Artificial intelligence in plant biotechnology: Advances in gene editing, stress tolerance, and predictive breeding

Kordrostami M.[1]; Ghasemi-Soloklui A.A.[1]; Moori S.[1]; Rahimi M.[2*]

Received: October 2025 Accepted: January 2026

Abstract

Artificial intelligence (AI) has become a powerful tool in plant biotechnology, transforming applications ranging from multi-omics data analysis to genome engineering, predictive breeding, and precision agriculture. The rapid expansion of biological datasets has created an increasing demand for computational models capable of identifying complex patterns, capturing nonlinear relationships, and predicting plant responses under diverse environmental conditions. This review synthesizes recent advances in machine learning and deep learning approaches for genomic, transcriptomic, proteomic, and metabolomic analyses, with particular emphasis on their integration into plant biotechnology. It highlights the growing role of AI in improving the precision and efficiency of gene-editing technologies, particularly CRISPR-based systems, and examines recent developments in digital phenotyping, machine vision, and automated detection of biotic and abiotic stresses. The review further explores AI-driven predictive breeding, multimodal genotype–environment modeling, plant digital twins, and systems biology approaches for modeling gene regulatory and metabolic networks. In addition, major challenges—including data quality, model interpretability, dataset bias, ethical considerations, and biosecurity issues—are critically discussed. Finally, future perspectives on computational genome design, robotic agriculture, autonomous breeding pipelines, and integrated AI-enabled crop management systems are presented. Overall, the convergence of AI and plant biotechnology offers significant opportunities to accelerate the development of resilient, high-yielding, and climate-adaptive crop varieties for sustainable agriculture.

Keywords: Artificial intelligence, Plant biotechnology, CRISPR gene editing, Digital phenotyping, Predictive breeding

---

1-Nuclear Agriculture Research School, Nuclear Science and Technology Research Institute (NSTRI), Karaj, Iran
2-Department of Biotechnology, Institute of Science and High Technology and Environmental Sciences, Graduate University of Advanced Technology, Kerman, Iran.
*Corresponding author's Email: mehdi83ra@gmail; me.rahimi@kgut.ac.ir

Introduction

The convergence of artificial intelligence (AI) and plant biotechnology is transforming our strategies on understanding, engineering, and optimizing plant systems for food security, climate resilience, and sustainable agriculture (Gao *et al*., 2025; Xu *et al*., 2025). The exponential growth of multi-omics data, including genomics, transcriptomics, proteomics, and metabolomics, has generated an urgent need for computational methods to extract meaningful biological knowledge from increasingly complex and high-dimensional datasets over the last decade (Yagaliyeva *et al*., 2025; Edukondalu *et al*., 2026). Conventional statistical approaches, although beneficial, are frequently inadequate to address the nonlinear associations, epistatic interactions, and dynamic regulatory networks characterizing plant responses to environmental stresses (Cai *et al*., 2025). In this context, deep learning has become a powerful tool to uncover latent patterns, predict molecular phenotypes and provide unprecedented guidance for genome editing (Raihan *et al*., 2026). Recent advances in convolutional neural networks and foundation models have enabled researchers to decode complex plant genomic landscapes, predict key genes associated with agronomic traits and estimate genetic values from compressed genome-wide polymorphisms while dealing with challenges such as polyploidy and highly repetitive sequences (Cai *et al*., 2025; Xu *et al*., 2025; Raihan *et al*., 2026). AI is transforming plant biotechnology, from the speed of genetic gain in crop plants through AI-guided genomic selection to digital plant phenotyping in the face of changing climatic and anthropogenic conditions (Yagaliyeva *et al*., 2025; Edukondalu *et al*., 2026). This review offers a holistic overview of the recent progress in AI-powered plant biotechnology, emphasizing multi-omics integration, deep learning-driven genomic prediction, precision breeding, digital phenotyping, and the nascent concept of plant foundation models (Gao *et al*., 2025; Xu *et al*., 2025). We also present a critical review of the persistent challenges such as the “black-box” nature of deep learning models, data heterogeneity, limitations of small-sample learning, and barriers to real-world deployment. We offer a forward-looking perspective on how explainable AI, inclusive data governance, and intelligent breeding loops will drive next-generation climate-resilient crop design (Cai *et al*., 2025; Amar *et al*., 2026).

A new perspective on artificial intelligence in plant biotechnology

Over the past decade, artificial intelligence has become one of the most fundamental driving forces behind progress in the life sciences, particularly in plant biotechnology, which demands speed, accuracy, and the management of massive datasets (Bhardwaj *et al.,* 2022). The exponential growth of multi-omics data, the rapid development of gene-editing tools, and the increasing complexity of biotic and abiotic stresses under climate change conditions have led scientists toward tools capable of revealing hidden patterns, multi-faceted relationships, and accurate predictions (Varadharajan *et al.,* 2025). In this context, AI—as a powerful toolbox ranging from machine learning to deep learning—has bridged the gap between massive datasets and practical knowledge, opening new pathways for understanding and engineering plant performance (Rane *et al.,* 2024) (Fig. 1).

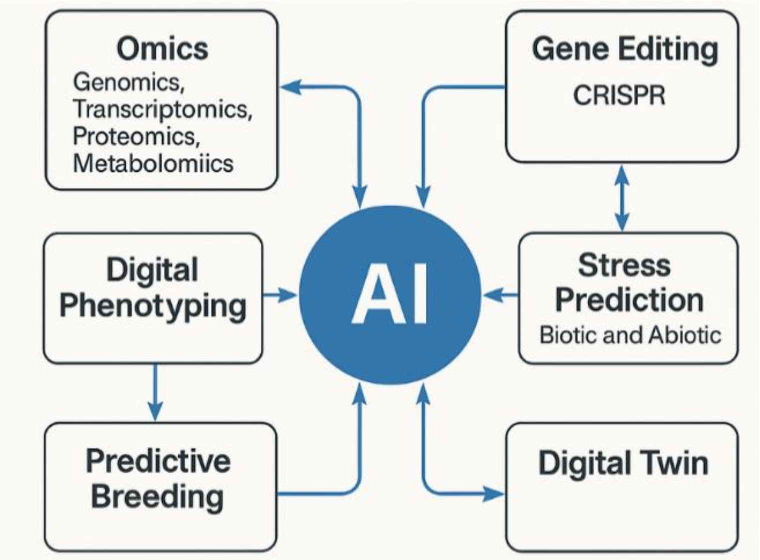

Figure 1: Conceptual map of the role of AI in plant biotechnology, showing connections among omics, gene editing, digital phenotyping, stress prediction, predictive breeding, and digital twin systems.

In the past, many plant biotechnology processes—including genetic selection, identification of genes involved in stress tolerance, phenotype analysis, and metabolic pathway design—were performed manually, empirically, and through trial and error (Stewart Jr, 2025). Today, however, the emergence of advanced data analysis models and deep neural networks has made it possible to identify extremely complex patterns in biological systems at scales previously impossible (Peng *et al.,* 2018). This shift is not merely about increasing analysis speed; it fundamentally changes the nature of research: from retrospective analyses to predictive analyses (Zaorsky *et al.,* 2022). In fact, AI can now predict the behavior of genes, molecular networks, metabolic pathways, and plant responses to stresses before they occur—and this is precisely the point at which science advances to the level of "biological intelligence" (Murmu *et al.,* 2024).

One of the most important reasons for AI's dominance in plant biotechnology is its adaptability to the multi-dimensional nature of biological data (Niazian and Niedbała, 2020). Genomic, transcriptomic, metabolomic, phenotypic images, spectroscopy, remote sensing, and climatic data all have multi-layered, non-linear structures (Liang *et al.,* 2025). Traditional statistical models have limited ability to analyze such data, but convolutional neural networks (CNNs), transformer models, XGBoost, Random Forest algorithms, and generative models can uncover hidden and non-obvious relationships (Abdulkareem *et al.,* 2024). This capability has opened new horizons for identifying target genes, precise genome editing, discovery of stress-tolerance markers, and design of engineered metabolic pathways (Mipeshwaree Devi *et al.,* 2023).

Simultaneously, AI has played a significant role in transforming plant measurement and observation methods (Tripodi *et al.,* 2022). The emergence of digital phenomics—including hyperspectral, thermal, 3D imaging, and LiDAR—has generated an unprecedented volume of phenotypic information (Mansoor *et al.,* 2025). Analysis of these data is only possible with intelligent algorithms capable of determining plant shape, growth, health, chlorophyll content, water status, and stress responses with exceptional accuracy (Islam *et al.,* 2024). Consequently, AI has not only automated plant monitoring but also enabled the creation of decision support systems for intelligent farm management and breeding programs (Su, 2020).

In the field of gene editing, AI is now a critical component of guide RNA design in CRISPR technology (Abbaszadeh and Shahlai, 2025). Prediction of off-targets, design of single-guide RNAs with highest efficiency, simulation of potential mutations, and even suggestion of genome modification pathways—these are now performed by machine learning models (Chuai *et al.,* 2018). This represents a fundamental transformation: gene editing has evolved from a trial-and-error process to an intelligent, predictable, and precise process (Dixit *et al.,* 2024a). Moreover, combining AI with gene regulatory network modeling has enabled researchers to predict the effects of each edit on molecular pathways, final phenotype, and environmental adaptation without performing wet-lab experiments (Xu *et al.,* 2022).

One of the most exciting emerging approaches is the development of plant digital twins—a virtual model capable of simulating growth, metabolism, yield, and stress responses under different conditions (Chen *et al.,* 2023). Such an approach, now expanding with AI assistance, allows researchers to simulate multiple climatic, management, and genetic scenarios and select the optimal breeding or management pathway. A plant digital twin can even predict the future performance of a genotype under climate change—a topic that will play a key role in the future of smart agriculture (Peladarinos *et al.,* 2023).

Despite all these advances, challenges remain. Shortages of high-quality data, heterogeneity of data sources, risk of bias in models, complexity of biological networks, and limitations in interpreting deep models are among the obstacles that must be addressed (Ching *et al.,* 2018). Additionally, ethical issues related to automated decision-making, data protection, and biosecurity are of particular importance (De Haro, 2024). Nevertheless, global trends indicate that the role of AI in plant biotechnology is not only expanding but will become the central core of designing intelligent plants, optimized genomes, and next-generation agricultural systems in the coming years.

Machine learning and deep learning models in omics data analysis

The rapid growth of sequencing technologies and molecular measurement tools has led the life sciences to face an explosion of omics data in the last decade (Dai and Shen, 2022). Genomic, transcriptomic, proteomic, metabolomic, and epigenomic data are now stored as massive, multi-dimensional, and complex datasets—datasets whose analysis using traditional statistical methods has become nearly impossible (Srivastava *et al.,* 2024). In such circumstances, AI—especially machine learning and deep learning—has become a key tool for extracting meaning, discovering hidden patterns, and generating predictive models in plant sciences. These technologies not only have the ability to analyze massive volumes of omics data but can also reconstruct complex molecular interactions and pathways in biological systems with unprecedented accuracy (Kalita *et al.,* 2025) (Fig. 2).

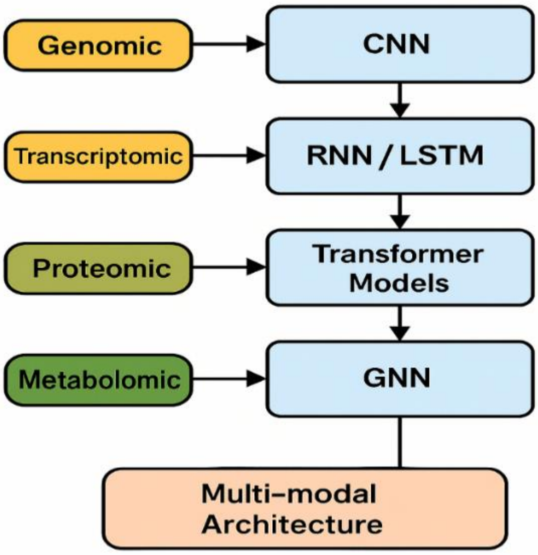

Figure 2: Architecture of machine learning and deep learning models used in omics data analysis, including convolutional neural networks (CNNs), recurrent and LSTM networks, transformer models, graph neural networks (GNNs), and multi-modal architectures for processing genomic, transcriptomic, proteomic, and metabolomic data.

