Frontiers in Bioscience Elite

Front. Biosci. (Elite Ed) **2025**; 17(1): 27936
https://doi.org/10.31083/FBE27936

Review

Transformation of Plant Breeding Using Data Analytics and Information Technology: Innovations, Applications, and Prospective Directions

Mohammad Muzahidur Rahman Bhuiyan[1], Inshad Rahman Noman[2], Md Munna Aziz[1], Md Mizanur Rahaman[1,*], Md Rashedul Islam[1], Mia Md Tofayel Gonee Manik[1], Kallol Das[3]

[1]College of Business, Westcliff University, Irvine, CA 92614, USA
[2]Department of Computer Science, California State University, Los Angeles, CA 90032, USA
[3]College of Agriculture, Food and Environmental Sciences, California Polytechnic State University, San Luis Obispo, CA 93407, USA
*Correspondence: M.rahaman.392@westcliff.edu (Md Mizanur Rahaman)
Academic Editor: Baohong Zhang
Submitted: 6 November 2024 Revised: 17 December 2024 Accepted: 3 January 2025 Published: 23 January 2025

Abstract

Our study focused on plant breeding, from traditional methods to the present most advanced genetic and data-driven concepts. Conventional breeding techniques, such as mass selection and cross-breeding, have been instrumental in crop improvement, although they possess inherent limitations in precision and efficiency. Advanced molecular methods allow breeders to improve crops quicker by more accurately targeting specific traits. Data analytics and information technology (IT) are crucial in modern plant breeding, providing tools for data management, analysis, and interpretation of large volumes of data from genomic, phenotypic, and environmental sources. Meanwhile, emerging technologies in machine learning, high-throughput phenotyping, and the Internet of Things (IoT) provide real-time insights into the performance and responses of plants to environmental variables, enabling precision breeding. These tools will allow breeders to select complex traits related to yield, disease resistance, and abiotic stress tolerance more precisely and effectively. Moreover, this data-driven approach will enable breeders to use resources judiciously and make crops resilient, thus contributing to sustainable agriculture. Data analytics integrated into IT will enhance traditional breeding and other key applications in sustainable agriculture, such as crop yield improvement, biofortification, and climate change adaptation. This review aims to highlight the role of interdisciplinary collaboration among breeders, data scientists, and agronomists in absorbing these technologies. Further, this review discusses the future trends that will make plant breeding even more effective with this new wave of artificial intelligence (AI), blockchain, and collaborative platforms, bringing new data transparency, collaboration, and predictability levels. Data and IT-based breeding will greatly contribute to future global food security and sustainable food production. Thus, creating high-performing, resource-efficient crops will be the foundation of a future agricultural vision that balances environmental care. More technological integration in plant breeding is needed for resilient and sustainable food systems to handle the growing population and changing climate challenges.

Keywords: breeding techniques; data analytics; information technology; plant breeding; technologies

1. Introduction

Plant breeding has been a cornerstone of agriculture, continuously evolving to meet the demands of food security, environmental adaptability, and crop improvement. Initially, plant breeding relied heavily on traditional methods, such as mass selection and crossbreeding, where plants with favorable traits were selected for propagation [1]. These methods allowed for the gradual accumulation of beneficial characteristics within a plant population, although they required multiple generations to achieve desired outcomes. Mass selection, one of the earliest techniques, involved selecting and cultivating plants with desirable traits from a large population, while pure-line selection improved genetic uniformity by using single superior plants as the basis for new crops [2]. The development of Mendelian genetics in the 19th century marked a turning point in plant breeding by introducing genetic principles that allowed for the prediction of trait inheritance [3]. Mendel's discoveries laid the groundwork for quantitative genetics, which enabled breeders to focus on complex traits like yield and disease resistance, accelerating crop improvement [4]. The advent of hybridization techniques further revolutionized plant breeding in the 20th century by combining desirable traits from different parent plants, leading to the development of hybrids with enhanced productivity and resilience [5]. With the rise of biotechnology in the latter half of the 20th century, genetic transformation technologies, such as Agrobacterium-mediated transformation, allowing for the insertion of specific genes into plant genomes, achieving precise trait modification [6]. The introduction of genome editing technologies like clustered regularly interspaced short palindromic repeats and

Copyright: © 2025 The Author(s). Published by IMR Press.
This is an open access article under the CC BY 4.0 license.

Publisher's Note: IMR Press stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

CRISPR-associated protein 9 (CRISPR-Cas9) in the 21st century has added another layer of precision, enabling the targeted alteration of plant DNA without introducing foreign genes [7]. These advanced tools have drastically reduced the time required for trait development and offer unparalleled control over crop characteristics [8].

However, the integration of data analytics and information technology (IT) into plant breeding represents a transformative shift in the field, allowing for more precise, data-driven approaches to crop improvement. The use of big data, artificial intelligence (AI), and machine learning (ML) in plant breeding has enabled breeders to analyze complex datasets from genomic, phenotypic, and environmental sources, facilitating marker-assisted selection and genomic selection [9,10]. By leveraging large-scale data, breeders can now predict and identify desirable traits with greater accuracy, accelerating the breeding process and reducing the cost associated with traditional breeding cycles [11]. Information technology supports real-time data collection and analysis, particularly through high-throughput phenotyping platforms that capture extensive phenotypic data, allowing breeders to evaluate a vast number of plants under various environmental conditions [12]. Additionally, IT facilitates data storage and sharing, enabling collaboration across institutions and geographic boundaries, thus enhancing the scalability of breeding programs [13]. The combination of data analytics and IT in plant breeding has ushered in an era of precision breeding, where specific genetic markers associated with advantageous traits can be targeted, creating crops that are not only high-yielding but also resilient to biotic and abiotic stresses. By synthesizing recent innovations in plant breeding and the role of data-driven technologies, this review seeks to provide a comprehensive perspective on how merging breeding techniques with data analytics and IT can contribute to global food security, improve crop resilience, and foster environmentally sustainable farming. The insights from this review aim to guide future research and applications in data-driven plant breeding, helping to create adaptive and efficient crop breeding strategies suited to the demands of a rapidly changing world.

This review aims to provide an in-depth examination of advancements in plant breeding techniques, with a specific focus on the integration of data analytics and IT. The scope of the review encompasses both traditional and modern breeding methods, exploring how these methodologies have evolved to meet the challenges of contemporary agriculture. Key objectives include (i) examining the history and development of plant breeding techniques, from mass selection to genome editing; (ii) discussing the transformative role of data analytics and IT in accelerating breeding cycles and enhancing precision; and (iii) highlighting the applications and potential of these integrated approaches for sustainable agricultural practices.

