GBLUP is efficient for highly polygenic traits, while Bayesian models excel when specific markers have substantial effects. Machine learning models may provide an edge in cases of complex gene interactions. Hybrid approaches, combining different models, can sometimes achieve better accuracy, capturing a broader range of genetic effects.

5.3 Adaptability of prediction models under different environmental conditions

One of the challenges in applying genomic prediction models in yam breeding is the adaptability of these models under diverse environmental conditions. Yam is cultivated in a variety of climatic regions, and traits like yield and disease resistance can vary significantly with environmental conditions. Prediction models must be capable of accounting for genotype-by-environment (GxE) interactions to be effective across different regions. GBLUP, for instance, can be extended with environmental interaction terms to model GxE effects, allowing predictions to consider both the genetic and environmental components. This is especially beneficial in crops like yam, where performance can be highly influenced by local climate and soil conditions (Crossa et al., 2017).

Models incorporating environmental covariates have shown improved accuracy in predicting trait performance across different environments. For instance, GBLUP models with interaction terms are used in crops like barley and wheat, where they have been shown to perform better than standard models in multi-environment trials. This adaptability is crucial in yam breeding, as traits such as drought resistance, tuber yield, and disease resistance are influenced by climatic variability. Researchers have also explored the use of multivariate models and mixed-model approaches to better capture GxE interactions, which can significantly improve prediction accuracy in yam (Ge et al., 2020).

Additionally, recent advances in high-throughput phenotyping and environmental monitoring provide data that can be incorporated into genomic prediction models. By integrating phenotypic and environmental data, machine learning models, like deep learning, can capture complex interactions that simpler models may miss. For example, CNNs and MLPs, when trained with environmental covariates, show promising accuracy improvements for yield predictions in crops with variable growing conditions (Sandhu et al., 2021). This approach holds potential for yam breeding, enabling breeders to make predictions that are more accurate and tailored to specific environments. Overall, the adaptability of genomic prediction models under diverse conditions enhances their utility, making them valuable tools for yam breeders working across various climates.

6 Strategies to Shorten the Yam Breeding Cycle

6.1 Application of speed breeding techniques

Speed breeding (SB) is an innovative technique designed to accelerate plant growth and development through the use of controlled environments, including extended photoperiods, greenhouses, and precise temperature and humidity settings. By adjusting these factors, SB allows multiple plant generations to complete their life cycles within a single year, significantly advancing the breeding process. Originally demonstrated in wheat, SB has shown potential for use in long-generation crops like yam, where it can reduce the time required for seed maturation and full plant development, ultimately speeding up the breeding cycle and improving genetic gains (Watson et al., 2019).

In yam breeding, speed breeding techniques could be applied by maintaining plants under artificial light for extended periods, simulating continuous day conditions to maximize photosynthesis and encourage rapid growth. Using greenhouse facilities, yam plants could undergo more frequent seed-to-seed cycles, reducing the time between generations. These methods allow breeders to focus on critical early traits such as seedling vigor, disease resistance, and drought tolerance under controlled conditions before field evaluation. Mondo et al. (2021) demonstrated that changes in calcium nitrate concentration significantly affect pollen germination rates in these two yam varieties, particularly with a notable increase in pollen germination at optimal concentrations. This finding supports the application of appropriate calcium nitrate concentrations in the breeding process to enhance yam pollen viability, thereby increasing the likelihood of successful pollination and fruit set (Mondo et al., 2021) (Figure 2). For example, studies in wheat have shown that by reducing the photoperiod, plants can go through several generations in one year, significantly expediting breeding cycles (Begna, 2022).