Foundation models (FMs) are becoming a paradigm in computational plant biology, referring to neural networks pre-trained on large unlabeled biological datasets using self-supervised learning, which can then be fine-tuned for diverse downstream tasks with minimal additional supervision (Bommasani *et al.*, 2021; Xu *et al.*, 2025). Plant-specific FMs like GPN, AgroNT, Plant Caduceus, and PlantRNA-FM have demonstrated great performance in predicting gene functions, regulatory elements, and mutation effects in various plant species (Xu *et al.*, 2025; Wang *et al.*, 2025b). These models solve problems often seen in plant genomes like polyploidy, high repetitive sequence content and environment-responsive regulatory elements (Farooq *et al.*, 2024). The future development of plant foundation models is likely to integrate sequence analysis, structure prediction and functional annotation into integrated computational frameworks that will greatly accelerate hypothesis generation and experimental prioritization (Xu *et al.*, 2025).

In genome analysis, machine learning plays an important role in identifying key genes, predicting their function, and finding genetic markers associated with target traits (Mahood *et al.,* 2020). Algorithms such as Random Forest, Support Vector Machines (SVM), XGBoost, and transformer models have been widely used for gene classification, identification of functional regions of the genome, and prediction of mutation effects (Shea *et al.,* 2023).

Table 1 provides a comparative overview of these and other AI architectures commonly employed in plant biotechnology, summarizing their primary applications, strengths, and limitations. One major development is the ability of machine learning to model non-linear relationships and epistatic interactions—relationships that often remain hidden from traditional methods (Tang *et al.,* 2020). This capability has made it possible to identify genes that interactively contribute to stress tolerance, growth, or production of specific metabolites (Shaik and Ramakrishna, 2014).

In transcriptomics, the analysis of gene expression patterns has been transformed by AI. Previously, RNA-seq data analysis relied mainly on statistical tests, but today deep neural networks can detect co-expression patterns, stress-activated pathways, and complex regulatory networks with higher accuracy (Yuan and Bar-Joseph, 2019). Models such as Autoencoders, Long Short-Term Memory (LSTM) networks, and Graph Neural Networks (GNNs) enable the analysis of time-structured data, dynamic pathways, and complex gene networks (Otal *et al.,* 2025). GNNs, in particular, have proven highly effective in understanding network structures such as gene co-expression networks because they can analyze gene interactions as nodes and edges in a graph, thereby providing a more precise picture of biological system behavior (Lazaros *et al.,* 2024).

In proteomics and metabolomics, the challenges are more complex because measurements in these fields vary greatly in accuracy, noise, and data volume (Smith *et al.,* 2014). Nevertheless, deep learning models have been able to reveal hidden patterns in mass spectrometry, chromatography, and metabolite image data (Liebal *et al.,* 2020). CNNs have been very successful in detecting spectrometry peaks, separating noise, and identifying proteins (Lin *et al.,* 2019). Furthermore, generative models such as Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs) can complete missing data, reconstruct unknown metabolic patterns, and even suggest new pathways for metabolic engineering (Alwani, 2024).

Table 1: Comparative analysis of artificial intelligence models in plant biotechnology.

Model Architecture	Primary applications in plant biotechnology	Key strengths	Key limitations	Representative references
Convolutional neural networks (CNNs)	Disease detection from leaf images; hyperspectral image analysis; phenotyping	Excellent spatial feature extraction; translation invariance; mature ecosystem (ResNet, EfficientNet)	Requires large labeled datasets; limited temporal modeling; high computational cost for 3D data	Mohanty *et al*., 2016; Kamilaris and Prenafeta-Boldú, 2018
Graph neural networks (GNNs)	Gene regulatory network inference; protein interaction modeling; metabolic pathway analysis	Naturally handles graph-structured biological data; captures relational dependencies; interpretable node-level predictions	Scalability challenges with very large graphs; over smoothing in deep architectures	Wu *et al*., 2020; Lazaros *et al*., 2024
Transformer models	Genome sequence analysis; multi-omics integration; protein structure prediction (AlphaFold)	Captures long-range dependencies; attention mechanism provides interpretability; excellent for sequential data	Very high computational requirements; needs massive pre-training data	Choi and Lee, 2023; Xu *et al*., 2025
Long short-term memory (LSTM) / RNNs	Time-series stress response prediction; growth trajectory modeling; climate impact forecasting	Handles variable-length temporal sequences; captures temporal dependencies	Vanishing gradient issues for very long sequences; slower training than Transformers	Wen *et al*., 2023; Bi *et al*., 2023
Generative adversarial networks (GANs)	Synthetic multi-omics data generation; missing data imputation; novel genome sequence design	Generates realistic synthetic data; unsupervised learning capabilities	Training instability; mode collapse; difficult evaluation metrics	Alwani, 2024
Variational autoencoders (VAEs)	Dimensionality reduction of omics data; latent representation learning; anomaly detection	Probabilistic latent space; smooth interpolation between samples; good for visualization	Blurrier generations than GANs; posterior collapse issues	Smith *et al*., 2014
Random forest / XGBoost	Feature importance ranking; genotype	High interpretability; handles mixed data	Limited for unstructured data	Shaik and Ramakrishna,

Model Architecture	Primary applications in plant biotechnology	Key strengths	Key limitations	Representative references
	prediction; marker discovery	types; robust to overfitting; low computational cost	(images); less accurate for very complex non-linear relationships	2014; Abdulkareem *et al*., 2024
Foundation models (e.g., GPN, AgroNT, PlantCaduceus)	Multi-task plant genomics; cross-species transfer learning; regulatory element prediction	Exceptional generalization; pre-trained on massive data; adaptable to many downstream tasks	Extremely high computational requirements; data-hungry; limited plant-specific pre-training data	Xu *et al*., 2025

One key application of AI in omics is multi-omics integration—a process in which genomic, transcriptomic, proteomic, metabolomic, and epigenomic data are analyzed simultaneously to obtain an integrated view of plant function (Cembrowska-Lech *et al.,* 2023). This integration is very difficult with traditional methods because each data layer has different structure, scale, and noise (Dong and Rekatsinas, 2018). However, AI algorithms—especially multi-modal models and deep neural networks—have enabled the combination of these layers (Ahmed *et al.,* 2023). For example, multi-input transformer models can simultaneously analyze DNA sequences, gene expression patterns, and metabolic profiles and discover relationships among them. This approach is highly effective in identifying active cellular pathways under stress, discovering complex quantitative trait loci (QTLs), and designing plant breeding strategies (Choi and Lee, 2023).

On the other hand, AI plays an essential role in accelerating the discovery of genes associated with biotic and abiotic stresses. The complex patterns indicating tolerance to drought, salinity, heat, or disease are often embedded in multi-gene interactions and multi-layered regulatory pathways (Zhang *et al.,* 2024). DL models can extract these patterns and identify genes that contribute most to plant stability (Wang *et al.,* 2025b). Meanwhile, machine learning is highly effective in ranking gene importance and shows researchers which genes should be targeted for editing or selection.

AI-based analyses have also transformed Genome-Wide Association Studies (GWAS) (Xie and Wang, 2025). In traditional GWAS, the relationship between SNPs and traits was usually examined linearly, whereas AI can also consider non-linear relationships and multi-way genetic interactions (Alamin *et al.,* 2022). Deep neural networks in GWAS can now identify influential SNPs with greater accuracy, especially for complex, polygenic traits (Bellot *et al.,* 2018). Another strength of AI in omics is the ability to predict genome and phenotype function based on molecular data. AI models can predict which genomic changes will lead to improved growth, yield, or resistance. This capacity is vital for areas such as designing new genotypes, planning experiments, and even guiding CRISPR technology (Yetgin, 2025).

Artificial intelligence in gene editing: Enhancing CRISPR precision and targeted gene design

The CRISPR-Cas technology has created an unprecedented revolution in plant genome editing over the past decade. However, despite its high ability to create targeted mutations, two fundamental challenges remain: prediction of off-target regions and precise design of guide RNAs to achieve high editing efficiency (Vats *et al.,* 2019). The emergence of AI—particularly machine learning and deep learning algorithms—has transformed these fundamental challenges and created a new wave of "intelligent gene editing". Today, AI models can not only predict the molecular behavior of the CRISPR system but also suggest new editing pathways and have brought guide RNA design to a level of precision previously thought impossible (Balajishanmugam, 2025) (Fig. 3).

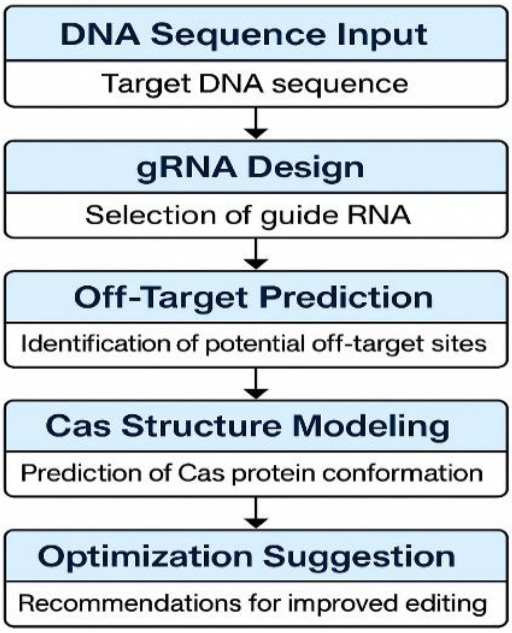

Figure 3: Intelligent flowchart of an AI-based gene editing system, including DNA sequence input, guide RNA design, off-target site prediction, Cas protein structure modeling, and provision of optimization suggestions to increase editing precision and efficiency.

One of the main strengths of AI in the CRISPR field is its ability to analyze DNA sequences at the whole-genome scale (Kim *et al.,* 2025). Models such as CNNs and transformer networks can learn sequence patterns associated with successful cleavage by Cas9 or Cas12, mutation rates, and the probability of unwanted cuts. These models are trained on millions of experimental examples and can predict guide RNA efficiency with much higher accuracy than rule-based methods (Fong and Wong, 2023). Tools such as Deep CRISPR, Elevation, CRISPR-Net, and CRISPRon are examples of AI-based systems that have dramatically increased the accuracy of gene knockouts or insertions (Dixit *et al.,* 2024a).

In guide RNA design, machine learning models have been highly effective by ranking influential features—such as RNA secondary structure, GC content, PAM position, and surrounding sequences. These models allow researchers to select the most suitable option with the highest probability of success from among thousands of potential sequences (Abbaszadeh and Shahlai, 2025). This is especially important in plant systems with complex genome structures—such as hexaploidy wheat or other polyploid species. AI can also predict the potential effects of each edit at the molecular and phenotypic levels. Deep learning models can analyze gene networks and metabolic pathways associated with the target gene and show what consequences deletion, addition, or alteration of a sequence will have for the cell, tissue, or whole plant. This capability is valuable because many complex traits (e.g., drought tolerance, heat resistance, or production of secondary metabolites) depend on a network of genes, and editing a single gene will be successful only if its effect on the entire system is understood (Farooq *et al.,* 2024).

Another application of AI in gene editing is predicting the structure and molecular behavior of Cas proteins. Models such as AlphaFold and RoseTTAFold can now simulate the three-dimensional structures of Cas9 or Cas12 with high accuracy. These simulations allow researchers to identify motifs responsible for enzyme specificity or efficiency and even design engineered versions of Cas with greater accuracy or better compatibility with plants. This is particularly important for improving CRISPR performance in species with complex or resistant genomes (Pan *et al.,* 2025).