2. Traditional and Advanced Plant Breeding Techniques

Traditional plant breeding methods have been the foundation of crop improvement for centuries, primarily involving mass selection and hybridization to enhance desirable traits. The progression of plant breeding techniques, from traditional methods like mass selection and cross-breeding to advanced methods such as CRISPR/Cas9 and synthetic biology (Fig. 1). Thus, the mass selection, one of the oldest methods, involves selecting a group of superior plants from a population based on observable characteristics, such as yield or resistance, and then cultivating seeds from these plants to enhance trait frequency in subsequent generations [1]. While effective for traits controlled by a few genes, the mass selection is limited by its low precision and is typically slow, often requiring multiple generations to achieve significant improvements [14]. Hybridization is another traditional breeding method that gained prominence in the 20th century and involves cross-breeding two genetically distinct parents to combine advantageous traits in their offspring [5]. Hybrid vigor, or heterosis, is frequently observed in these crosses, resulting in hybrids with higher productivity, disease resistance, or stress tolerance compared to their parents [4]. However, hybridization can be labor-intensive and time-consuming, as it may take several generations to stabilize desired traits within a new hybrid population (Fig. 2). Moreover, both mass selection and hybridization are limited by their reliance on the natural genetic variation within a species, restricting the potential for introducing entirely new traits. The development of molecular biology and biotechnology has introduced a new era of modern breeding techniques that include genetic transformation, genome editing, and marker-assisted selection (Fig. 1). These techniques enable breeders to manipulate plant genomes directly, allowing for precise trait selection that is often impossible to achieve through traditional breeding alone. Genetic transformation, particularly Agrobacterium-mediated transformation, allows for the insertion of specific genes into plant genomes, bypassing the need for conventional breeding cycles [6]. This technique has enabled the development of genetically modified (GM) crops with traits such as herbicide tolerance, pest resistance, and improved nutritional content, which has had a significant impact on global agriculture [15]. While effective, genetic transformation has faced public resistance and regulatory scrutiny in many regions, limiting its widespread adoption.

Furthermore, genome editing technologies, especially CRISPR-Cas9, have further refined genetic manipulation by allowing targeted modifications to specific genes within a plant's DNA [16]. Unlike genetic transformation, genome editing can make precise alterations without introducing foreign DNA, making it both a faster and more publicly acceptable method for crop improvement [7]. Applications of CRISPR in plant breeding include enhancing disease resis-

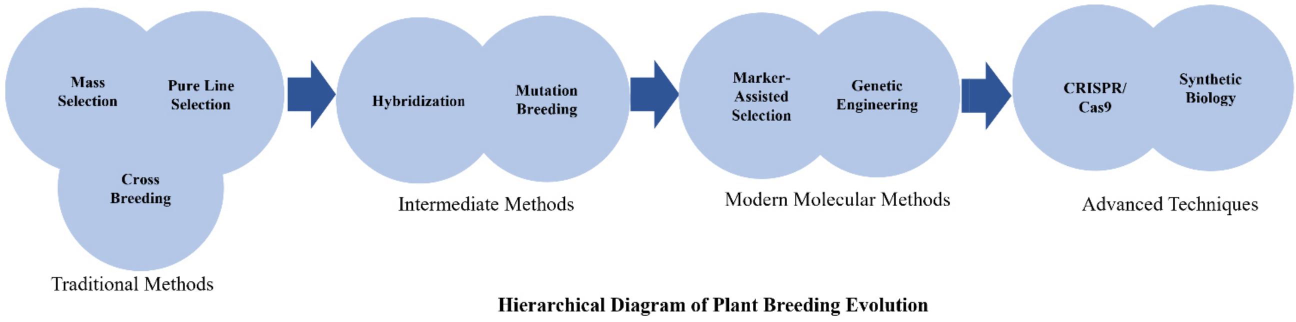

Fig. 1. Showing the hierarchical diagram of plant breeding evolution. CRISPR, clustered regularly interspaced short palindromic repeats. Created using PowerPoint.