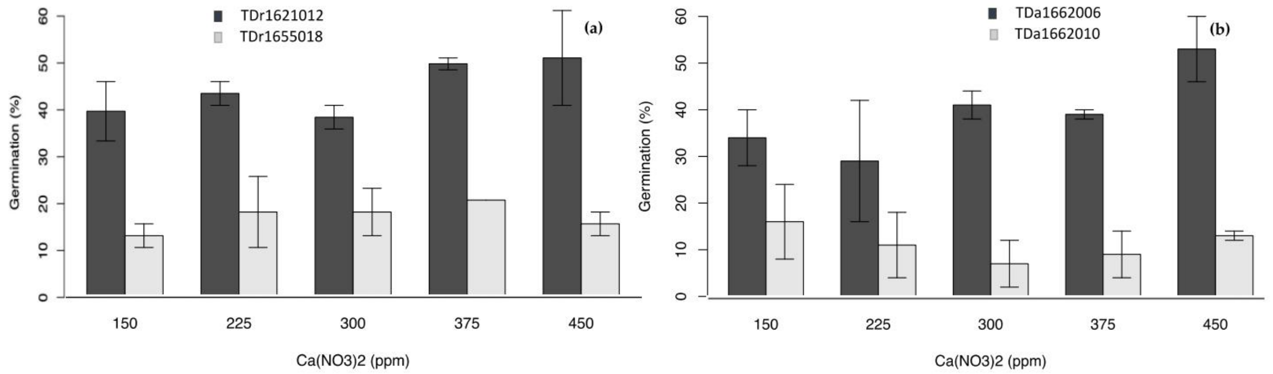

Figure 2 Effects of calcium nitrate on pollen germination of genotypes of (a) D. rotundata, (b) D. alata (Adopted from Mondo et al., 2021)

6.2 Potential of integrating genomic selection with speed breeding

The integration of GS with speed breeding creates a synergistic approach, often referred to as SpeedGS, which further accelerates breeding cycles. Genomic selection leverages genome-wide markers to predict the breeding value of individuals, enabling selection decisions without the need for full phenotypic data. When combined with speed breeding, GS allows for rapid generation advancement based on genomic predictions, reducing the breeding cycle length and increasing selection intensity (Jighly et al., 2019).

SpeedGS has proven effective in crops like wheat and maize, where rapid generation cycling and genomic predictions have resulted in increased genetic gains for traits with low heritability. In yam, integrating SpeedGS could facilitate the selection of complex traits, such as tuber yield and quality, which are difficult to evaluate in traditional breeding due to the long generation time. By using genomic estimated breeding values (GEBVs) to predict the performance of each generation in a speed-breeding setting, breeders can maximize genetic gain per unit of time while bypassing some of the time-consuming field trials. This combination allows for faster development of varieties with improved resilience to stresses and higher yields (Watson, 2019).

6.3 Impact of shortened breeding cycles on genetic gain

Shortened breeding cycles have a direct impact on genetic gain by allowing more rapid accumulation of favorable alleles in the breeding population. Traditional breeding approaches, limited by lengthy cycles, can delay the realization of genetic progress, especially for complex traits controlled by multiple genes. However, by reducing the breeding cycle through SpeedGS, breeders can accelerate the rate of genetic improvement, which is essential for achieving high-yield, resilient yam varieties (Das et al., 2020).

Studies indicate that SpeedGS increases genetic gain per year by shortening the cycle time and increasing selection intensity, especially for traits with low heritability, such as disease resistance and tuber quality. The rapid cycling enabled by SpeedGS permits breeders to explore broader genetic variability in a shorter time, allowing for faster adaptation to environmental challenges and market demands. However, one challenge of SpeedGS is the potential for increased inbreeding, which may reduce long-term genetic diversity. To address this, strategies such as random mating within selection rounds or controlled hybridization can be implemented to maintain genetic diversity while maximizing genetic gain (Li et al., 2021).

7 Challenges and Countermeasures in Implementing Genomic Selection

Implementing genomic selection (GS) in yam breeding programs is a complex process, requiring careful consideration of genetic, environmental, and logistical factors. While GS offers substantial potential for accelerating breeding cycles and improving trait selection accuracy, challenges such as marker density, sample size, and environmental interactions affect its efficacy. This section outlines these challenges and proposes countermeasures to enhance GS implementation in yam breeding.