In recent years, a new generation of algorithms known as AI-guided genome engineering has been developed that can fully design genome modification pathways (Li *et al.,* 2025). These models can suggest which genes should be edited, what types of changes are desirable, and even what combinations of edits can produce the best phenotype. This approach, which will revolutionize the future of plant breeding, essentially assigns the role of "genome designer" to AI—that is, the intelligent system can suggest the optimal version of a genome for a given goal (e.g., salinity resistance or increased biomass) (Malone *et al.,* 2010).

One of the most important applications of AI in gene editing is the identification of off-targets—a serious challenge for CRISPR technology. AI models can predict regions prone to unwanted cuts with high accuracy by analyzing sequence features, chromatin structure, DNA accessibility, and Cas binding patterns. Deep neural networks such as CRISPR-Net have reduced off-target rates by more than 50%. This advance is particularly important for gene editing in plants where even a small unwanted change can affect growth, metabolism, or resistance pathways (Wang *et al.,* 2025a).

The combination of AI with CRISPR has also accelerated the development of new gene-editing systems. For example, AI plays a vital role in designing highly accurate guide RNAs in RNA-editing systems (CRISPR-Cas13). Machine learning models can analyze complex RNA structures and RNA-Cas13 interactions and identify patterns that are not detectable by experimental methods. These applications have opened new pathways for controlling gene expression in plants, which can be highly effective in managing transient stresses or modifying temporary traits (Abbaszadeh and Shahlai, 2025). Furthermore, AI plays a significant role in the development of prime editing and base editing systems. These emerging technologies, which have higher precision than nuclease-based CRISPR cleavage, can optimize guide RNAs, prime editing guide RNAs (pegRNAs), or molecular triggers for each genomic location with the help of AI. This level of precision has made point edits and precise modifications more reliable for plant applications (Jiang *et al.,* 2025).

Predictive systems for biotic and abiotic stress tolerance in plants

The interaction of plants with environmental stresses—both biotic (diseases, pests, viruses) and abiotic (drought, salinity, heat, cold, nutrient deficiency)—is one of the most complex dimensions of plant biology (Dixit *et al.,* 2024b). Environmental stresses are usually multi-factorial, dynamic, and time-dependent, and the plant response is the result of a series of network interactions among genes, hormonal pathways, metabolic pathways, and physiological processes. This complexity has made analysis of plant resistance and prediction of its performance very difficult with traditional methods. However, the emergence of AI has been able to overcome these limitations and create a new generation of stress tolerance prediction systems—systems capable of integrating multi-source data, discovering hidden patterns, and predicting the final plant response with high accuracy (Hoque *et al.,* 2020; Hallajian *et al.,* 2024) (Fig. 4).

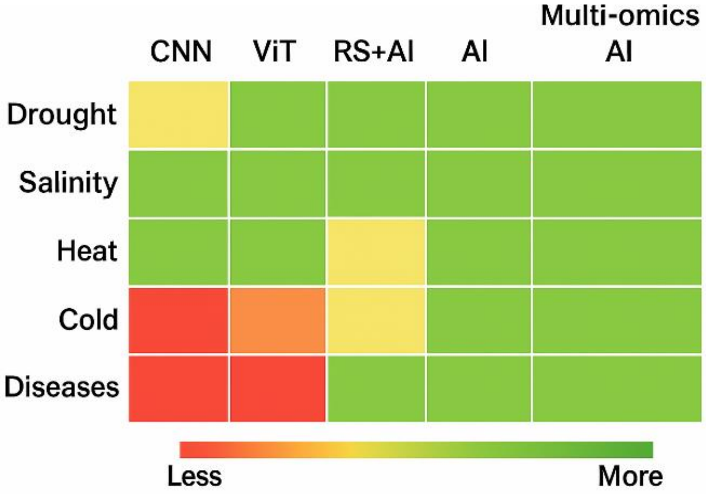

Figure 4: Heat map of the application of AI techniques in managing various plant stresses, including drought, salinity, heat, cold, and diseases, against models such as CNNs, Vision Transformers (ViT), remote sensing-based AI systems (RS+AI), intelligent multi-omics analysis, and digital twin-based predictions.

The basic structure of stress prediction systems is based on multi-layered data. These systems typically use a combination of genomic, transcriptomic, metabolomic, phenotypic, climatic, and remote sensing data (Dai *et al.,* 2024). In the past, analysis of these massive and heterogeneous datasets was not possible, but with the development of machine learning and deep learning algorithms, these layers can now be integrated into multi-modal models. Multi-input transformer models, multi-channel CNNs, and GNNs can extract complex patterns of plant performance under stress conditions without the need for handcrafted rules.

Machine learning created the first wave of transformation in stress tolerance prediction. Algorithms such as Random Forest, SVM, and XGBoost have a high ability to analyze environmental and phenotypic data and can identify the most important traits or genes involved in resistance by ranking features (Sarkar *et al.,* 2024). For example, a Random Forest model can show which traits (e.g., proline content, leaf temperature index, or root-to-shoot ratio) play the greatest role in plant stability under drought conditions. These algorithms can also incorporate complex combinations of environmental conditions—such as simultaneous increases in temperature, decreased humidity, and increased radiation—into the analysis and estimate plant response under such conditions (Asaari *et al.,* 2022).

However, deep learning created a second, more powerful wave. CNNs have been able to detect subtle but important changes in plant phenotypic images—such as RGB photos, thermal images, and hyperspectral images—that are signs of stress initiation. These models can detect diseases before the appearance of visible symptoms, or detect drought and salinity stress at very early stages. This capability is important for intelligent farm management because it enables rapid intervention and damage reduction (Wen *et al.,* 2023).

In the area of biotic stress, AI has been particularly transformative in disease diagnosis. Deep CNNs can diagnose diseases from images of leaves, stems, or fruits with over 95% accuracy—even diseases that resemble each other or have different symptom severities. Models such as ResNet, EfficientNet, and Vision Transformers (ViT) are now used in crop disease detection systems and have dramatically reduced the cost of disease monitoring and the need for manual inspection (Henry *et al.,* 2025). In the area of abiotic stress, remote sensing-based models play a vital role. The combination of satellite, drone, and ground-based data with ML/DL algorithms has enabled accurate prediction of water status, thermal balance, and chlorophyll levels. These systems can detect drought and salinity stress at the field or regional scale before yield loss occurs (Bagherian, 2022). Models such as LSTM and GRU, which can analyze time-series data, also predict the future state of stress—for example, they can predict what level of plant stress will be reached within a week or month if current soil and weather conditions continue.

Explainable AI: Opening the black box in plant biotechnology

Although deep learning models have shown promising predictive power in plant biotechnology applications, their "black box" nature presents a significant barrier to scientific acceptance, regulatory approval, and end-user confidence (Castelvecchi, 2016; Samek *et al.,* 2021). Biologists, plant breeders and farmers are understandably reluctant to take action on predictions from models that have no way to explain how decisions were made. In order to address this limitation, Explainable Artificial Intelligence (XAI) has become a major research direction, offering solutions to make the model predictions transparent and interpretable without substantially sacrificing performance (Arrieta *et al*., 2020; Linardatos *et al*., 2021). The XAI methods applicable to plant biotechnologies can be categorised into three major approaches (Belle and Papantonis, 2021; Pai *et al*., 2025):

1. Feature Attribution Methods: Methods such as SHapley Additive exPlanations (SHAP) (Lundberg and Lee, 2017) and Local Interpretable Model-agnostic Explanations (LIME) (Ribeiro *et al*., 2016) provide importance scores for input features, identifying the genes, spectral bands or image sections that were most influential in a specific prediction. For example, SHAP values can be used to identify the most predictive transcription factors or metabolic markers of drought tolerance in a deep learning model, providing breeders with testable hypotheses about molecular mechanisms (Rudin, 2019; Murmu *et al*., 2024).

2. Visualization-Based Methods: Gradient-weighted Class Activation Mapping (Grad-CAM) and its variants produce heatmaps that highlight the most important regions of an input image for a model's decision (Selvaraju *et al*., 2017). In plant disease diagnosis, Grad-CAM can show if a CNN identified a lesion pattern specific to a pathogen, rather than background artefacts or leaf orientation, which is an important validation step before field deployment (Ghosal *et al*., 2018; Barbedo, 2019).

3. Concept-based and Rule-based Explanations: These approaches extract human readable rules or concepts from complex models (Kim *et al*., 2018). Random Forest models can generate decision trees or rule lists that offer simple "if-then" logic that is easily understood and verified by farmers and extension officers (Letham *et al*., 2015; Molnar, 2022). Applications of XAI in plant biotechnology are rapidly increasing (Pai *et al*., 2025): Disease Diagnosis: XAI techniques have been applied to plant disease detection systems to ensure that models are focusing on biologically relevant symptoms (e.g., chlorotic lesions, sporulation patterns) and not spurious correlations with background soil, lighting conditions, or leaf orientation (Singh *et al*., 2020; Bansal *et al*., 2021). Grad-CAM visualizations have been used to identify model failures such as

reliance on leaf edges instead of pathogen-specific patterns and thus have provided guidance for targeted data augmentation and model improvement (Wang *et al*., 2021; Zhang *et al*., 2022).

Genomic Prediction. Deep learning models are used to predict complex traits from genomic sequences, and SHAP analysis is used to identify which specific SNPs or regulatory motifs are driving predictions (Wang *et al*., 2025b). This confirms model logic and also reveals new candidate markers for breeding programs (Farooq *et al*., 2024). Within the multi-omic integration framework, XAI delivers the insight into the contribution of different molecular layers (transcriptomic, proteomic, metabolomic) to the phenotype predictions that can be used to guide experimental validation efforts (Ahmed *et al*., 2023; Cembrowska-Lech *et al*., 2023). Stress Phenotyping: Models based on hyperspectral or thermal imaging can be used to identify the spectral wavelengths or temperature patterns most diagnostic of early-stage drought or salinity stress, using (XAI) (Asaari *et al*., 2022; Wen *et al*., 2023). This knowledge helps in developing simpler and more interpretable sensors and reduces the computational requirements for edge-deployment (Bagherian, 2022; Islam *et al*., 2024). The benefits of XAI are significant for stakeholder trust and adoption (Ghazal *et al*., 2025; Pai *et al*., 2025). For plant scientists, explainable models produce testable biological hypotheses, not uninterpretable correlation (Rudin, 2019). In the context of breeders, XAI provides confidence that the selection decisions are based on relevant genetic or phenotypic features (Farooq *et al*., 2024). Simple rule-based explanations (e.g., “the model predicted disease because it detected yellow spots covering >15% of leaf area”) enable farmers and extension officers to ground-truth and build trust necessary for adoption (Aker, 2011; Fabregas *et al*., 2019). However, there are still several challenges for XAI in plant biotechnology (Arrieta *et al*., 2020; Samek *et al*., 2021). First, validation of explanation faithfulness, i.e., whether the attributed features do indeed represent the decision process of the model, is difficult in biological contexts where the ground-truth causal mechanisms are incompletely known (Rudin, 2019). Second, different XAI techniques may provide different explanations for the same prediction, which may confuse a user as to which explanation to trust (Krishnan, 2020). Third, the computational cost of extracting explanations could be too high for real-time field applications on resource-limited edge devices (Ghazal *et al*., 2025). Finally, XAI does not solve other problems such as data bias or model robustness directly. An explainable but biased model is still a problem (Mehrabi *et al*., 2022).

Future directions for XAI in plant biotechnology include (i) the development of plant-specific XAI benchmarks and evaluation metrics, (ii) the integration of explanation generation directly into model architectures (explainable-by-design models), and

(iii) the development of standardized visualization interfaces tailored to different end-user groups (scientists, breeders, farmers) (Pai *et al*., 2025; Xu *et al*., 2025). The combination of XAI with federated learning, which allows privacy-conscious and interpretable collaborative model training across distributed agricultural datasets, is a notably promising avenue for scalable, trustworthy AI in global plant biotechnology (Li *et al*., 2021; Ghazal *et al*., 2025).