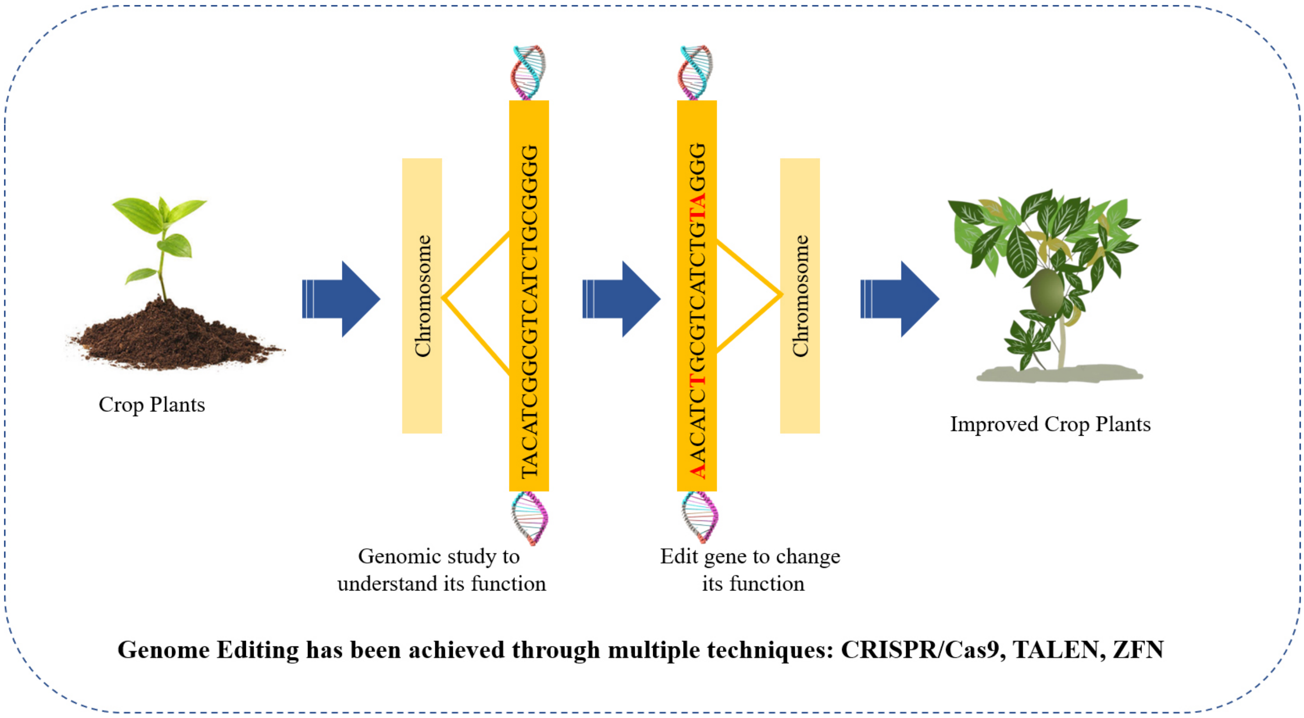

Fig. 2. Using genome editing through multiple techniques for crop improvement. TALEN, transcription activator-like effector nucleases; ZFN, zinc-finger nuclease. Created using PowerPoint.

tance, improving stress tolerance, and increasing nutritional quality, with significant potential to address challenges associated with climate change and food security [8].

However, the integration of data analytics and information technology (IT) in plant breeding has further enhanced the precision and efficiency of modern breeding techniques. Data science, powered by big data and machine learning algorithms, enables breeders to analyze vast datasets from genomic, phenotypic, and environmental sources to optimize breeding decisions and predict trait performance with high accuracy [17]. For instance, genomic selection is highly dependent on statistical models that leverage large datasets to predict an individual plant's performance based on its genotype, a process made more efficient through IT-driven analytics [18]. Gene editing involved with marker-assisted selection (MAS) and genomic selection (GS) represents additional modern techniques that use molecular markers to track and select desirable traits [19]. MAS involves identifying and selecting specific DNA markers that are associated with traits of interest, thereby accelerating the breeding process and increasing selection accuracy [9]. Genomic selection, a more recent advancement, uses statistical models to predict the breeding value of plants based on their genetic profiles, enhancing efficiency for traits influenced by multiple genes [20]. Together, these modern breeding techniques have enabled more rapid and precise development of crop varieties suited to the demands of modern agriculture.

High-throughput phenotyping (HTP), enabled by IT and data analytics, allows breeders to collect and analyze phenotypic data on a large scale, providing valuable insights into trait expression under different environmental conditions [21]. IT tools also facilitate the real-time monitoring and analysis of breeding trials, allowing breeders to make data-driven selections that are less susceptible to environmental variation. Through machine learning and AI, data analytics can identify complex patterns in breeding data, improving the ability to select traits such as yield, disease resistance, and stress tolerance with greater precision [22].

The combination of advanced breeding techniques with data analytics and IT has led to a new paradigm known as precision breeding, where digital tools and genetic insights converge to create optimized crop varieties tailored to specific agricultural needs. This integration not only enhances the speed and accuracy of breeding but also contributes to sustainable agriculture by enabling the development of resilient, high-performing crops suited to the challenges of modern farming.

3. Data Analytics and Information Technology in Plant Breeding

Data analytics plays a crucial role in modern plant breeding by leveraging big data, data mining, and statistical analysis to identify and predict beneficial crop traits. The primary steps in plant genetic transformation and genome editing include large construct transformation, gene expression regulation, organelle transformation, transgene removal and confinement, and genome editing (Fig. 3). These steps aim to improve precision, safety, and efficiency in plant biotechnology. Big data analytics allows breeders to analyze vast datasets collected from genomic, phenotypic, and environmental sources, facilitating informed decision-making in crop improvement [18]. For instance, by analyzing large-scale genomic data, breeders can identify molecular markers associated with desirable traits such as disease resistance or drought tolerance, thereby accelerating the selection process [23]. Additionally, data mining techniques enable breeders to extract meaningful patterns from complex datasets, aiding in the identification of gene-trait associations and optimizing breeding strategies for specific environments [24]. With the rapid advancements in high-throughput sequencing and phenotyping, data analytics has become essential for managing and interpreting the vast amount of data generated in breeding programs. Statistical models such as genomic selection rely heavily on data analytics to predict the breeding value of individual plants, allowing breeders to make accurate selections based on genotype information. This approach has proven effective in improving complex traits controlled by multiple genes, thus enhancing the accuracy of crop improvement programs [22].

Information technology (IT) has become integral to plant breeding by supporting data management, real-time monitoring, and high-throughput phenotyping. IT infrastructure facilitates the storage, retrieval, and sharing of vast datasets, enabling collaboration across institutions and geographic regions [12]. Data management platforms, such as integrated databases and cloud computing, allow breeders to store and access genomic and phenotypic data efficiently, promoting efficient workflow and knowledge sharing [25]. Real-time monitoring tools in plant breeding leverage sensors, remote sensing technologies, and IoT (Internet of Things) devices to gather continuous data on environmental conditions and plant performance [21]. These technologies enable breeders to monitor crop growth and detect stress conditions in real-time, providing insights into plant responses under diverse environments. Furthermore, high-throughput phenotyping platforms, often powered by IT, capture phenotypic data on a large scale, helping breeders to evaluate traits like yield, biomass, and stress resilience quickly and accurately [26].

Machine learning (ML) and artificial intelligence (AI) have introduced a new era of precision breeding by enabling predictive modeling and data-driven decision-making. Predictive models built using ML algorithms analyze patterns in genomic and phenotypic data, allowing breeders to make accurate predictions about plant performance and identify optimal parent combinations for crossbreeding [27]. For example, random forest and support vector machines are popular algorithms that help predict complex traits by analyzing data from previous breeding cycles [17]. AI tools, such as deep learning, facilitate the analysis of high-dimensional data, such as images from high-throughput phenotyping, to assess plant health and growth characteristics [28]. By automating trait evaluation, AI reduces the time and cost associated with phenotyping and increases the accuracy of trait selection. The integration of ML and AI in plant breeding also enhances resource efficiency by enabling precise applications of water, fertilizers, and other inputs based on predicted plant needs, supporting sustainable agricultural practices [29]. The convergence of data analytics, IT, and AI in plant breeding underscores the shift towards precision agriculture, where technology-driven insights inform every stage of crop development. This approach enables breeders to address complex challenges, such as climate change and food security, by creating resilient and high-performing crop varieties suited to specific environments.