7.1 Limitations of marker density and sample size

One of the primary challenges in genomic selection is achieving sufficient marker density to capture the genetic variation in yam. Marker density refers to the number of genetic markers (e.g., SNPs) used across the genome to predict breeding values accurately. Higher marker density generally improves prediction accuracy, but genotyping at high density can be cost-prohibitive, particularly for programs with limited resources. Low-density markers may miss essential genetic information, reducing the accuracy of genomic estimated breeding values (GEBVs) (Norman et al., 2018).

Sample size is another critical factor that influences the accuracy of genomic predictions. Smaller sample sizes limit the statistical power of GS models, leading to lower predictive accuracy. In GS, the training population must be representative of the breeding population to ensure that GEBVs are reliable when applied to new genotypes. Studies in crops like wheat have shown that increasing the sample size in training populations significantly enhances prediction accuracy (Grenier et al., 2015).

To address these limitations, one strategy is to use a moderate-density SNP panel, where markers are strategically spaced to capture as much genetic variation as possible with fewer markers. Alternatively, combining low-density genotyping with imputation techniques can fill in gaps and approximate high-density data cost-effectively (Lillehammer et al., 2013). Increasing the training population size is another solution, especially by pooling data from multiple breeding programs, which can improve prediction accuracy by diversifying the genetic representation.

7.2 Complexity of environment and gene interactions

Genotype-by-environment (GxE) interactions are particularly complex in yam breeding due to the crop’s cultivation across diverse environments. GxE interactions occur when different genotypes respond variably to environmental conditions, affecting the stability of trait predictions across locations and climates. In genomic selection, failing to account for GxE interactions can lead to biased predictions, as traits expressed in one environment may not translate consistently to another (Reyna et al., 2021).

To manage GxE complexity, models that incorporate environment-specific terms and allow for marker-by-environment interactions have been developed. For instance, mixed models that include GxE interaction terms improve prediction accuracy in multiple-environment trials. Studies in other crops, such as barley, show that these models capture spatial and environmental variability, allowing for more reliable trait predictions across diverse growing conditions (Oakey et al., 2016). Incorporating environmental covariates, like temperature and rainfall data, can also improve the precision of predictions by directly modeling environmental effects on trait expression. Furthermore, stratifying the breeding program to target specific environments or selecting genotypes within environment-specific sub-populations can reduce the impact of GxE interactions. This approach allows for the development of varieties tailored to particular environmental conditions, improving the robustness and applicability of genomic predictions.

7.3 Strategies to improve prediction accuracy (e.g., increasing data scale and marker diversity)

Improving the accuracy of GS predictions in yam breeding requires strategies that enhance both the scale and diversity of the data used in training models. One approach is to increase the training population size by combining data from multiple breeding programs or by incorporating historical phenotype and genotype data. Studies indicate that expanding the dataset improves model robustness and predictive power, especially when the genetic diversity within the training population reflects that of the target breeding population (Arruda et al., 2015).

Marker diversity is another crucial factor. High-density SNP arrays are traditionally used, but alternative marker types, such as microsatellites or multi-allelic markers, may capture additional genetic variation. Employing a mix of markers or using haplotype-based approaches can capture a broader range of genetic diversity, improving the power of GS models to detect quantitative trait loci (QTL) linked to important traits (Solberg et al., 2008).

Lastly, advanced machine learning models, such as random forests or deep learning approaches, offer potential for capturing complex genetic architectures, including epistatic interactions that are often missed in linear models. These models, although computationally intensive, can provide significant accuracy improvements for traits with complex inheritance patterns. Developing hybrid models that combine the simplicity of linear models with the adaptability of machine learning techniques may represent an optimal strategy for improving prediction accuracy in yam breeding (Norman et al., 2018).

8 Impact of Genomic Selection on Yam Breeding Efficiency

Genomic selection (GS) has introduced transformative changes in yam breeding efficiency by accelerating breeding cycles, enhancing genetic gain, and streamlining the breeding process. By utilizing genome-wide markers to predict the breeding value of individuals, GS enables earlier selection of superior genotypes, reducing the time and resources needed for field evaluations.