Application of artificial intelligence in predictive breeding and accelerating plant improvement

Plant breeding, as one of the fundamental bases of food security and sustainable production, has always relied on identifying superior genotypes, favorable genetic diversity, and complex traits. However, classical breeding methods have serious limitations despite their historical effectiveness: long time requirements, high costs, and difficulty in analyzing polygenic and complex traits are among the main obstacles (Hafeez *et al.,* 2023). The emergence of AI—especially machine learning, deep learning, and multi-modal modeling algorithms—has transformed this traditional process and created the foundation for predictive breeding; an approach in which breeding decisions are made not based on trial and error but on precise predictions, genotype-phenotype modeling, and analysis of plant growth dynamics (Xu *et al.,* 2022) (Fig. 5).

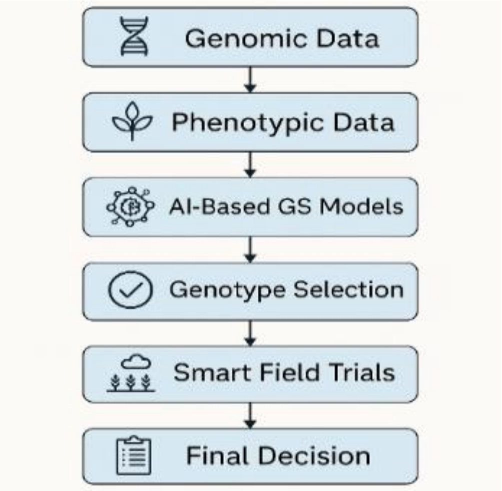

Figure 5: Workflow of AI-based predictive plant breeding, including integration of genomic and phenotypic data, intelligent genomic selection (GS) models, selection of superior genotypes, remote sensing-based field trials, and final decision-making to predict and optimize plant performance.

Predictive breeding is fundamentally based on the use of multi-source data: genome, transcriptome, metabolomic data, phenotypic images, climatic data, soil properties, and management information. AI can combine these complex layers into integrated models and recover the hidden relationship between genotype and phenotype—something that classical statistical methods are often unable to do (Amin *et al.,* 2025). Random Forest, XGBoost, CNNs, GNNs, and multi-channel transformers are now widely used to predict genotype performance in different environments. These models can estimate the effect of environmental conditions on trait expression and suggest the best genotypes for specific environments—a topic of vital importance in the era of climate change (Choi and Lee, 2023).

One of the most important roles of AI in plant breeding is drastically reducing the time required to select superior genotypes. In traditional systems, screening hundreds or thousands of genotypes takes years and requires multiple field trials. However, AI models can rank high-potential genotypes with high accuracy by analyzing initial data—whether genomic, phenotypic, or climatic. For example, recent studies have shown that CNNs trained on early growth images can predict end-of-season yield. This rapid prediction can reduce the breeding cycle by up to 50% or more (Xu *et al.,* 2022). Another key role of AI is in genomic selection (GS). In genomic selection, models estimate the overall genomic effect on a trait and, based on that, superior genotypes are selected. Machine learning algorithms—especially Bayesian models, kernel-based methods, and neural networks—have significantly increased the accuracy of genomic selection (Heslot *et al.,* 2012). In many strategic crops such as wheat, maize, rice, and soybean, the use of AI-based genomic selection has made it possible to select superior genotypes in early generations and has dramatically reduced the costs of breeding programs (Zhang *et al.,* 2025). Another important application of AI is predicting G×E interactions—i.e., genotype-by-environment interaction. Deep learning models can predict the effects of climate change, temperature, humidity, and crop management on plant traits. This capability is vital for designing genotypes adapted to future climates—one of the greatest challenges facing humanity (Malosetti *et al.,* 2016).

Integration of AI with bioengineering and systems biology: Modeling gene regulatory, metabolic, and cellular networks

Bioengineering and systems biology are based on the principle that the behavior of a cell, tissue, or whole plant is the result of complex interactions among thousands of molecules, pathways, regulators, and dynamic networks. These networks are usually non-linear, time-dependent, and influenced by environmental conditions. Therefore, analyzing and engineering them with traditional methods is often limited and incomplete (Sheth and Thaker, 2014). The emergence of AI—especially deep learning models, graph networks, and generative systems—has overcome past limitations and ushered bioengineering into a new stage where genetic and metabolic pathway design is based on accurate, predictive simulations (Naskar *et al.,* 2025) (Fig. 6).

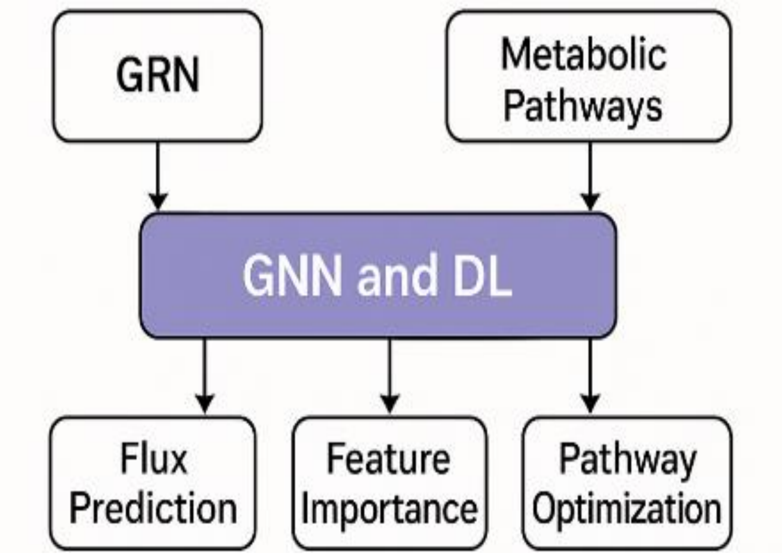

Figure 6: Modeling of gene regulatory networks and metabolic pathways using graph neural networks (GNNs) and deep learning, including reconstruction of gene regulatory networks (GRNs), metabolic pathways, prediction of metabolic fluxes, feature importance analysis, and optimization of biological pathways.

One of the most important applications of AI in systems biology is the modeling of gene regulatory networks (GRNs). Gene networks in plants are very complex and involve thousands of genes, transcription factors, and feedback interactions. Classical models such as Bayesian networks or ordinary differential equation (ODE)-based systems can explain parts of network behavior but face limitations when faced with genomic scale and multi-layered interactions (Dasgupta and De, 2023). In contrast, graph neural networks have been able to model the graph-like structure of GRNs with high accuracy. GNNs can detect connection patterns, key nodes, regulatory pathways, and the role of genes in signal transduction, and even predict the effect of deleting or adding a gene. This capability is vital for designing stress-resistant genotypes or plants with optimized metabolic pathways (Wu *et al.,* 2020).

The digital twin revolution in plants: Simulating growth, metabolism, and stress tolerance

The concept of the digital twin, first introduced in aerospace and complex systems engineering, is now rapidly expanding in the life sciences, especially plant sciences (Fig. 7). A digital twin is essentially a dynamic, intelligent computational model that reconstructs a virtual version of a real system and can predict its behavior in real time or through complex simulations. The application of this technology in plants has created an emerging revolution: for the first time, the growth, physiology, metabolic pathways, and stress responses of a plant can be fully reconstructed in a digital environment, and the effect of any genetic, management, or environmental change can be assessed before it occurs (Chen *et al.,* 2023; Iranshahi *et al.,* 2025).

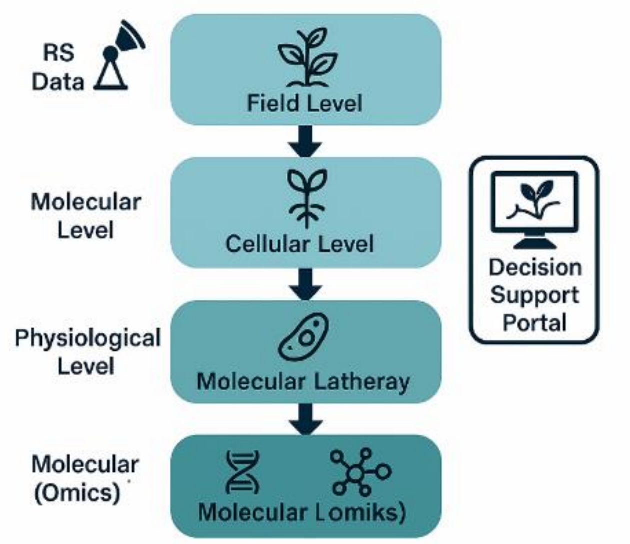

Figure 7: Architecture of a plant digital twin, including molecular layer (omics), cellular layer, physiological layer, field layer, remote sensing (RS) data, and decision portal for simulating and predicting plant behavior under different conditions.

A plant digital twin combines omics, phenotypic, physiological, climatic, and remote sensing data. Machine learning and deep learning algorithms integrate these data to create a dynamic model that not only reflects the current state of the plant but also predicts its future behavior. In this model, every biological layer—from the genome and gene networks to metabolic pathways, physiological exchanges, and environmental responses—is connected through precise computational relationships. Thus, the digital twin is not a static model but a living computational version of the plant (Cembrowska-Lech *et al.,* 2023).

One of the most important applications of the digital twin is simulating plant growth and development. Deep models—especially LSTMs and transformer-based predictors—can predict growth trends from seed to maturity. For example, a digital twin can estimate future growth patterns, leaf development rate, biomass amount, and even final yield based on initial plant images, soil moisture status, and climatic conditions. Such simulations are of vital value for crop planning, plant breeding, and intelligent farm management (Bi *et al.,* 2023).

On the other hand, the digital twin plays a fundamental role in reconstructing plant metabolic pathways. Metabolic flux models, when combined with deep learning, can predict metabolic fluxes at the cellular level. In these systems, the behavior of thousands of enzymes, reactions, intermediates, and branching pathways is modeled in parallel and dynamically. The digital twin can show how a genetic mutation in an enzyme changes metabolic flux and how this change affects growth, stress tolerance, or the production of valuable metabolites. This capability is essential for designing medicinal plants, industrial plants, or crops with high nutritional value (Li *et al.,* 2022).

One of the most advanced capacities of the digital twin is modeling plant responses to biotic and abiotic stresses. Under drought, salinity, or heat stress, the molecular and physiological interactions of the plant are very complex and multi-layered. The digital twin can analyze thousands of internal signals—from hormones to transcription factors and water pathway regulators—and accurately predict how the plant will behave when faced with stress. These models can even run hypothetical scenarios—for example, if proline content, antioxidant enzyme activity, or the expression of ion transport-related genes changes, how will the plant modify its resistance? (Xu *et al.,* 2025).

The combination of the digital twin with remote sensing (RS) data is also a major advance. Drone, satellite, or ground-based images are continuously fed into the model, and the digital twin updates the plant's status in real time. If a plant in the field experiences heat stress, this change is reflected in spectral reflectance or leaf surface temperature in the images, and the digital twin immediately simulates the possible metabolic and physiological changes. This feature has made real-time monitoring possible for large-scale farms—something previously difficult to imagine (Wu *et al.,* 2023).

In recent years, the digital twin has found amazing applications in plant breeding. The model can predict the effect of any allelic combination, mutation, or CRISPR edit on phenotype and yield without the need for costly and time-consuming experiments. This capability has given rise to a new generation of plant breeding known as *in silico* breeding, in which the best genotypes are selected solely through simulation before actual cultivation. This approach is particularly effective for complex traits such as drought tolerance or improving nutritional quality (Peladarinos *et al.,* 2023). Another application is using the digital twin to optimize farm management. The model can simulate different irrigation, nutrition, planting distance, or plant density scenarios and suggest the best management combination for maximum yield. This novel precision agriculture will play a major role in water, input, and energy productivity.