4. Applications of Modern Plant Breeding Enhanced by Data Analytics and IT

The various objectives of plant breeding focus on enhancing traits like disease resistance, higher yield, improved quality, and stress tolerance (Fig. 4). These objectives aim to optimize plant growth, adaptability, and productivity under diverse environmental conditions. Data analytics plays a critical role in enhancing crop yield and optimizing resource efficiency, particularly by enabling breeders to identify traits that maximize productivity with minimal input. Through the analysis of big data from field trials, breeders can assess and select high-yielding varieties

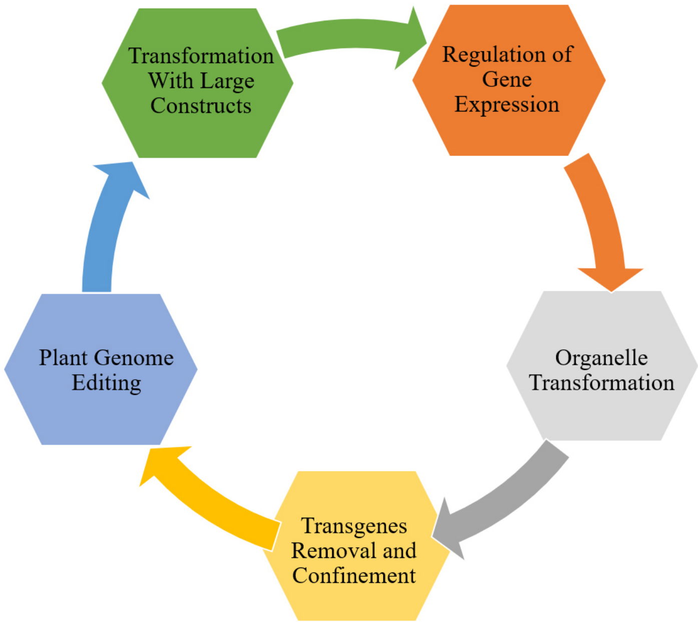

Fig. 3. Key components of plant genetic transformation and genome editing. Created using PowerPoint.

that require less water, fertilizer, and other inputs [18]. By combining data on plant genotypes with phenotypic outcomes, predictive models help breeders pinpoint genetic markers associated with traits like drought tolerance and nitrogen-use efficiency, which can lead to crops that require fewer inputs for optimal yield [23]. Additionally, precision agriculture tools supported by data analytics allow for real-time monitoring of crop health and growth, enabling farmers to apply resources only where needed, thereby improving resource efficiency and reducing waste [24]. Abiotic stresses, such as drought and extreme temperatures, alongside biotic stresses from pests and diseases, significantly impact crop productivity. Modern plant breeding enhanced by data analytics and IT offers innovative solutions to these challenges by identifying genetic traits linked to stress resilience. For example, genomic selection models use environmental data and genomic information to predict plant performance under specific stress conditions, enabling breeders to develop stress-tolerant varieties faster and more effectively. Technologies like CRISPR and machine learning assist in identifying and editing genes associated with pest and disease resistance, providing crops that are more resilient to pathogens without the need for chemical interventions [22,30,31].

Furthermore, real-time data collection via remote sensing and IoT devices allows breeders to monitor stress responses in crops and make data-driven decisions that improve plant resilience to various stressors [21]. Data-driven approaches are essential for enhancing the nutritional quality of crops and achieving biofortification. By integrating data from molecular, phenotypic, and environmental sources, breeders can identify genetic markers linked to nutrient content, such as vitamins and minerals, and select for these traits in new crop varieties [9]. For instance, bioinformatics tools aid in the analysis of genetic pathways that influence the synthesis of micronutrients, allowing breeders to develop crops with higher levels of iron, zinc, and other essential nutrients [32]. Data analytics also enables breeders to track nutrient retention across generations, ensuring that biofortified traits remain stable. Such advancements contribute to the development of nutritionally superior crops that address global malnutrition challenges [24].

The integration of data analytics and IT in plant breeding supports sustainability goals by enabling the development of crops that are more adaptable to changing environmental conditions and less reliant on chemical inputs. Using big data from climate models, breeders can predict future growing conditions and develop varieties adapted to specific ecological niches [23]. Furthermore, genomic tools enable the selection of traits associated with soil health and carbon sequestration, supporting sustainable practices that enhance environmental resilience [12]. Precision breeding approaches driven by machine learning optimize crop management, reducing the environmental footprint of agriculture by minimizing inputs like water and pesticides [17]. By fostering crops that are both productive and ecologically sustainable, data-driven plant breeding contributes to a resilient agricultural system that aligns with environmental and societal needs.

5. Key Areas of Innovation and Fine-Tuning Breeding Technologies

Advancements in data-driven technologies have transformed genetic transformation techniques, allowing for precise and targeted modifications that were previously challenging. Genetic transformation, which involves intro-

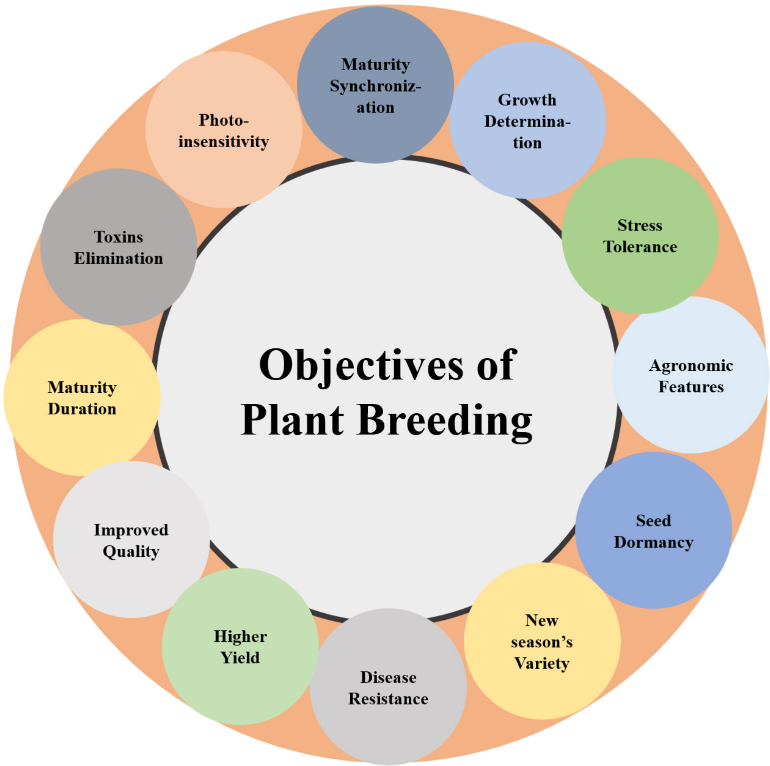

Fig. 4. Applications of plant breeding to fulfill different objectives. Created using PowerPoint.