8.1 Expected effects of genomic selection on accelerating breeding

One of the primary benefits of GS is its potential to significantly reduce breeding cycles by allowing early selection based on genetic potential rather than waiting for full phenotypic evaluations. This is particularly valuable in yam, where conventional breeding cycles can span several years due to the plant's long growth period. GS enables rapid generation turnover by selecting genotypes with high breeding values in early stages, thus expediting the breeding process. Studies in other crops have demonstrated the effectiveness of GS in reducing breeding cycle time by as much as 50%, a benefit that can directly translate to yam breeding programs (Jighly et al., 2019). By minimizing cycle duration, GS makes it possible to develop yam varieties more quickly, which is essential for responding to changing agricultural demands and environmental challenges.

Furthermore, when GS is combined with other accelerated breeding techniques, such as speed breeding, the effect on breeding cycle reduction becomes even more pronounced. This combined approach has shown success in other crops, demonstrating that both GS and accelerated growth conditions (e.g., controlled photoperiods) can synergistically reduce the breeding cycle duration and speed up the introduction of new varieties (Watson, 2019).

8.2 Enhancement of genetic gain and acceleration of breeding speed

Genomic selection significantly enhances genetic gain by increasing the selection intensity and accuracy of breeding decisions. In traditional breeding, selection relies on observable traits, which are influenced by environmental variability. GS, however, uses genomic data to predict the genetic potential of individuals with high accuracy, thus enabling breeders to focus on candidates with the highest potential. This approach has proven successful in various crops, showing increased genetic gain rates compared to phenotypic selection. For example, studies in rice and wheat have reported up to a seven-fold increase in genetic gain when GS was incorporated alongside traditional breeding methods (Xu et al., 2019).

In yam breeding, GS can increase genetic gain by allowing the continuous selection of favorable alleles across multiple generations. This continuous selection helps accumulate advantageous traits, such as disease resistance and yield, at a faster rate than conventional methods. By integrating GS into yam breeding, it becomes feasible to achieve higher genetic gain per generation, which is critical for improving complex traits and adapting to environmental stresses. The increased genetic gain ultimately contributes to higher productivity and resilience in yam varieties, benefiting both breeders and farmers.

8.3 Case studies of successful yam breeding applications

Several case studies highlight the successful application of GS in plant breeding, underscoring its potential in yam. While specific yam-focused studies remain limited, research in other crops provides valuable insights into the potential benefits of GS in yam breeding. For instance, a study at the Bangladesh Rice Research Institute demonstrated that GS coupled with trait-specific marker-assisted selection reduced the breeding cycle by approximately 1.5 years and increased yield improvement by 117 kg per hectare per year—a seven-fold gain over baseline rates (Biswas et al., 2023). Similar approaches in yam breeding could lead to comparable advances, particularly in regions where yam is a staple food.

Another example is seen in macadamia breeding, where GS was shown to double genetic gain in nut yield by reducing the generation time from eight to four years (O'Connor et al., 2021). Applying these principles to yam, which has a similarly long breeding cycle, could expedite the development of high-yield, stress-resilient varieties. In each of these cases, the application of GS led to improved breeding efficiency, reduced time to market for new varieties, and enhanced genetic gain. These studies illustrate the promising impact of GS on breeding efficiency and underscore the potential for similar gains in yam breeding, where the combination of GS with advanced breeding techniques could revolutionize the development of improved varieties.