At the scientific level, the digital twin has created an important paradigm shift: from one-dimensional analyses to multi-layered analyses, from simple predictions to network-based predictions, and from experimental breeding to fully computational breeding. This technology is still in the early stages of development in plant sciences, but global trends indicate that in the next decade, it will become one of the main pillars of plant biotechnology, plant breeding, smart agriculture, and genome design (Alves *et al.,* 2023).

Future pathways: Toward automated plant biotechnology and the next generation of smart agriculture

The combination of AI with plant biotechnology over the past decade has opened new pathways in research, plant breeding, agriculture, and bioengineering—a pathway that is now rapidly moving toward the next generation of intelligent and automated systems (Niazian and Niedbała, 2020). The future of plant biotechnology is a future in which predictive models, farm robots, self-learning systems, and computation-based genome design will play a central role. This process is not merely a technological upgrade; it is a paradigmatic shift that will transform the very nature of plant research and production.

One of the most important future pathways is the development of fully automated plant biotechnology. In this vision, data from farm sensors, drones, weather stations, and digital phenotyping tools are automatically fed into a central system. This system, using deep learning, analyzes the information and detects plant growth patterns, stress, or nutritional needs. Then, decisions—from irrigation and nutrition to genotype selection and even design of genetic mutations—are made automatically. This represents a transition from manual monitoring to automated, intelligent management (Sahoo *et al.,* 2026).

Another important future pathway is the design of novel recombinant genomes with the help of AI. Generative models, graph networks, and biological optimization algorithms can now suggest genomes that do not exist in nature but are computationally stable, efficient, and adapted to future environmental conditions. This new stage in biotechnology—called computational genome design—will play a vital role in producing plants that simultaneously possess tolerance to drought, salinity, heat, or nutrient deficiencies. Combining this approach with next-generation CRISPR will change plant breeding from modifying existing genes to designing entirely new genomes (Dixit *et al.,* 2024a).

Enhancing the capabilities of plant digital twins will also play an important role in the future. The digital twin will not only simulate growth and stress responses but also, in more future-oriented versions, will be able to model the effects of climate change, changes in farm management, or genetic mutations over long-term time scales. Digital twins equipped with continual learning will be able to update themselves with new data and provide an increasingly accurate model of the plant biosystem (Liu *et al.,* 2023).

Bridging the lab-to-farm gap: Real-world deployment challenges and farmer-centered solutions

The computational advances described above are revolutionizing plant biotech but their translation to the field will ultimately depend on successful deployment in agricultural practice, which is not without significant technical, socio-economic and institutional barriers (Aker, 2011; Cole and Fernando, 2021). These barriers have led to low adoption rates of digital agricultural technologies in many parts of the world, especially among smallholder farmers in low- and middle-income countries (Nakasone and Torero, 2016; Fabregas *et al*., 2019). Infrastructural and technical barriers still exist in plenty (Aker, 2011; Tsan *et al*., 2019). First, a lot of rural agricultural areas don't have dependable internet access, so cloud-dependent AI applications don't work (Kremer *et al*., 2019). Second, the computational needs of deep learning models (e.g., GPUs for Transformer or large CNN inference) go beyond the capacities of standard smartphones or affordable field-deployable edge devices for farmers (Mohanty *et al*., 2016; Barbedo, 2019). Third, for most crops, and particularly for orphan crops and regionally important varieties, there is a lack of high-quality training data annotated by experts for a large diversity of environmental conditions, crop varieties and disease strains (Barbedo, 2019; Kamilaris and Prenafeta-Boldú, 2018). Models trained on images from controlled greenhouses or on publicly available datasets often fail catastrophically when deployed in fields with variable lighting, complex backgrounds and novel disease presentations (Mohanty *et al*., 2016; Barbedo, 2019).

Socio-economic and farmer-centered barriers are also significant (Tsan *et al*., 2019; Amar *et al*., 2026). Many AI-driven agricultural tools are created without meaningful input from the farmers who use them, resulting in "one-size-fits-all" answers that do not fit local farming practices, literacy rates, language choices, and cultural settings (Mansur *et al*., 2022; Amar *et al*., 2026). Involvement of farmers in the design process, from problem identification to interface testing, is critical but rarely systematic (Fabregas *et al*., 2019; Cole and Fernando, 2021). The expense of AI-enabled sensors, drones, or subscription-based software platforms is also prohibitive for smallholder farmers who operate on thin margins (Aker, 2011; Tsan *et al*., 2019). Without subsidies or innovative financing mechanisms, these technologies exacerbate, not alleviate, agricultural inequalities (Fabregas *et al*., 2019). Other challenges include data governance and trust barriers (Wiseman *et al*., 2019; Amar *et al*., 2026). Farmers are increasingly asked to share farm and personal data without clear information on who can access it, how it is used, or whether it affects eligibility for subsidies, insurance, or credit (Taylor and Bhasme, 2021; Amar *et al*., 2026). In many countries, trust in digital agricultural platforms has been eroded by opaque data practices and fragile legal structures for data privacy and consent (Wiseman *et al*., 2019). Proprietary platforms that lock farmers into specific tools, impede data export or inhibit the ability to customise recommendations further erode farmer agency and long-term adoption (Zimmer, 2018; Amar *et al*., 2026). Successful case studies have determined strategies to overcome deployment barriers (Aker, 2011; Fabregas *et al*., 2019; Tsan *et al*., 2019): Participatory Design and Co-Creation: Farmer involvement throughout the entire process of technology development, from needs assessment and prototyping to field testing, greatly increases usability, relevance, and adoption (Cole and Fernando, 2021; Amar *et al*., 2026). For example, the engagement of Irish farmers in a participatory design process resulted in the development of a smartphone app for agricultural reporting that achieved greater adoption rates than traditional top-down initiatives, as it fitted into farmers' normal routines and addressed their pain points (Healy *et al*., 2021).

Appropriate Technology and Modality Selection: In low-connectivity, low-literacy settings, voice-based interfaces, SMS services and interactive voice response (IVR) systems may be more effective than smartphone applications that require reading proficiency and access to the internet (Aker, 2011; Nakasone and Torero, 2016). Digital literacy gaps are addressed through the integration of digital tools with local "resource farmers" or extension agents that provide supplemental support to build community trust (Fabregas *et al*., 2019; Cole and Fernando, 2021). Privacy-preserving and decentralized architectures: Technologies like federated learning allow for collaborative model training without centralizing raw farm data, alleviating data sovereignty issues while still harnessing AI capabilities (Li *et al*., 2021; Ghazal *et al*., 2025). Clear and enforceable data governance frameworks that define data ownership, consent procedures, usage restrictions, and benefit-sharing mechanisms are critical to building trust among farmers (Wiseman *et al*., 2019; Amar *et al*., 2026).

Capacity Building and Digital Literacy: Ongoing investment in farmer training programs, digital literacy initiatives, and peer-learning networks is crucial to ensure the effective usage and adaptation of AI tools by farmers (Aker, 2011; Fabregas *et al*., 2019). This is especially important for women farmers and other marginalized groups who could encounter additional barriers to access technology (Tsan *et al*., 2019; Amar *et al*., 2026). Policy and Institutional Support: Government subsidies for AI-enabled sensors and connectivity infrastructure, tax incentives for aggrotech companies serving smallholders, and regulatory frameworks that mandate transparency and data protection can accelerate equitable adoption (Fabregas *et al.*, 2019; Kremer *et al.*, 2019). In particular, public-private partnerships that combine private sector innovation with public sector reach and trust have shown promise (Zimmer, 2018; Tsan *et al.*, 2019). Another key pathway is integration with existing agricultural extension systems (Aker, 2011; Cole and Fernando, 2021). AI tools are most effective when they are designed to augment the capabilities of human extension agents rather than replace them, by providing decision support, helping to prioritise farm visits and enabling more personalised advice at scale (Fabregas *et al.*, 2019; Tsan *et al.*, 2019). The notion of "digital extension" as a complement to traditional face-to-face extension needs careful attention to the roles, incentives and training of extension officers as intermediaries between AI systems and farming communities (Aker, 2011; Cole and Fernando, 2021). A deployment-oriented plant biotechnology research agenda should include (Fabregas *et al.*, 2019; Tsan *et al.*, 2019; Amar *et al.*, 2026): (1) development of lightweight AI models suitable for edge-computing on inexpensive hardware; (2) development of robust, diverse, and openly available field-annotated datasets for multiple environments, seasons, and cultural practices; (3) systematic evaluation of XAI methods for farmer-facing explanations; (4) randomized controlled trials to measure the agronomic, economic, and social impacts of AI deployment across different farming systems; and (5) comparative studies of participatory design, governance, and financing mechanisms to determine context-appropriate implementation strategies.

Conclusion and future outlook

The integration of artificial intelligence with plant biotechnology in recent years has created a fundamental transformation, steering research and development in plant sciences toward more precise, faster, and more predictive systems. Deep learning-based multi-omics analysis, increased precision of genome editing with off-target prediction models, field-scale digital phenotyping, early detection of biotic and abiotic stresses, and the development of predictive breeding are among the most important outcomes of this technological integration. The investigations conducted in this article show that AI is capable of extracting complex biological patterns that were previously undetectable with traditional methods and provides insights that contribute to the development of high-yielding, resilient plants adapted to climate change. Furthermore, AI plays an important role in increasing efficiency, reducing costs, and accelerating research and breeding processes, and can bridge the gap between big biological data and practical decision-making.

Despite these opportunities, challenges such as data quality, heterogeneity of sources, model interpretability, biosecurity, and ethical issues remain and require the development of standardized frameworks, more transparent algorithms, and more responsible approaches to deploying this technology. The overall conclusion is that AI is not merely an auxiliary tool but a fundamental pillar in the transformation of the next generation of plant biotechnology. Alongside data quality and model interpretability, the convergence of AI with gene editing and synthetic biology presents dual-use concerns (De Haro, 2024) – where the same technologies developed for beneficial crop improvement could potentially be misused to develop agricultural pathogens or disrupt food systems (Dixit *et al*., 2024a). Biosecurity frameworks for AI-designed genetic sequences, screening protocols for DNA synthesis orders, and international governance mechanisms for AI-powered biotechnologies are urgently needed but still underdeveloped (De Haro, 2024; Kim *et al*., 2025).

The future outlook for intelligent plant biotechnology is a multi-layered, interconnected future based on advanced computational models. In the coming decade, we expect foundation plant models—similar to powerful language models—to be developed as a unified computational platform for analyzing, predicting, and designing plant biosystems. Computational genome design with the help of generative models will accelerate the move toward computational bio-design, enabling the production of plants with optimized traits for future climates (Fig. 8).

Alongside this, the plant digital twin will become one of the main tools for predicting growth, yield, stress responses, and evaluating management scenarios, enabling researchers to conduct virtual experiments with very high precision. At the farm level, robotic agriculture and self-learning systems will play a decisive role. Ground robots and drones equipped with multi-spectral sensors and machine vision models will upgrade farm management from the plot level to the individual plant level. On the other hand, the integration of omics data, remote sensing, meteorology, and phenotyping into intelligent cloud systems will create large digital networks that bring together all biological and environmental components in a single ecosystem, revolutionizing plant breeding and bioengineering.