ducing new genes into plant genomes to impart desirable traits, has traditionally used methods like Agrobacterium-mediated transformation and particle bombardment [6]. However, these methods often lacked precision, resulting in unintended genetic changes. The integration of data analytics and IT, particularly bioinformatics, now allows breeders to predict gene expression outcomes and improve transformation efficiency by selecting optimal gene insertion sites [33]. Moreover, CRISPR-Cas9 and other genome-editing tools have added another layer of precision to genetic transformation, enabling specific modifications at targeted genome sites. CRISPR-based methods combined with data analytics help breeders identify candidate genes associated with advantageous traits, optimize guide RNA design, and predict off-target effects, significantly improving transformation success [8]. Through the use of computational tools, these techniques reduce the risk of unintended genetic disruptions, making genetic transformation more accurate and acceptable for regulatory frameworks [34]. Data-driven approaches have enhanced mass selection by facilitating the identification of genetic markers linked to desirable traits, allowing for more targeted selection while maintaining genetic variability within a population [35]. Using data analytics and high-throughput phenotyping, breeders can now evaluate thousands of plants simultaneously, selecting individuals with optimal genetic profiles that also retain genetic diversity, which is essential for crop adaptability and resilience [24]. Statistical and machine learning models assist in assessing phenotypic traits across different environmental conditions, streamlining the mass selection process and enabling breeders to focus on specific traits while ensuring that other beneficial traits are not lost [23].

Furthermore, despite significant advancements, several technical, regulatory, and environmental challenges limit the full integration of data analytics and IT into plant breeding. Technical challenges include the high costs of infrastructure, such as high-throughput phenotyping platforms, and the need for specialized expertise in bioinformatics and data science [12]. Moreover, data integration remains challenging, as phenotypic, genomic, and environmental data often come from different sources and require standardized formats and analysis tools for meaningful insights [17]. Regulatory challenges also impact the adoption of advanced breeding technologies. For instance, regulatory bodies in different regions have varying guidelines for genetically modified organisms (GMOs) and gene-edited crops, complicating the approval process for novel traits developed using genetic transformation and genome editing [34]. Additionally, the public perception around genetic modification and privacy concerns regarding the use of big data in breeding can hinder adoption and innovation [36].

Environmental challenges are significant, as the deployment of genetically modified or precision-bred crops may have ecological impacts, such as gene flow to wild relatives and effects on biodiversity [37]. Therefore, developing strategies to minimize these impacts is essential for the responsible use of these technologies. As plant breeding continues to advance, addressing these challenges through improved regulatory frameworks, technical solutions, and public engagement will be critical to ensuring sustainable and widespread adoption.

6. Challenges and Limitations of Data-Driven Plant Breeding

The implementation of data-driven approaches in plant breeding faces several technical and infrastructural limitations, primarily related to data integration, high costs, and skill gaps (Fig. 5). Data integration is a major hurdle, as breeding programs generate massive amounts of data from diverse sources, including genomic, phenotypic, and environmental datasets, often in disparate formats. Integrating and analyzing these heterogeneous datasets requires sophisticated bioinformatics infrastructure, which is frequently inaccessible to smaller breeding programs [9]. Furthermore, the high costs associated with high-throughput phenotyping platforms, sequencing technologies, and computational tools can be prohibitive, especially for developing countries and resource-limited institutions [23].

Moreover, there is a significant skill gap in bioinformatics and data science within the plant breeding sector, which limits the ability of breeding programs to leverage data-driven technologies fully. Advanced data analytics requires specialized expertise in genomics, data science, and machine learning, which are not traditionally part of plant breeders' training [17]. To address these limitations, investments in capacity building and infrastructure, as well as collaborations between breeding institutions and technology providers, are essential. Data-driven plant breeding raises important ethical and regulatory concerns, particularly related to data privacy, intellectual property rights, and the ethical use of gene editing. Data privacy is a significant issue, as breeding data often includes sensitive genetic information that must be handled responsibly. Inadequate data governance frameworks may lead to unauthorized access and misuse of proprietary breeding data [38]. Intellectual property (IP) rights also play a crucial role in data-driven breeding. Innovations, such as genetically engineered and gene-edited crops, are frequently subject to IP restrictions, potentially limiting access for small-scale breeders and farmers [36].

Additionally, the ethical implications of gene editing technologies like CRISPR have prompted regulatory scrutiny. Concerns over potential off-target effects and the long-term impact of gene editing on ecosystems and food systems lead to strict regulations in some regions, which can delay or prevent the deployment of advanced breeding technologies [34]. As data-driven breeding expands, developing globally consistent regulatory frameworks that balance innovation with ethical considerations will be critical.

Data-driven plant breeding must address potential environmental concerns, particularly regarding biodiversity loss and unintended ecological impacts. Genetic diversity is essential for crop resilience and adaptability, but the focus on developing high-yielding, data-optimized varieties can lead to genetic homogeneity within crops, making them vulnerable to pests, diseases, and changing climates [39]. This loss of diversity could reduce the resilience of agricultural systems, as monocultures are more susceptible to widespread losses from biotic and abiotic stresses [40]. There are also concerns over the unintended ecological impacts of genetically modified or gene-edited plants, such as gene flow to wild relatives and potential disruptions to local ecosystems [37]. As precision breeding and genetic transformation techniques become more widespread, it is crucial to monitor and manage these ecological risks. Implementing best practices, such as promoting genetic diversity in breeding programs and conducting thorough environmental impact assessments, can help mitigate these risks while enabling the responsible use of data-driven technologies.