9 Concluding Remarks

The implementation of genomic selection (GS) in yam breeding has shown immense potential for accelerating breeding cycles, enhancing genetic gain, and making the breeding process more efficient. This concluding section summarizes the key findings of this research, insights for practical application in yam breeding, and recommendations for future research. This research has highlighted the transformative impact of GS on yam breeding efficiency. By using genome-wide markers to predict the breeding values of individual genotypes, GS allows breeders to make informed selection decisions earlier in the breeding cycle. This shift from phenotypic selection to genomic prediction has been shown to reduce the breeding cycle duration, increasing the pace at which new, high-yielding, and resilient yam varieties can be developed. Additionally, integrating GS with other technologies, such as speed breeding, further enhances genetic gain by allowing for rapid generation turnover and continuous selection of favorable alleles. These combined approaches hold the potential to significantly improve trait expression in yam, particularly for complex traits such as disease resistance and tuber quality.

For practical application in yam breeding, several strategies should be considered. First, optimizing marker density and sample size is crucial for maximizing the accuracy of GS models. Sufficiently dense markers ensure that the genetic variation across the genome is well-captured, while a large sample size provides a robust training population, increasing the reliability of genomic predictions. Additionally, integrating environmental data into GS models can help manage genotype-by-environment interactions, improving prediction accuracy in diverse cultivation conditions. This is particularly relevant for yam, which is grown across various ecological zones. Finally, the establishment of collaborative breeding programs that pool resources and data could help mitigate challenges such as limited marker availability and genotyping costs, making GS accessible to more yam breeding programs globally.

Future research in GS for yam breeding should focus on several key areas. First, the development of more sophisticated models that account for non-additive genetic effects and complex GxE interactions is necessary. Machine learning and deep learning models, though computationally intensive, could provide significant accuracy improvements for predicting complex traits in yam. Additionally, expanding genomic resources for yam, such as creating more comprehensive SNP arrays and reference genomes, would support more accurate GS models. Lastly, exploring the integration of new phenotyping technologies, such as hyperspectral imaging and remote sensing, could enhance trait measurement accuracy and improve the overall efficiency of yam breeding programs. By addressing these areas, future research can ensure that GS becomes a staple method in yam breeding, fostering the development of resilient, high-yield yam varieties suited to diverse environments and enhancing global food security.

Acknowledgments

The authors would like to thank the research group for their suggestions on my 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

Agre P., Norman P., Asiedu R., and Asfaw A., 2021, Identification of quantitative trait nucleotides and candidate genes for tuber yield and mosaic virus tolerance in an elite population of white guinea yam (*Dioscorea rotundata*) using genome-wide association scan, BMC Plant Biology, 21: 552.
https://doi.org/10.1186/s12870-021-03314-w

An S.F., Yu Y.L., Tan Z.W., Li C.M., Yang Q., Yang H.Q., Su X.Y., Sun Y., and Liang H.Z., 2024, Genomic analysis of yam: understanding its adaptive evolution and medicinal properties, Medicinal Plant Research, 14(1): 31-44.
https://doi.org/10.5376/mpr.2024.14.0003

Arruda M.P., Brown P., Lipka A., Krill A., Thurber C., and Kolb F., 2015, Genomic selection for predicting fusarium head blight resistance in a wheat breeding program, The Plant Genome, 8(3): plantgenome2015.01.0003.
https://doi.org/10.3835/plantgenome2015.01.0003

Asfaw A., Aderonmu D.S., Darkwa K., De Koeyer D., Agre P., Abe A., Olasanmi B., Adebola P., and Asiedu R., 2020, Genetic parameters, prediction, and selection in a white guinea yam early-generation breeding population using pedigree information, Crop Science, 61(2): 1038-1051.
https://doi.org/10.1002/csc2.20382

Begna T., 2022, Speed breeding to accelerate crop improvement, International Journal of Agricultural Science and Food Technology, 8(2): 178-186.
https://doi.org/10.17352/2455-815x.000161

Biswas P., Ahmed M., Afrin W., Rahman A., Shalahuddin A., Islam R., and Islam M.R., 2023, Enhancing genetic gain through the application of genomic selection in developing irrigated rice, Frontiers in Genetics, 14: 1083221.
https://doi.org/10.3389/fgene.2023.1083221