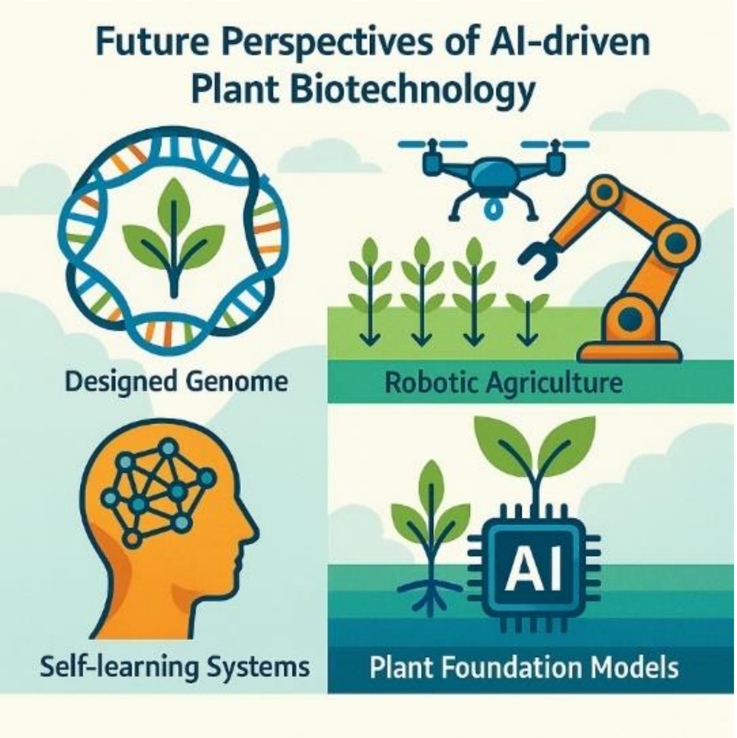

Figure 8: Future outlook of intelligent plant biotechnology, including novel genome design, robotic agriculture, self-learning AI-based systems, and foundation models specifically for plant sciences

Parallel to the development of these technologies, the need to design ethical, regulatory, and biosecurity frameworks becomes increasingly important. Establishing global standards for data quality, model interpretability, algorithmic transparency, and protection of biological data is essential to ensure public acceptance and foster sustainable innovation. In summary, the future of intelligent plant biotechnology is one in which research, breeding, prediction, and plant management are all conducted within a fully intelligent digital-biological cycle—a cycle that can transform global food security, agricultural productivity, and environmental sustainability at an unprecedented level.

References

Abbaszadeh, A. and Shahlai, A., 2025. *Artificial Intelligence for CRISPR Guide RNA Design: Explainable Models and Off-Target Safety*. arXiv. https://doi.org/10.48550/arXiv.2508.20130

Abdulkareem, A., Jimada-Ojuolape, B., Balogun, M.O. and Adesina, L.M., 2024. Comparative Analysis of XGBoost and Random Forest Algorithms for Transformer Failure Prediction. *Annals of Faculty Engineering Hunedoara – International Journal of Engineering*, 22(**4**), 137-146. (No DOI - ISSN: 1584-2665)

Ahmed, S.F., Alam, M.S.B., Hassan, M., Rozbu, M.R., Ishtiak, T., Rafa, N., Mofijur, M., Shawkat Ali, A.B.M. and Gandomi, A.H., 2023. Deep learning modelling techniques: current progress, applications, advantages, and challenges. *Artificial Intelligence Review*, 56(**11**), 13521-13617. https://doi.org/10.1007/s10462-023-10466-8

Aker, J. C. (2011). Dial 'A' for agriculture: A review of information and communication technologies for agricultural extension in developing countries. *Agricultural Economics*, 42 (6), 631-647. https://doi.org/10.1111/j.1574-0862.2011.00545.x

Alamin, M., Sultana, M. H., Lou, X., Jin, W. and Xu, H., 2022. Dissecting complex traits using omics data: A review on the linear mixed models and their application in GWAS. *Plants*, 11(**23**), 3277. https://doi.org/10.3390/plants11233277

Alamin, M., Sultana, M. H., Lou, X., Jin, W., & Xu, H. (2022). Dissecting complex traits using omics data: A review on the linear mixed models and their application in GWAS. *Plants*, 11(23), 3277. https://doi.org/10.3390/plants11233277

Alves, R.G., Maia, R.F. and Lima, F., 2023. Development of a Digital Twin for smart farming: Irrigation management system for water saving. *Journal of Cleaner Production*, 388, 135920. https://doi.org/10.1016/j.jclepro.2023.135920

Alwani, N.M., 2024. *Synthetic data generation for the optimization of strains in metabolic engineering using generative adversarial networks* [Master's thesis]. Delft University of Technology.

Amar, S., Bori, N. and Cörvers, R., 2026. From collection to control: Data governance, digital technologies, and the politics of inclusion in the digitalisation of smallholder agriculture. *Outlook on Agriculture*, 55(**1**), 19-27. https://doi.org/10.1177/00307270251410944

Amin, A., Zaman, W. and Park, S., 2025. Harnessing Multi-Omics and Predictive Modeling for Climate-Resilient Crop Breeding: From Genomes to Fields. *Genes*, 16(**7**), 809. https://doi.org/10.3390/genes16070809

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... & Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82-115. https://doi.org/10.1016/j.inffus.2019.12.012

Asaari, M.S.M., Mertens, S., Verbraeken, L., Dhondt, S., Inzé, D., Bikram, K. and Scheunders, P., 2022. Non-destructive analysis of plant physiological traits using hyperspectral imaging: A case study on drought stress. *Computers and Electronics in Agriculture*, 195, 106806. https://doi.org/10.1016/j.compag.2022.106806

Bagherian, K., 2022. *Rapid peanut phenotyping and water quality monitoring using remote sensing and machine learning techniques* [Doctoral dissertation]. Auburn University.

Balajishanmugam, V., 2025. Smart Gene Editing. In *Advanced AI and Data Science Applications*, CRC Press, chapter 7,14 pages

Bansal, A., Kaur, A., & Kaur, A. (2021). Explainable AI for plant disease detection using Grad-CAM. *2021 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE)*, 78-82. https://doi.org/10.1109/ICCIKE51210.2021.9410733

Barbedo, J. G. A. (2019). Plant disease identification from individual lesions and spots: A review. *Biosystems Engineering*, 180, 30-45. https://doi.org/10.1016/j.biosystemseng.2018.12.009

Belle, V., & Papantonis, I. (2021). Principles and practice of explainable machine learning. *Frontiers in Big Data*, 4, 688969. https://doi.org/10.3389/fdata.2021.688969

Bellot, P., de Los Campos, G. and Pérez-Enciso, M., 2018. Can deep learning improve genomic prediction of complex human traits? *Genetics*, 210(**3**), 809-819. https://doi.org/10.1534/genetics.118.301298

Bhardwaj, A., Kishore, S. and Pandey, D.K., 2022. Artificial intelligence in biological sciences. *Life*, 12(**9**), 1430. https://doi.org/10.3390/life12091430

Bi, L., Wally, O., Hu, G., Tenuta, A.U., Kandel, Y.R. and Mueller, D.S., 2023. A transformer-based approach for early prediction of soybean yield using time-series images. *Frontiers in Plant Science*, 14, 1173036. https://doi.org/10.3389/fpls.2023.1173036

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., ... & Liang, P. (2021). On the opportunities and risks of foundation models. *arXiv preprint*, arXiv:2108.07258. https://doi.org/10.48550/arXiv.2108.07258

Cai, S., Sun, C. and Tian, J., 2025. Towards smart agriculture: AI-driven prediction of key genes for revolutionizing crop breeding. *Planta*, 262, 116. https://doi.org/10.1007/s00425-025-04841-8

Castelvecchi, D. (2016). Can we open the black box of AI? *Nature*, 538(7623), 20-23. https://doi.org/10.1038/538020a

Cembrowska-Lech, D., Krzemińska, A., Miller, T., Nowakowska, A., Adamski, C., Radaczyńska, M., Mikiciuk, G. and Mikiciuk, M., 2023. An integrated multi-omics and artificial intelligence framework for advance plant phenotyping in horticulture. *Biology*, 12(**10**), 1298. https://doi.org/10.3390/biology12101298

Chen, K., Zhao, B., Zhou, H., Zhou, L., Niu, K., Jin, X., Li, R., Yuan, Y.

and Zheng, Y., 2023. Digital twins in plant factory: A five-dimensional modeling method for plant factory transplanter digital twins. *Agriculture*, 13(**7**), 1336. https://doi.org/10.3390/agriculture13071336

Ching, T., Himmelstein, D.S., Beaulieu-Jones, B.K., Kalinin, A.A., Do, B.T., Way, G.P., Ferrero, E., Agapow, P.M., Zietz, M. and Hoffman, M.M., 2018. Opportunities and obstacles for deep learning in biology and medicine. *Journal of the Royal Society Interface*, 15(**141**), 20170387. https://doi.org/10.1098/rsif.2017.0387

Choi, S.R. and Lee, M., 2023. Transformer architecture and attention mechanisms in genome data analysis: a comprehensive review. *Biology*, 12(**7**), 1033. https://doi.org/10.3390/biology12071033

Chuai, G., Ma, H., Yan, J., Chen, M., Hong, N., Xue, D., Zhou, C., Zhu, C., Chen, K. and Duan, B., 2018. DeepCRISPR: optimized CRISPR guide RNA design by deep learning. *Genome Biology*, 19(**1**), 80. https://doi.org/10.1186/s13059-018-1458-4

Cole, S., & Fernando, A. N. (2021). 'Mobile'izing agricultural advice: Technology adoption, diffusion, and sustainability. *The Economic Journal*, 131(633), 192-219. https://doi.org/10.1093/ej/ueaa084

Dai, Q., Chen, Z., Li, Z., Song, S., Xue, X., Lv, S., Wang, Y. and Guo, Y., 2024. A Monitoring, Evaluation, and Prediction System for Slight Water Stress in Citrus Seedlings Based on an Improved Multilayer Perceptron Model. *Agronomy*, 14(**4**), 808. https://doi.org/10.3390/agronomy14040808

Dai, X. and Shen, L., 2022. Advances and trends in omics technology development. *Frontiers in Medicine*, 9, 911861. https://doi.org/10.3389/fmed.2022.911861

Dasgupta, A. and De, R.K., 2023. Artificial intelligence in systems biology. In *Handbook of Statistics* (Vol. 49, pp. 153-201). Elsevier. https://doi.org/10.1016/bs.host.2023.06.004

De Haro, L.P., 2024. Biosecurity risk assessment for the use of artificial intelligence in synthetic biology. *Applied Biosafety*, 29(2), 96-107. https://doi.org/10.1089/apb.2024.0015

Dixit, S., Kumar, A., Srinivasan, K., Vincent, P.D.R. and Ramu Krishnan, N., 2024a. Advancing genome editing with artificial intelligence: opportunities, challenges, and future directions. *Frontiers in Bioengineering and Biotechnology*, 11, 1335901. https://doi.org/10.3389/fbioe.2023.1335901

Dixit, S., Sivalingam, P.N., Baskaran, R.M., Senthil-Kumar, M. and Ghosh, P.K., 2024b. Plant responses to concurrent abiotic and biotic stress: unravelling physiological and morphological mechanisms. *Plant*

Physiology Reports, 29(**1**), 6-17. https://doi.org/10.1007/s40502-023-00767-1

Dong, X.L. and Rekatsinas, T., 2018. Data integration and machine learning: A natural synergy. In *Proceedings of the 2018 International Conference on Management of Data* (pp. 1645-1650). ACM. (No DOI - conference proceeding)

Edukondalu, B., Aswini, N., Amaresh, Singh, B., Nare, B., Veni, K., Sneha, M. and Devi, K.R., 2026. Accelerating genetic gain through integrated genomic selection in crop plants. *Journal of Applied Genetics*, 67, 249-269. https://doi.org/10.1007/s13353-025-01034-7

Fabregas, R., Kremer, M., & Schilbach, F. (2019). Realizing the potential of digital development: The case of agricultural advice. *Science*, 366(6471), eaay3038. https://doi.org/10.1126/science.aay3038

Farooq, M.A., Gao, S., Hassan, M.A., Huang, Z., Rasheed, A., Hearne, S., Prasanna, B., Li, X. and Li, H., 2024. Artificial intelligence in plant breeding. *Trends in Genetics*, 40(**10**), 891-908. https://doi.org/10.1016/j.tig.2024.08.001