7. Future Directions and Prospects

As data-driven plant breeding evolves, several emerging technologies, including artificial intelligence (AI), Internet of Things (IoT), and blockchain, are paving the way for new methodologies and efficiencies in crop improvement. AI, particularly machine learning and deep learning, offers predictive capabilities that can revolutionize trait selection by identifying complex patterns in genetic and phenotypic data, thus expediting the breeding process [28]. IoT, through remote sensing and smart sensors, enables real-time monitoring of plant growth and environmental conditions, providing breeders with valuable data that can inform breeding decisions based on real-time insights [41]. Blockchain technology, although still emerging in agriculture, holds the potential for data transparency and traceability, enabling secure and decentralized data sharing across breeding programs, which could enhance collaboration and data integrity [42]. These technologies offer promising prospects for accelerating breeding efficiency and ensuring data security and accessibility. As these technologies become more integrated into breeding workflows, they have the potential to create a connected and intelligent breeding ecosystem that supports faster, more precise crop improvement. The future of data-driven plant breeding hinges on interdisciplinary collaboration across fields such as genomics, bioinformatics, environmental science, and computer science. Collaborative research can address the complexity of agricultural systems by combining expertise from diverse disciplines, allowing for a holistic approach to crop breeding challenges [35]. For instance, partnerships between plant scientists and computer scientists can lead to the development of more robust machine-learning algorithms tailored to agricultural data, improving predictions of trait performance and resilience under various conditions [43].

Moreover, international collaboration is also critical in pooling resources and sharing knowledge across regions. Projects like the International Wheat Genome Sequencing

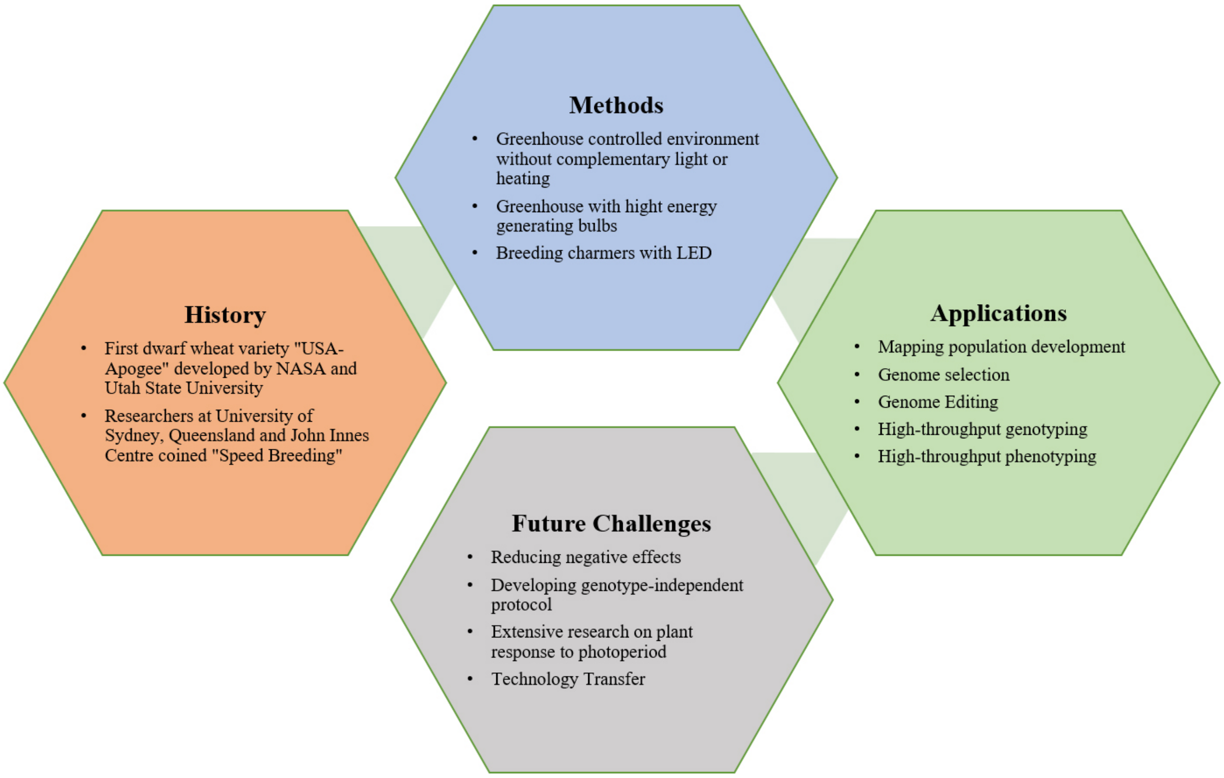

Fig. 5. Current methods, applications, and challenges of plant breeding. LED, light-emitting diode. Created using PowerPoint.

Consortium (IWGSC) and the Consultative Group on International Agricultural Research (CGIAR) platforms for crop improvement exemplify how global collaboration can drive innovations that benefit plant breeding worldwide [44]. Increased investment in research and development and fostering partnerships between academia, industry, and government institutions can enhance the scope and impact of data-driven breeding technologies. Data-driven plant breeding holds significant promise for addressing sustainable agriculture and global food security challenges [45,46]. Enhanced breeding technologies can contribute to sustainability by creating crops that require fewer inputs, such as water and fertilizers, while maintaining high productivity. This is critical for adapting to climate change, where resilient crop varieties can mitigate the impacts of extreme weather events [23]. Furthermore, precision breeding enables the development of biofortified crops that improve nutritional quality, supporting food security initiatives that address malnutrition in low-resource regions [32]. By aligning data-driven breeding efforts with sustainable agriculture goals, breeders can develop crops that are not only high-yielding but also environmentally resilient. This approach can help meet the United Nations' Sustainable Development Goals (SDGs) for zero hunger and responsible consumption and production by creating a balance between high productivity and environmental stewardship [47]. As the global population continues to rise, data and IT-enhanced breeding will be essential for creating a sustainable and secure food system for future generations.