Cappetta E., Andolfo G., Di Matteo A., Barone A., Frusciante L., and Ercolano M.R., 2020, Accelerating tomato breeding by exploiting genomic selection approaches, Plants, 9(9): 1236.
https://doi.org/10.3390/plants9091236

Crossa J., Pérez-Rodríguez P., Cuevas J., Montesinos-López O., Jarquín D., de los Campos G., Burgueño J., González-Camacho J., Pérez-Elizalde S., Beyene Y., Dreisigacker S., Singh R., Zhang X., Gowda M., Roorkiwal M., Rutkoski J., and Varshney R., 2017, Genomic selection in plant breeding: methods, models, and perspectives, Trends in Plant Science, 22(11): 961-975.
https://doi.org/10.1016/j.tplants.2017.08.011

Das R.R., Vinayan M.T., Patel M., Phagna R.K., Singh S., Shahi J., and Zaidi P.H., 2020, Genetic gains with rapid-cycle genomic selection for combined drought and waterlogging tolerance in tropical maize (*Zea mays* L.), The Plant Genome, 13.
https://doi.org/10.1002/tpg2.20035

Diouf M., Zoclanclounon Y.A.B., Mboup P.A., Diouf D., Malédon E., Rivallan R., Chair H., and Dossa K., 2023, Genome-wide development of intra- and inter-specific transferable SSR markers and construction of a dynamic web resource for yam molecular breeding: Y2MD, The Plant Genome, 17(2): e20428.
https://doi.org/10.1101/2023.03.09.531889

Ge F., Jia C., Bao P., Wu X., Liang C., and Yan P., 2020, Accuracies of genomic prediction for growth traits at weaning and yearling ages in yak, Animals, 10(10): 1793.
https://doi.org/10.3390/ani10101793

Grenier C., Cao T., Ospina Y., Quintero C., Chatel M., Tohme J., Courtois B., and Ahmadi N., 2015, Accuracy of genomic selection in a rice synthetic population developed for recurrent selection breeding, PLoS ONE, 10(8): e0136594.
https://doi.org/10.1371/journal.pone.0136594

Heslot N., Yang H.P., Sorrells M., and Jannink J.L., 2012, Genomic selection in plant breeding: a comparison of models, Crop Science, 52(1): 146-160.
https://doi.org/10.2135/CROPSCI2011.06.0297

Jighly A., Lin Z., Pembleton L., Cogan N., Spangenberg G., Hayes B., and Daetwyler H., 2019, Boosting genetic gain in allogamous crops via speed breeding and genomic selection, Frontiers in Plant Science, 10: 1364.
https://doi.org/10.3389/fpls.2019.01364

Li Y., Kaur S., Pembleton L., Valipour-Kahrood H., Rosewarne G., and Daetwyler H., 2021, Strategies of preserving genetic diversity while maximizing genetic response from implementing genomic selection in pulse breeding programs, Theoretical and Applied Genetics, 135: 1813-1828.
https://doi.org/10.1007/s00122-022-04071-6

Lillehammer M., Meuwissen T., and Sonesson A., 2013, A low-marker density implementation of genomic selection in aquaculture using within-family genomic breeding values, Genetics, Selection, Evolution, 45: 39.
https://doi.org/10.1186/1297-9686-45-39

Lin Z., Cogan N., Pembleton L., Spangenberg G., Forster J., Hayes B., and Daetwyler H., 2016, Genetic gain and inbreeding from genomic selection in a simulated commercial breeding program for perennial ryegrass, The Plant Genome, 9(1): plantgenome2015.06.0046.
https://doi.org/10.3835/plantgenome2015.06.0046

Mondo J., Agre P., Asiedu R., Akoroda M., and Asfaw A., 2021, Optimized protocol for in vitro pollen germination in yam (*Dioscorea* spp.), Plants, 10(4): 795.
https://doi.org/10.3390/plants10040795