Fong, J.H. and Wong, A.S., 2023. Advancing CRISPR/Cas gene editing with machine learning. *Current Opinion in Biomedical Engineering*, 28, 100477. https://doi.org/10.1016/j.cobme.2023.100477

Gao, S., Yu, T., Rasheed, A., Wang, J., Crossa, J., Hearne, S. and Li, H., 2025. Fast-forwarding plant breeding with deep learning-based genomic prediction. *Journal of Integrative Plant Biology*, 67(**7**), 1700-1705. https://doi.org/10.1111/jipb.13914

Ghazal, A., Al-Masri, E., & Al-Madi, N. (2025). A Systematic Review of AI-Driven Wearable Plant Disease Recognition Systems. *2025 12th International Conference on Future Internet of Things and Cloud (FiCloud)*, 424-430. https://doi.org/10.1109/FiCloud63104.2025.00069

Ghosal, S., Blystone, D., Singh, A. K., Ganapathysubramanian, B., Singh, A., & Sarkar, S. (2018). An explainable deep machine vision framework for plant stress phenotyping. *Proceedings of the National Academy of Sciences*, 115(18), 4613-4618. https://doi.org/10.1073/pnas.1716999115

Hafeez, A., Ali, B., Javed, M. A., Saleem, A., Fatima, M., Fathi, A., Afridi, M.S., Aydin, V., Oral, M.A. and Soudy, F.A., 2023. Plant breeding for harmony between sustainable agriculture, the environment, and global food security: an era of genomics-assisted breeding. *Planta*, 258(**5**), 97. https://doi.org/10.1007/s00425-023-04243-8

Hallajian, M.T., Ebadi, A.A. and Kordrostami, M., 2024. Advancing rice breeding for drought tolerance: a comprehensive study of traditional and mutant lines through agronomic performance and drought tolerance indices. *BMC Plant Biology*, 24(**1**), 1087. https://doi.org/10.1186/s12870-024-05724-8

Healy, S., O'Donoghue, C., & Ryan, M. (2021). Participatory design of a smartphone app for agricultural reporting. *Journal of Rural Studies*, 86, 127-138. https://doi.org/10.1016/j.jrurstud.2021.05.018

Henry, A., Daniel, E., Seetha, S. and Durga, S., 2025. Bridging CNNs and Transformers: A Comparative Study of EfficientNet, ResNet, and ViT for Plant Disease Detection. In *2025 6th International Conference on Data Intelligence and Cognitive Informatics (ICDICI)* (pp. 718-724). IEEE. (No DOI - conference proceeding)

Heslot, N., Yang, H.P., Sorrells, M.E. 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

Hoque, T.S., Sohag, A.A.M., Kordrostami, M., Hossain, M.A., Islam, M.S., Burritt, D.J. and Hossain, M.A., 2020. The effect of exposure to a combination of stressors on rice productivity and grain yields. In A. Roychoudhury (Ed.), *Rice research for quality improvement: genomics and genetic engineering: volume 1: breeding techniques and abiotic stress tolerance* (pp. 675-727). Springer. https://doi.org/10.1007/978-981-15-4120-9_28

Iranshahi, K., Brun, J., Arnold, T., Sergi, T. and Müller, U.C., 2025. Digital twins: Recent advances and future directions in engineering fields. *Intelligent Systems with Applications*, 26, 200516. https://doi.org/10.1016/j.iswa.2025.200516

Islam, S., Reza, M.N., Samsuzzaman, S.A., Cho, Y.J., Noh, D.H., Chung, S.O. and Hong, S.J., 2024. Machine vision and artificial intelligence for plant growth stress detection and monitoring: A review. *Precision Agriculture Science and Technology*, 6(**1**), 33-57. https://doi.org/10.12972/pastj.20240003

Jiang, B., An, Z., Niu, L. and Qin, D., 2025. Precise genome editing process and its applications in plants driven by AI. *Functional and Integrative Genomics*, 25(**1**), 109. https://doi.org/10.1007/s10142-025-01582-9

Kalita, I., Bhattacharjee, S. and Saharia, M., 2025. Advancements in Medicinal Plant Research: Harnessing Artificial Intelligence, Machine Learning, Deep Learning, and Bioinformatics. In *Biotechnology, Multiple Omics, and Precision Breeding in Medicinal Plants* (pp. 135-145). CRC Press. (No DOI - book chapter)

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147, 70-90. https://doi.org/10.1016/j.compag.2018.02.016

Kim, B., Wattenberg, M., Gilmer, J., Cai, C., Wexler, J., & Viegas, F. (2018). Interpretability beyond feature attribution: Quantitative testing with concept activation vectors (TCAV). *International Conference on Machine Learning*, 2668-2677. https://proceedings.mlr.press/v80/kim18d.html

Kim, M.G., Go, M.J., Kang, S.H., Jeong, S.H. and Lim, K., 2025. Revolutionizing CRISPR technology with artificial intelligence. *Experimental and Molecular Medicine*, 57, 1419-1431. https://doi.org/10.1038/s12276-025-01462-9

Kremer, M., Lubega, P., & Myers, S. (2019). *Digital agricultural extension: A review of the evidence*. World Bank Policy Research Working Paper No. 9020. https://doi.org/10.1596/1813-9450-9020

Krishnan, M. (2020). Against interpretability: A critical examination of the interpretability problem in machine learning. *Philosophy & Technology*, 33(3), 487-502. https://doi.org/10.1007/s13347-020-00411-7

Lazaros, K., Koumadorakis, D.E., Vlamos, P. and Vrahatis, A.G., 2024. Graph neural network approaches for single-cell data: a recent overview. *Neural Computing and Applications*, 36(17), 9963-9987. https://doi.org/10.1007/s00521-024-09656-8

Letham, B., Rudin, C., McCormick, T. H., & Madigan, D. (2015). Interpretable classifiers using rules and Bayesian analysis: Building a better stroke prediction model. *The Annals of Applied Statistics*, 9(3), 1350-1371. https://doi.org/10.1214/15-AOAS848

Li, L., Fan, Y., Tse, M., & Lin, K. Y. (2021). A review of applications in federated learning. *Computers & Industrial Engineering*, 149, 106854. https://doi.org/10.1016/j.cie.2020.106854

Li, W., Zhu, D. and Wang, Q., 2022. A single view leaf reconstruction method based on the fusion of ResNet and differentiable render in plant growth digital twin system. *Computers and Electronics in Agriculture*, 193, 106712. https://doi.org/10.1016/j.compag.2021.106712

Li, Z., Khan, W.U., Bai, G., Dong, C., Wang, J., Zhang, Y., Wang, C., Zhang, H., Wang, W. and Luo, M., 2025. From Code to Life: The AI-Driven Revolution in Genome Editing. *Advanced Science*, e17029. https://doi.org/10.1002/advs.202401702

Liang, X., Yu, S., Ju, Y., Wang, Y. and Yin, D., 2025. Multi-scale remote-sensing phenomics integrated with multi-omics: Advances in crop drought–heat stress tolerance mechanisms and perspectives for climate-smart agriculture. *Plants*, 14(**18**), 2829. https://doi.org/10.3390/plants14182829

Liebal, U.W., Phan, A.N., Sudhakar, M., Raman, K. and Blank, L.M., 2020. Machine learning applications for mass spectrometry-based metabolomics. *Metabolites*, 10(**6**), 243. (DOI not found - verify at https://doi.org/10.3390/metabo10060243

Lin, Y.M., Chen, C.T. and Chang, J.M., 2019. MS2CNN: predicting MS/MS spectrum based on protein sequence using deep convolutional neural networks. *BMC Genomics*, 20(**Suppl 9**), 906. (DOI not found - verify at https://doi.org/10.1186/s12864-019-6251-2

Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (2021). Explainable AI: A review of machine learning interpretability methods. *Entropy*, 23(1), 18. https://doi.org/10.3390/e23010018

Liu, J., Wang, L., Wang, Y., Xu, S. and Liu, Y., 2023. Research on the interface of sustainable plant factory based on digital twin. *Sustainability*, 15(**6**), 5010. (DOI not found - verify at https://doi.org/10.3390/su15065010

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765-4774. https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html

Mahood, E.H., Kruse, L.H. and Moghe, G.D., 2020. Machine learning: a powerful tool for gene function prediction in plants. *Applications in Plant Sciences*, 8(**7**), e11376. (DOI not found - verify at https://doi.org/10.1002/aps3.11376

Malone, T.W., Laubacher, R. and Dellarocas, C., 2010. The collective intelligence genome. *MIT Sloan Management Review*. (https://sloanreview.mit.edu/article/the-collective-intelligence-genome/)

Malosetti, M., Bustos-Korts, D., Boer, M.P. and van Eeuwijk, F.A., 2016. Predicting responses in multiple environments: issues in relation to genotype×environment interactions. *Crop Science*, 56(**5**), 2210-2222. https://doi.org/10.2135/cropsci2015.05.0316

Mansoor, S., Karunathilake, E., Tuan, T.T. and Chung, Y.S., 2025. Genomics, Phenomics, and Machine Learning in Shaping the Future of Plant Research: Advancements and Challenges. *Horticultural Plant Journal*. 11(**2**), 486-503. https://doi.org/10.1016/j.hpj.2023.09.005

Mansur, D. I., Hossain, M., Sow, D. M., & Rahman, M. (2022). Local language interfaces and agricultural technology adoption: Evidence from rural Malaysia. *Agricultural*

Economics, 53(2), 287-301. https://doi.org/10.1111/agec.12697

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2022). A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 54(6), 1-35. https://doi.org/10.1145/3457607

Mipeshwaree Devi, A., Khedashwori Devi, K., Premi Devi, P., Lakshmipriyari Devi, M. and Das, S., 2023. Metabolic engineering of plant secondary metabolites: prospects and its technological challenges. *Frontiers in Plant Science*, 14, 1171154. https://doi.org/10.3389/fpls.2023.1171154

Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, 7, 1419. https://doi.org/10.3389/fpls.2016.01419

Molnar, C. (2022). *Interpretable machine learning: A guide for making black box models explainable* (2nd ed.). Independently published. https://christophm.github.io/interpretable-ml-book/

Murmu, S., Sinha, D., Chaurasia, H., Sharma, S., Das, R., Jha, G. K. and Archak, S., 2024. A review of artificial intelligence-assisted omics techniques in plant defense: current trends and future directions. *Frontiers in Plant Science*, 15, 1292054. https://doi.org/10.3389/fpls.2024.1292054

Nakasone, E., & Torero, M. (2016). A text message away: ICTs as a tool to improve food security. *Applied Economic Perspectives and Policy*, 38 (2), 175-196. https://doi.org/10.1093/aepp/ppw007

Naskar, S., Sharma, S., Kuotsu, K., Halder, S., Pal, G., Saha, S., Mondal, S., Biswas, U. K., Jana, M. and Bhattacharjee, S., 2025. The biomedical applications of artificial intelligence: an overview of decades of research. *Journal of Drug Targeting*, 33(**5**), 717-748. https://doi.org/10.1080/1061186X.2025.2457045

Niazian, M. and Niedbała, G., 2020. Machine learning for plant breeding and biotechnology. *Agriculture*, 10(**10**), 436. https://doi.org/10.3390/agriculture10100436

Otal, H.T., Subasi, A., Kurt, F., Canbaz, M.A. and Uzun, Y., 2025. Analysis of gene regulatory networks from gene expression using graph neural networks. In *Digital Healthcare, Digital Transformation and Citizen Empowerment in Asia-Pacific and Europe for a Healthier Society* (pp. 249-270). Elsevier. https://doi.org/10.1016/B978-0-443-30168-1.00011-6