8. Conclusion

Data analytics and information technology (IT) have fundamentally transformed modern plant breeding, enabling unprecedented precision, efficiency, and adaptability in crop improvement. These technologies facilitate the analysis of vast genomic, phenotypic, and environmental datasets, allowing breeders to identify and enhance beneficial traits that were once difficult to capture through traditional methods. By integrating high-throughput phenotyping, machine learning algorithms, and genomic tools, data-driven breeding accelerates the development of resilient and productive crop varieties. This shift has implications far beyond the breeding process itself; it supports the creation of crops that can withstand environmental stresses, contribute to food security, and promote sustainable agricultural practices. The role of data analytics and IT in modern breeding demonstrates how technology can address complex agricultural challenges in a climate-stressed world. The future of plant breeding will depend on continued innovation and interdisciplinary collaboration. As more advanced technologies such as artificial intelligence (AI), the Internet of Things (IoT), and blockchain become available, their integration into breeding will further optimize trait selection and data management, improving decision-making and breeding efficiency. AI and machine learning are already transforming trait prediction, while IoT devices enable real-time monitoring of crop health and environmental conditions, providing invaluable insights for breeders. Meanwhile, blockchain has the potential to secure data sharing and enhance collaboration across institutions by creating transparent, decentralized breeding data networks. To fully realize these benefits, breeders, data scientists, agronomists, and regulatory bodies must work together to bridge the gap between technology and agriculture, fostering a collaborative environment where expertise from diverse fields converges for sustainable innovation. As the global population continues to grow, the demand for sustainable food systems has never been more urgent. Future plant breeding must not only focus on increasing yield but also consider the broader environmental impact of agriculture. Data-driven breeding can play a crucial role in achieving this goal by developing crop varieties that require fewer resources, such as water, fertilizers, and pesticides, and are more resilient to climate change impacts. By supporting the United Nations' Sustainable Development Goals (SDGs), data-enhanced breeding can contribute to sustainable agricultural practices that balance productivity with ecological stewardship. This approach aligns with the vision of resilient agriculture that not only meets food demands but also respects natural ecosystems.

Looking forward, the integration of technology in agriculture offers exciting possibilities for developing a resilient, data-informed agricultural landscape. Plant breeding will increasingly rely on precision-based methodologies, driven by advanced analytics and real-time data, to adapt to rapidly changing environmental conditions. By embedding sustainability into breeding objectives, we can work toward a future where agriculture is both productive and environmentally harmonious. In this vision, data analytics and IT do not merely serve as tools for improvement; they redefine the scope and potential of plant breeding, transforming it into a discipline that addresses global food security and environmental resilience through science and technology. Ultimately, the convergence of breeding technologies with data analytics and IT paves the way for a sustainable, adaptive agricultural system capable of thriving in an uncertain future.

Author Contributions

MMRB, IRN, MMA, and MMR designed the research study and wrote an initial draft. MMRB, and MRI performed the research, MMR supervised research plan and editing manuscript. KD and MMTGM supported figure drawing, writing initial draft, and helped in resources. All authors contributed to editorial changes in the manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work. All authors read and approved the final manuscript.

Ethics Approval and Consent to Participate

Not applicable.

Acknowledgment

We would like to express our gratitude to all the anonymous reviewers for their contribution and critical reviews to improve the manuscript.

Funding

This research received no external funding.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] Allard RW. Principles of plant breeding. 2nd edn. John Wiley & Sons: USA. 1999.

[2] Hayes HK, Immer FR, Smith DC. Methods of plant breeding. McGraw-Hill: USA. 1955.

[3] Mendel G. Versuche über Pflanzenhybriden. Verhandlungen des naturforschenden Vereines in Brünn. 1866; 4: 3–47.

[4] Hallauer AR, Carena MJ, Miranda Filho JB. Quantitative genetics in maize breeding. Springer: Germany. 2010.

[5] Sleper DA, Poehlman JM. Breeding field crops. 5th edn. Blackwell Publishing: UK. 2006.

[6] Gelvin SB. Agrobacterium-mediated plant transformation: the biology behind the "gene-jockeying" tool. Microbiology and Molecular Biology Reviews: MMBR. 2003; 67: 16–37. https://doi.org/10.1128/MMBR.67.1.16-37.2003.

[7] Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science. 2014; 346: 1258096. https://doi.org/10.1126/science.1258096.

[8] Gao C. Genome engineering for crop improvement and future agriculture. Cell. 2021; 184: 1621–1635. https://doi.org/10.1016/j.cell.2021.01.005.

[9] Varshney RK, Roorkiwal M, Sorrells ME. Genomic selection for crop improvement: challenges and opportunities. The Plant Genome. 2019; 12: 180108. https://doi.org/10.1080/21645698.2015.1129937.

[10] Jeon D, Kang Y, Lee S, Choi S, Sung Y, Lee TH, *et al*. Digitalizing breeding in plants: A new trend of next-generation breeding based on genomic prediction. Frontiers in Plant Science. 2023; 14: 1092584. https://doi.org/10.3389/fpls.2023.1092584.

[11] Nerkar G, Devarumath S, Purankar M, Kumar A, Valarmathi R, Devarumath R, Appunu C. Advances in Crop Breeding Through Precision Genome Editing. Frontiers in Genetics. 2022; 13: 880195. https://doi.org/10.3389/fgene.2022.880195.

[12] Pauli D, Chapman SC, Bart R, Topp CN, Lawrence-Dill CJ, Poland J, *et al*. The Quest for Understanding Phenotypic Variation via Integrated Approaches in the Field Environment. Plant Physiology. 2016; 172: 622–634. https://doi.org/10.1104/pp.16.00592.

[13] Xu L, Naylor D, Dong Z, Simmons T, Pierroz G, Hixson KK, *et al*. Drought delays development of the sorghum root microbiome and enriches for monoderm bacteria. Proceedings of the National Academy of Sciences of the United States of America. 2018; 115: E4284–E4293. https://doi.org/10.1073/pnas.1717308115.

[14] Acquaah G. Principles of plant genetics and breeding. John Wiley & Sons: USA. 2012.

[15] James C. Global status of commercialized biotech/GM crops: 2013. ISAAA Brief No. 46. International Service for the Acquisition of Agri-biotech Applications (ISAAA). 2013. Available at: https://www.isaaa.org/ (Accessed: 30 October 2024).

[16] Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease

in adaptive bacterial immunity. Science. 2012; 337: 816–821. https://doi.org/10.1126/science.1225829.

[17] Hickey JM, Chiurugwi T, Mackay I, Powell W, Implementing Genomic Selection in CGIAR Breeding Programs Workshop Participants. Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery. Nature Genetics. 2017; 49: 1297–1303. https://doi.org/10.1038/ng.3920.

[18] Crossa J, Pérez P, Hickey J, Burgueño J, Ornella L, Cerón-Rojas J, *et al*. Genomic prediction in CIMMYT maize and wheat breeding programs. Heredity. 2014; 112: 48–60. https://doi.org/10.1038/hdy.2013.16.

[19] Xu Y, Crouch JH. Marker-assisted selection in plant breeding: from publications to practice. Crop Science. 2008; 48: 391–407. https://doi.org/10.2135/cropsci2007.04.0191.

[20] Heffner EL, Sorrells ME, Jannink JL. Genomic selection for crop improvement. Crop Science. 2009; 49: 1–12. https://doi.org/10.2135/cropsci2008.08.0512.