Mulualem T., Mekbib F., Shimelis H., Gebre E., and Amelework B., 2018, Genetic diversity of yam (*Dioscorea* spp.) landrace collections from ethiopia using simple sequence repeat markers, Australian Journal of Crop Science, 12(8): 1223-1230.
https://doi.org/10.21475/AJCS.18.12.08.PNE885

Norman A., Taylor J., Edwards J., and Kuchel H., 2018, Optimising genomic selection in wheat: effect of marker density, population size and population structure on prediction accuracy, G3: Genes|Genomes|Genetics, 8: 2889-2899. https://doi.org/10.1534/g3.118.200311

Norman P., Agre P., Asiedu R., and Asfaw A., 2022, Multiple-traits selection in white guinea yam (*Dioscorea rotundata*) genotypes, Plants, 11: 3003. https://doi.org/10.3390/plants11213003

Oakey H., Cullis B., Thompson R., Comadran J., Halpin C., and Waugh R., 2016, Genomic selection in multi-environment crop trials, G3: Genes|Genomes|Genetics, 6(5): 1313-1326. https://doi.org/10.1534/g3.116.027524

O'Connor K., Hayes B., Hardner C., Alam M., Henry R., and Topp B., 2021, Genomic selection and genetic gain for nut yield in an australian macadamia breeding population, BMC Genomics, 22: 370. https://doi.org/10.1186/s12864-021-07694-z

Reyna M., Grondona D., and Jarquín D., 2021, Development of a genomic prediction pipeline for maintaining comparable sample sizes in training and testing sets across prediction schemes accounting for the genotype-by-environment interaction, Agriculture, 11(10): 932. https://doi.org/10.3390/agriculture11100932

Sandhu K., Lozada D.N., Zhang Z., Pumphrey M., and Carter A., 2021, Deep learning for predicting complex traits in spring wheat breeding program, Frontiers in Plant Science, 11: 613325. https://doi.org/10.3389/fpls.2020.613325

Saski C., Bhattacharjee R., Scheffler B., and Asiedu R., 2015, Genomic resources for water yam (*Dioscorea* alata L.): analyses of EST-sequences, de novo sequencing and GBS libraries, PLoS ONE, 10(7): e0134031. https://doi.org/10.1371/journal.pone.0134031

Solberg T., Sonesson A., Woolliams J., and Meuwissen T., 2008, Genomic selection using different marker types and densities, Journal of Animal Science, 86(10): 2447-2454. https://doi.org/10.2527/jas.2007-0010

Tamiru M., Natsume S., Takagi H., White B., Yaegashi H., Shimizu M., Yoshida K., Uemura A., Oikawa K., Abe A., Urasaki N., Matsumura H., Babil P., Yamanaka S., Matsumoto R., Muranaka S., Girma G., Lopez-Montes A., Gedil M., and Terauchi R., 2017, Genome sequencing of the staple food crop white guinea yam enables the development of a molecular marker for sex determination, BMC Biology, 15(1): 86. https://doi.org/10.1186/s12915-017-0419-x

Wang W., 2024, The role of isoenzymatic variation in delineating phylogenetic relationships within Zea genus, Maize Genomics and Genetics, 15(1): 25-33. https://doi.org/10.5376/mgg.2024.15.0004

Watson A., Hickey L., Christopher J., Rutkoski J., Poland J., and Hayes B., 2019, Integrating genomic selection and speed breeding to increase genetic gain in spring wheat, Crop Science. https://doi.org/10.2135/CROPSCI2018.12.0757

Xu Y., Liu X., Fu J., Wang H., Wang J., Huang C., Prasanna B., Olsen M., Wang G., and Zhang A., 2019, Enhancing genetic gain through genomic selection: from livestock to plants, Plant Communications, 1: 100005. https://doi.org/10.1016/j.xplc.2019.100005

Zhao W., Lai X., Liu D., Zhang Z., Ma P., Wang Q., Zhang Z., Pan Y., 2020, Applications of support vector machine in genomic prediction in pig and maize populations, Frontiers in Genetics, 11: 598318. https://doi.org/10.3389/fgene.2020.598318