Biological Times Volume No. 5 | Issue No. 1 | Pages 7-8

Predictive Breeding: An Overview of Concepts, Tools, and Future Directions

Muhammad Arham*, Muhammad Talha Abdur Rehman, Hafiz Muhammad Hamza Aziz Chaudary and Marriam Riaz

Department of Plant Breeding and Genetics, University of Agriculture, Faisalabad

*Corresponding Author: **marham2203@gmail.com**

ABSTRACT

Plant breeding and genetics contribute significantly to meeting global food requirements, but the current threat of climate change, high population growth, and scarcity of natural resources require classical breeding to become faster, more accurate, and predictive. Predictive breeding is an emerging strategy that integrates genomic, phenotypic, and environmental data with statistical and computational models to predict plant performance prior to field testing. It helps breeders select favorable lines earlier, reducing the need for multiple seasons of field-based experiments. Despite some challenges related to the availability of data, transferability of models, and insufficient infrastructure, predictive breeding is increasingly being explored in modern crop improvement programs. This review highlights the basic concepts, key tools, and applications of predictive breeding.

Keywords: Predictive Breeding, Genomic Selection, Climate resilience, Prediction-based breeding

To cite this article: Arham M, MTA Rehman, HMHA Chaudary & M Riaz. Predictive Breeding: An Overview of Concepts, Tools, and Future Directions. Biological Times. 2026. January 5(1): 7-8.

Introduction

Plant Breeding and Genetics has been playing a vital role in securing global food production, since the selection made by prehistoric farmers, evolving through mendelian genetics to modern molecular breeding tools. Prehistoric domestication was primarily based on the perceptions of the farmer's visible attributes such as production levels, taste, susceptibility to diseases and adaptability to new conditions (1). These purely phenotype-based selections gradually shifted towards genotype-informed improvement after Mendel's laws of inheritance followed by their rediscovery with other key innovations in the mid to late nineteenth century (2) (3). In early 20th century, the first successful hybrid corn and the Green Revolution transforming global food production (4). By the end of 20th century, marker-assisted selection (MAS) as an emerging tool that offered breeders to select plants according to the DNA makers associated with important characteristics(5). But MAS is mainly effective for oligogenic traits but shows limited efficiency for polygenic traits. Moreover, it relies on biparental mapping populations with limited size and genetic diversity. (6)

Despite these advancements, modern plant breeding faces major challenges, including climate change, increasing population, and low availability of resources (1). Rising temperatures, droughts, floods, erratic weather patterns resulting in difficulties for ensuring global food security and sustainable agriculture. These challenges indicate that traditional and molecular breeding alone are insufficient to feed a massively growing population under limited resources, resulting in a solution which is plant breeding must be faster, more precise, and predictive (7) (8).

Predictive Breeding, Concepts and Fundamentals

Predictive breeding is a modern approach that integrates genomic, phenotypic, and environmental data with statistical and computational models to predict how plants will perform before they are tested in the field (9) (10). It helps breeders to identify the most favorable candidate line earlier, faster without multiple seasons of field evaluation.

The foundation of predictive breeding lies in combining three major data sources. Genomic information refers to genome wide DNA variation using dense molecular markers or sequencing, allowing both major and minor gene effects to be considered (11). Phenomics data is measured from plant physical traits using high drones, sensors and imaging systems (12). Environmental data is detailed information about Climate, soil, and management data of crop physical traits taken for every location (7). Statistical genomic selection models (9) as well as machine learning and deep learning models (13) are used to integrate all these data sets to predict the genetic potential for a trait, which is Genomic Estimated Breeding Value (GEBVs).

Predictive breeding uses complete genome information to predict complex quantitative traits (11) while MAS relies only on few specific DNA linked markers (6). Predictive breeding models utilize multi environmental variables, it estimates the genotype X environment (G x E) interaction that highly influences crop performance, beneficial for climate smart crop breeding (8). It is the beginning of the digital breeding in which artificial intelligence and other automation tools help breeders manage these data sets and improve selection procedures (10). This method will succeed with the help of capable tools that are described in the following section.

Tools and Technologies

Predictive breeding practically allows breeders to make selections faster and more accurately using a broad range of tools. These are categories such as genomics, high-throughput plant phenotyping, crop simulation models, and statistical or computational models as shown in Fig.1 (14).

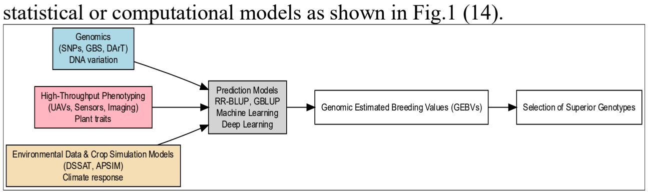

Figure 1: Conceptual framework of Predictive Breeding

Plants are genotyped using several genomics tools including single nucleotide polymorphism (SNP) arrays, Diversity Arrays Technology (DArT), and Genotyping-by-Sequencing (GBS), they provide the DNA level information of the plant (15). High-throughput plant phenotyping (HTPP) technologies i.e. drones (UAVs), ground sensors, and thermal and hyperspectral imaging system is used to measure plant physical traits (12) (16). DSSAT (17) and APSIM (18) are the crop simulation models that assume why plants show variations in their physiological behavior under environmental conditions, these tools are very beneficial to understand genotype X environment (G x E) interactions. Once these data are collected; statistical models (Ridge Regression Best Linear Unbiased Prediction and Genomic BLUP) (11), as well as machine learning (Random Forests and Support Vector Machines) (8), and deep learning models (13) are applied for prediction to estimate the GEBVs. These computational methods that make predictive breeding data driven. All these big data sets handled by bioinformatics integration and artificial intelligence (AI), further models analyzed these sets to support in practical application of predictive breeding (8).