[21] Araus JL, Cairns JE. Field high-throughput phenotyping: the new crop breeding frontier. Trends in Plant Science. 2014; 19: 52–61. https://doi.org/10.1016/j.tplants.2013.09.008.

[22] Spindel JE, McCouch SR. When more is better: how data sharing would accelerate genomic selection of crop plants. The New Phytologist. 2016; 212: 814–826. https://doi.org/10.1111/nph.14174.

[23] Xu Y, Li P, Zou C, Lu Y, Xie C, Zhang X, *et al*. Enhancing genetic gain in the era of molecular breeding. Journal of Experimental Botany. 2017; 68: 2641–2666. https://doi.org/10.1093/jxb/erx135.

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

[25] Fahlgren N, Gehan MA, Baxter I. Lights, camera, action: high-throughput plant phenotyping is ready for a close-up. Current Opinion in Plant Biology. 2015; 24: 93–99. https://doi.org/10.1016/j.pbi.2015.02.006.

[26] Yang W, Duan L, Chen G, Xiong L, Liu Q. Plant phenomics and high-throughput phenotyping: accelerating plant breeding and crop improvement. Frontiers in Plant Science. 2020; 11: 516919. https://doi.org/10.1016/j.pbi.2013.03.005.

[27] Battenfield SD, Guzmán C, Gaynor RC, Singh RP, Peña RJ, Dreisigacker S, *et al*. Genomic Selection for Processing and End-Use Quality Traits in the CIMMYT Spring Bread Wheat Breeding Program. The Plant Genome. 2016; 9: 1–12. https://doi.org/10.3835/plantgenome2016.01.0005.

[28] Kamilaris A, Prenafeta-Boldú FX. Deep learning in agriculture: a survey. Computers and Electronics in Agriculture. 2018; 147: 70–90. https://doi.org/10.1016/j.compag.2018.02.016.

[29] van Dijk ADJ, Kootstra G, Kruijer W, de Ridder D. Machine learning in plant science and plant breeding. iScience. 2020; 24: 101890. https://doi.org/10.1016/j.isci.2020.101890.

[30] Das K, Ayim BY, Borodynko-Filas N, Das SC, Aminuzzaman F. Genome editing (CRISPR/Cas9) in plant disease management: challenges and future prospects. Journal of Plant Protection Research. 2023; 63: 159–172. https://doi.org/10.24425/jppr.2023.145761.

[31] Rani S, Das K, Aminuzzaman F, Ayim BY, Borodynko-Filas N. Harnessing the future: cutting-edge technologies for plant disease control. Journal of Plant Protection Research. 2023; 63: 387–398. https://doi.org/10.24425/jppr.2023.147829.

[32] Bouis HE, Saltzman A. Improving nutrition through biofortification: A review of evidence from HarvestPlus, 2003 through 2016. Global Food Security. 2017; 12: 49–58. https://doi.org/10.1016/j.gfs.2017.01.009

[33] Chen K, Wang Y, Zhang R, Zhang H, Gao C. CRISPR/Cas Genome Editing and Precision Plant Breeding in Agriculture. Annual Review of Plant Biology. 2019; 70: 667–697. https://doi.org/10.1146/annurev-arplant-050718-100049.

[34] Metje-Sprink J, Menz J, Modrzejewski D, Sprink T. DNA-free genome editing: past, present and future. Frontiers in Plant Science. 2020; 11: 112. https://doi.org/10.3389/fpls.2018.01957.

[35] Ronald PC. Enhancing sustainable development through plant genetics. Nature Reviews Genetics. 2023; 24: 659–660. https://doi.org/10.1038/s41576-023-00651-4.

[36] Ewa WG, Agata T, Milica P, Anna B, Dennis E, Nick V, *et al*. Public perception of plant gene technologies worldwide in the light of food security. GM Crops & Food. 2022; 13: 218–241. https://doi.org/10.1080/21645698.2022.2111946.

[37] Snow AA, Andersen B, Jørgensen RB. Costs of transgenic herbicide resistance introgression from crops to wild relatives: gene flow impacts on fitness and adaptation. Molecular Ecology. 2005; 14: 2347–2358. http://dx.doi.org/10.1046/j.1365-294x.1999.00596.x.

[38] Ågren J, Oakley CG, Lundemo S, Schemske DW. Adaptive divergence in flowering time among natural populations of Arabidopsis thaliana: Estimates of selection and QTL mapping. Evolution. 2017; 71: 550–564. https://doi.org/10.1111/evo.13126.

[39] Ewens WJ. James F. Crow and the stochastic theory of population genetics. Genetics. 2012; 190: 287–290. https://doi.org/10.1534/genetics.111.135194.

[40] Esquinas-Alcázar J. Science and society: protecting crop genetic diversity for food security: political, ethical and technical challenges. Nature Reviews. Genetics. 2005; 6: 946–953. https://doi.org/10.1038/nrg1729.

[41] Sundmaeker H, Verdouw C, Wolfert S, Perez Freire L. Internet of food and farm 2020. In Digitising the industry (pp. 129–150). River Publishers: Denmark. 2016.

[42] Tripoli M, Schmidhuber J. Emerging opportunities for the application of blockchain in the agri-food industry. Food and Agriculture Organization of the United Nations (FAO). 2018. Available at: https://openknowledge.fao.org/handle/20.500.14283/ca9934en (Accessed: 30 October 2024).

[43] Benos L, Tagarakis AC, Dolias G, Berruto R, Kateris D, Bochtis D. Machine Learning in Agriculture: A Comprehensive Updated Review. Sensors. 2021; 21: 3758. https://doi.org/10.3390/s21113758.

[44] International Wheat Genome Sequencing Consortium (IWGSC). Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science. 2018; 361: eaar7191. https://doi.org/10.1126/science.aar7191.

[45] Yoosefzadeh Najafabadi M, Hesami M, Eskandari M. Machine Learning-Assisted Approaches in Modernized Plant Breeding Programs. Genes. 2023; 14: 777. https://doi.org/10.3390/genes14040777.

[46] Farooq MA, Gao S, Hassan MA, Huang Z, Rasheed A, Hearne S, *et al*. Artificial intelligence in plant breeding. Trends in Genetics: TIG. 2024; 40: 891–908. https://doi.org/10.1016/j.tig.2024.07.001.

[47] United Nations. Transforming our world: the 2030 agenda for sustainable development. 2015. Available at: https://sdgs.un.org/2030agenda (Accessed: 30 October 2024).