Pai, D. G., Balachandra, M., & Kamath, R. (2025). Explainable AI in agriculture: review of applications, methodologies, and future directions. *Engineering Research Express*, 7 (3), 032202. https://doi.org/10.1088/2631-8695/ae086a

Pan, L., Wang, A., Sang, R., Lu, X., Chen, W., Ma, Y. and Deng, F., 2025. AlphaFold 3 sheds insights into chemical enhancer-induced structural changes in Cas12a RNPs. *Health Nanotechnology*, 1(**1**), 1. https://doi.org/10.1186/s44301-024-00003-z

Peladarinos, N., Piromalis, D., Cheimaras, V., Tserepas, E., Munteanu, R.A. and Papageorgas, P., 2023. Enhancing smart agriculture by implementing digital twins: A comprehensive review. *Sensors*, 23(**16**), 7128. https://doi.org/10.3390/s23167128

Peng, L., Peng, M., Liao, B., Huang, G., Li, W. and Xie, D., 2018. The advances and challenges of deep learning application in biological big data processing. *Current Bioinformatics*, 13(**4**), 352-359. https://doi.org/10.2174/1574893612666170710121522

Raihan, T., Kim, C.H., Shimono, H., Kimura, A. and Iwata, H., 2026. ConvCGP: A convolutional neural network to predict genetic values of agronomic traits from compressed genome-wide polymorphisms. *The Plant Genome*, 19(**2**), e70223. https://doi.org/10.1002/tpg2.70223

Rane, N.L., Paramesha, M., Choudhary, S.P. and Rane, J., 2024. Machine learning and deep learning for big data analytics: A review of methods and applications. *Partners Universal International Innovation Journal*, 2(**3**), 172-197. https://doi.org/10.5281/zenodo.12271006

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135-1144. https://doi.org/10.1145/2939672.2939778

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206-215. https://doi.org/10.1038/s42256-019-0048-x

Sahoo, J.P., Mahapatra, S., Tripathy, S., Das, S., Mahapatra, S.R. and Dweh, T.J., 2026. Applications of artificial intelligence in plant genomics, genome editing, and biotechnology. In *AI Technologies for Crop Breeding* (pp. 129-145). *Elsevier*. https://doi.org/10.1016/B978-0-443-33633-1.00008-3

Samek, W., Montavon, G., Lapuschkin, S., Anders, C. J., & Müller, K. R. (2021). Explaining deep neural networks and beyond: A review of methods and applications. *Proceedings of the IEEE*, 109(3), 247-278. https://doi.org/10.1109/JPROC.2021.3060483

Sarkar, S., Dey, A., Pradhan, R., Sarkar, U.M., Chatterjee, C., Mondal, A. and Mitra, P., 2024. Crop yield prediction using multimodal meta-transformer and temporal graph neural networks.

IEEE Transactions on Agrifood Electronics. 2(**2**), 545 – 553. https://doi.org/10.1109/TAFE.2024.3438330

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE International Conference on Computer Vision*, 618-626. https://doi.org/10.1109/ICCV.2017.74

Shaik, R. and Ramakrishna, W., 2014. Machine learning approaches distinguish multiple stress conditions using stress-responsive genes and identify candidate genes for broad resistance in rice. *Plant Physiology*, 164(**1**), 481-495. https://doi.org/10.1104/pp.113.229146

Shea, A., Bartz, J., Zhang, L. and Dong, X., 2023. Predicting mutational function using machine learning. *Mutation Research-Reviews in Mutation Research*, 791, 108457. https://doi.org/10.1016/j.mrrev.2023.108457

Sheth, B.P. and Thaker, V.S., 2014. Plant systems biology: insights, advances and challenges. *Planta*, 240(**1**), 33-54. https://doi.org/10.1007/s00425-014-2081-9

Singh, D., Jain, N., Jain, P. and Kayal, P., 2020. PlantDoc: A dataset for visual plant disease detection. *Proceedings of the 7th ACM IKDD CoDS and 25th COMAD*, 249-253. https://doi.org/10.1145/3371158.3371196

Smith, R., Mathis, A.D., Ventura, D. and Prince, J.T., 2014. Proteomics, lipidomics, metabolomics: a mass spectrometry tutorial from a computer scientist's point of view. *BMC Bioinformatics*, 15(**Suppl 7**), S9. https://doi.org/10.1186/1471-2105-15-S7-S9

Srivastava, U., Kanchan, S., Kesheri, M., Gupta, M.K. and Singh, S., 2024. Types of omics data: genomics, metagenomics, epigenomics, transcriptomics, proteomics, metabolomics, and phenomics. In *Integrative Omics* (pp. 13-34). *Elsevier*. https://doi.org/10.1016/B978-0-443-16092-9.00002-3

Stewart, C.N., Jr., 2025. *Plant biotechnology and genetics: principles, techniques, and applications*. John Wiley and Sons. https://doi.org/10.1002/9780470282014

Su, W.H., 2020. Crop plant signaling for real-time plant identification in smart farm: A systematic review and new concept in artificial intelligence for automated weed control. *Artificial Intelligence in Agriculture*, 4, 262-271. https://doi.org/10.1016/j.aiia.2020.10.001

Tang, Y., Kurths, J., Lin, W., Ott, E. and Kocarev, L., 2020. Introduction to focus issue: When machine learning meets complex systems: Networks, chaos, and nonlinear dynamics. *Chaos: An*

Interdisciplinary Journal of Nonlinear Science, 30(**6**). https://doi.org/10.1063/5.0015454

Taylor, L., & Bhasme, S. (2021). Data justice and the digital transformation of agriculture. *Journal of Peasant Studies*, 48(5), 1028-1050. https://doi.org/10.1080/03066150.2021.1903429

Tripodi, P., Nicastro, N. and Pane, C., 2022. Digital applications and artificial intelligence in agriculture toward next-generation plant phenotyping. *Crop and Pasture Science*, 74(**6**), 597-614. https://doi.org/10.1071/CP22168

Tsan, M., Totapally, S., Hailu, M., & Addom, B. K. (2019). *The digital divide in agricultural extension: A review of ICT adoption among smallholder farmers*. ICTforAg Working Paper. (No DOI available)

Varadharajan, V., Rajendran, R., Muthuramalingam, P., Runthala, A., Madhesh, V., Swaminathan, G., Murugan, P., Srinivasan, H., Park, Y. and Shin, H., 2025. Multi-Omics Approaches Against Abiotic and Biotic Stress—A Review. *Plants*, 14(**6**), 865. https://doi.org/10.3390/plants14060865

Vats, S., Kumawat, S., Kumar, V., Patil, G.B., Joshi, T., Sonah, H., Sharma, T.R. and Deshmukh, R., 2019. Genome editing in plants: exploration of technological advancements and challenges. *Cells*, 8(**11**), 1386. https://doi.org/10.3390/cells8111386

Wang, H., Hu, R., Zhang, Y., & Zhang, X. (2021). Explainable deep learning for plant disease detection. *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 1842-1849. https://doi.org/10.1109/BIBM52615.2021.9669690

Wang, Y., Kong, S., Zhou, C., Wang, Y., Zhang, Y., Fang, Y. and Li, G., 2025a. A review of deep learning models for the prediction of chromatin interactions with DNA and epigenomic profiles. *Briefings in Bioinformatics*, 26(**1**), bbae651. https://doi.org/10.1093/bib/bbae651

Wang, Z., Yuan, H., Yan, J. and Liu, J., 2025b. Identification, characterization, and design of plant genome sequences using deep learning. *The Plant Journal*, 121(**1**), e17190. https://doi.org/10.1111/tpj.17190

Wen, T., Li, J.H., Wang, Q., Gao, Y.Y., Hao, G.F. and Song, B.A., 2023. Thermal imaging: The digital eye facilitates high-throughput phenotyping traits of plant growth and stress responses. *Science of the Total Environment*, 899, 165626. https://doi.org/10.1016/j.scitotenv.2023.165626

Wiseman, L., Sanderson, J., Zhang, A., & Jakku, E. (2019). Farmers and their data: An examination of farmers' willingness to share data. *Agriculture and Human Values*, 36(2), 289-302. https://doi.org/10.1007/s10460-019-09916-7

Wu, X., Lu, G. and Wu, Z., 2023. Remote sensing technology in the construction of digital twin basins: Applications and prospects. *Water*, 15(**11**), 2040. https://doi.org/10.3390/w15112040

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C. and Yu, P.S., 2020. A comprehensive survey on graph neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 32(**1**), 4-24. https://doi.org/10.1109/TNNLS.2020.3045381

Xie, L. and Wang, Y., 2025. The big data processing in agricultural genome-wide association studies: Challenges, technological advances, and future application prospects. *Advances in Resources Research*, 5(**4**), 1956-1976. https://doi.org/10.50908/arr.5.4_1956

Xu, F., Wu, T., Cheng, Q., Wang, X. and Yan, J., 2025. Foundation models in plant molecular biology: advances, challenges, and future directions. *Frontiers in Plant Science*, 16, 1611992. https://doi.org/10.3389/fpls.2025.1611992

Xu, X., Gao, F., Xiong, D., Fan, Z., Xiong, S., Dong, P., Qiao, H. and Ma, X., 2025. Digital twin-based winter wheat growth simulation and optimization. *Field Crops Research*, 329, 109953. https://doi.org/10.1016/j.fcr.2025.109953

Xu, Y., Zhang, X., Li, H., Zheng, H., Zhang, J., Olsen, M.S., Varshney, R.K., Prasanna, B.M. and Qian, Q., 2022. Smart breeding driven by big data, artificial intelligence, and integrated genomic-enviromic prediction. *Molecular Plant*, 15(**11**), 1664-1695. https://doi.org/10.1016/j.molp.2022.09.001

Yagaliyeva, B., Ivashchuk, O. and Goncharov, D., 2025. Modeling approaches for digital plant phenotyping under dynamic conditions of natural, climatic and anthropogenic factors. *Algorithms*, 18(**11**), 720. https://doi.org/10.3390/a18110720

Yetgin, A., 2025. Revolutionizing multi-omics analysis with artificial intelligence and data processing. *Quantitative Biology*, 13(**3**), e70002. https://doi.org/10.1002/qub2.70002

Yuan, Y. and Bar-Joseph, Z., 2019. Deep learning for inferring gene relationships from single-cell expression data. *Proceedings of the National Academy of Sciences*, 116(**52**), 27151-27158. https://doi.org/10.1073/pnas.1911797116

Zaorsky, N.G., Wang, X., Lehrer, E.J., Tchelebi, L.T., Yeich, A., Prasad, V.K., Chinchilli, V.M. and Wang, M., 2022. Retrospective comparative effectiveness research: Will changing the analytical methods change the results? *International Journal of Cancer*, 150(**12**), 1933-1940. https://doi.org/10.1002/ijc.33971

Zhang, X., Ibrahim, Z., Khaskheli, M.B., Raza, H., Zhou, F. and Shamsi, I.H., 2024. Integrative approaches to abiotic stress management in crops: combining bioinformatics educational tools and artificial intelligence applications. *Sustainability*, 16(**17**), 7651. https://doi.org/10.3390/su16177651

Zhang, Y., Huang, G., Zhao, Y., Lu, X., Wang, Y., Wang, C., Guo, X. and Zhao, C., 2025. Revolutionizing crop breeding: Next-generation artificial intelligence and big data-driven intelligent design. *Engineering*, 44, 245-255. https://doi.org/10.1016/j.eng.2025.01.015

Zhang, Y., Huang, G., Zhao, Y., Lu, X., Wang, Y., Wang, C., Guo, X., & Zhao, C. (2022). Revolutionizing crop breeding: Next-generation artificial intelligence and big data-driven intelligent design. *Engineering*, 44, 245-255. https://doi.org/10.1016/j.eng.2022.01.015

Zimmer, M. (2018). Addressing data governance in agricultural technology. *Data & Policy*, 1, e6. https://doi.org/10.1017/dap.2019.6