Applications

Predictive breeding is adopted increasingly in breeding programs to improve yield and stress resistance across regions. Genomic selection reduces half of the time to develop cultivars in major crops using predictive breeding methods (9). It assists breeders to choose breeding lines timely without much expenditure and consumption of resources. It is very beneficial to develop climate resilient crop varieties, integration of environmental variables predicts genotypes under stress conditions (19) (7). It is helpful in countries that are mostly affected by climate change.

Challenges

Predictive breeding faces several challenges. It requires high quality and well-organized data, but limited availability of phenotypic data becomes challenging (12, 20). Moreover, its wide adaptability is also limited; models are trained in specific population or environment, may perform poorly in another due to variations in genetic makeup or climate. Advanced machine learning and deep learning models can capture complex and non-linear relationships, but they require large datasets (13). Moreover, predictive breeding is data driven so highly trained personnel, infrastructure needed, which are not equally available across institutions and countries (8, 21).

Future Perspectives

Apart from these limitations. Development in phenotypic tools and standard data collection protocols make it appropriate for availability of organized and accurate data. Improvements in bioinformatics platforms, and training programs for researchers lower the technical barrier. Efficient machine learning and deep learning algorithms improve the prediction accuracy with limited data availability. Multi-environmental and diverse population structures integration will increase the transferability of models. With these improvements, predictive breeding will make important breakthroughs in future crop breeding.

Conclusion

Predictive breeding is an innovative technique that supports breeders to predict the future of their desired genotype. It is very beneficial to address the upcoming food security problems and climate challenges by integrating all the modern tools and artificial intelligence (AI) techniques. Although some limitations still exist, further application of biology and technology is likely to make predictive breeding a major processing component in next generation plant breeding.

References

[1] Ceccarelli S. Efficiency of plant breeding. Crop science. 2015;55(1):87-97.
[2] Allard RW. Principles of plant breeding: John Wiley & Sons; 1999.
[3] Fehr W. Principles of cultivar development: theory and technique. 1991.
[4] Evenson RE, Gollin D. Assessing the impact of the Green Revolution, 1960 to 2000. science. 2003;300(5620):758-62.
[5] Collard BC, Mackill DJ. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Philosophical Transactions of the Royal Society B: Biological Sciences. 2008;363(1491):557-72.
[6] Heffner EL, Sorrells ME, Jannink JL. Genomic selection for crop improvement. Crop Science. 2009;49(1):1-12.
[7] Cooper M, Messina CD, Podlich D, Totir LR, Baumgarten A, Hausmann NJ, et al. Predicting the future of plant breeding: complementing empirical evaluation with genetic prediction. Crop and Pasture Science. 2014;65(4):311-36.
[8] Xu Y, Zhang X, Li H, Zheng H, Zhang J, Olsen MS, et al. Smart breeding driven by big data, artificial intelligence, and integrated genomic-enviromic prediction. Molecular Plant. 2022;15(11):1664-95.
[9] Crossa J, Pérez-Rodríguez P, Cuevas J, Montesinos-López O, Jarquín D, De Los Campos G, et al. Genomic selection in plant breeding: methods, models, and perspectives. Trends in plant science. 2017;22(11):961-75.
[10] Jeon D, Kang Y, Lee S, Choi S, Sung Y, Lee T-H, et al. Digitalizing breeding in plants: A new trend of next-generation breeding based on genomic prediction. Frontiers in Plant Science. 2023;14:1092584.
[11] Meuwissen TH, Hayes BJ, Goddard M. Prediction of total genetic value using genome-wide dense marker maps. genetics. 2001;157(4):1819-29.
[12] Araus JL, Cairns JE. Field high-throughput phenotyping: the new crop breeding frontier. Trends in plant science. 2014;19(1):52-61.
[13] Gao S, Yu T, Rasheed A, Wang J, Crossa J, Hearne S, et al. Fast-forwarding plant breeding with deep learning-based genomic prediction. Journal of Integrative Plant Biology. 2025.
[14] Fritsche-Neto R, Resende RT, Olivoto T, Garcia-Abadillo J, Nascimento M, Bahia MAM, et al. Prediction-based breeding: Modern tools to optimize and reshape programs. Crop Science. 2025;65(5):e70175.
[15] Poland JA, Rife TW. Genotyping-by-sequencing for plant breeding and genetics. The plant genome. 2012;5(3).
[16] Yang W, Feng H, Zhang X, Zhang J, Doonan JH, Batchelor WD, et al. Crop phenomics and high-throughput phenotyping: past decades, current challenges, and future perspectives. Molecular plant. 2020;13(2):187-214.
[17] Jones J, Hoogenboom G, PORTER C, BOOTE K, BATCHELOR W, HUNT L, et al. The DSSAT cropping system model. European Journal of Agronomy. 2003.
[18] Holzworth DP, Huth NI, deVoil PG, Zurcher EJ, Herrmann NI, McLean G, et al. APSIM–evolution towards a new generation of agricultural systems simulation. Environmental Modelling & Software. 2014;62:327-50.
[19] Ndlovu N. Application of genomics and phenomics in plant breeding for climate resilience. Asian Plant Res J. 2020:53-66.
[20] Fiorani F, Schurr U. Future scenarios for plant phenotyping. Annual review of plant biology. 2013;64(1):267-91.
[21] Hu H, Scheben A, Edwards D. Advances in integrating genomics and bioinformatics in the plant breeding pipeline. Agriculture. 2018;8(6):75.