genes

MDPI

Review

Harnessing Multi-Omics and Predictive Modeling for Climate-Resilient Crop Breeding: From Genomes to Fields

Adnan Amin †, Wajid Zaman † and SeonJoo Park *

Department of Life Sciences, Yeungnam University, Gyeongsan 38541, Republic of Korea; adnan.amin@yu.ac.kr (A.A.); wajidzaman@yu.ac.kr (W.Z.)
* Correspondence: sjpark01@ynu.ac.kr
† These authors contributed equally to this work.

Abstract

The escalating impacts of climate change pose significant threats to global agriculture, necessitating a rapid development of climate-resilient crop varieties. The integration of multi-omics technologies—such as genomics, transcriptomics, proteomics, metabolomics, and phenomics—has revolutionized our understanding of the intricate molecular networks that govern plant stress responses. Coupled with advanced predictive modeling approaches such as machine learning, deep learning, and multi-omics-assisted genomic selection, these integrated frameworks enable accurate genotype-to-phenotype predictions that accelerate breeding for augmented stress tolerance. This review comprehensively synthesizes the current strategies for multi-omics data integration, highlighting computational tools, conceptual frameworks, and challenges in harmonizing heterogeneous datasets. We examine the contribution of digital phenotyping platforms and environmental data in dissecting genotype-by-environment interactions critical for climate adaptation resilience. Further, we discuss technical, biological, and ethical challenges, encompassing computational bottlenecks, trait complexity, data standardization, and equitable data sharing. Finally, we outline future directions that prioritize scalable infrastructures, interpretability, and collaborative platforms to facilitate the deployment of multi-omics-guided breeding in diverse agroecological contexts. This integrative approach possesses transformative potential for the development of resilient crops, ensuring agricultural sustainability amidst increasing environmental volatility.

Keywords: multi-omics integration; predictive modeling; climate-resilient crops; genotype to phenotype; digital phenotyping; genotype-by-environment interaction; machine learning; genomic selection

check for updates

Academic Editor: Quan Zou

Received: 18 June 2025
Revised: 4 July 2025
Accepted: 8 July 2025
Published: 10 July 2025

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

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Climate change has been increasingly threatening global food security through the exacerbation of abiotic stresses such as drought, heat, salinity, and flooding, which substantially reduce crop productivity and yield stability worldwide [1,2]. The frequency and severity of these environmental challenges are escalating, thereby compromising the resilience of many staple crops essential for human nutrition. Such stress conditions impose complex physiological and molecular constraints on plant growth and development, triggering dynamic alterations across multiple biological layers [3]. The pressing need to maintain agricultural productivity amid changing climate conditions drives the search for innovative strategies to enhance crop resilience beyond the capabilities of conventional breeding.

Conventional breeding techniques, while foundational to crop improvement over the past century, often struggle to keep pace with the rapid environmental changes and the complex genetic architecture governing stress tolerance traits [4]. These methods rely heavily on phenotypic selection and relatively simple marker-assisted approaches, which are constrained by limited throughput, the low resolution of polygenic trait architecture, and prolonged breeding cycles [5]. Moreover, environmental variability frequently masks genotype performance, hindering the identification and fixation of desirable alleles [6]. The polygenic and environmentally sensitive attributes of climate-adaptive traits necessitate a paradigm shift toward more precise, data-driven breeding frameworks capable of dissecting trait complexity at the molecular, biochemical, and physiological levels [7].

The integration of systems biology approaches, especially multi-omics methodologies, has transformed our understanding of plant responses to environmental stressors. Genomics, transcriptomics, proteomics, metabolomics, and phenomics collectively facilitate the comprehensive profiling of the molecular and phenotypic landscapes, elucidating the complex networks that regulate stress adaptation [8]. Advances in high-throughput sequencing, mass spectrometry, and imaging technologies generate vast, multidimensional datasets that reflect dynamic biological processes across scales [9]. Concurrent advancements in computational biology and machine learning offer powerful tools to integrate heterogeneous datasets, uncover hidden patterns, and build predictive models that link genotype to phenotype [10,11]. These integrative frameworks augment our ability to discern key regulators, biomarkers, and pathways that contribute to climate resilience, thereby accelerating the breeding pipeline.

Predictive modeling approaches, leveraging both classical machine learning algorithms and deep learning architectures, utilize multi-omics datasets to predict and model complex traits such as yield stability and stress tolerance under diverse environments [12,13]. The integration of digital phenotyping with environmental covariates refines the accuracy of such models, enabling the capture of genotype-by-environment interactions that majorly influence plant performance [14]. These data-driven strategies reduce the reliance on trial-and-error practices and enable breeders to prioritize candidate genotypes with enhanced accuracy. Moreover, multi-omics-informed genomic selection frameworks provide extensive insights by integrating genetic markers with transcriptomic and metabolic profiles, thereby increasing the prediction accuracy of breeding values [15].

This review synthesizes current advances in multi-omics technologies and predictive modeling as applied to climate-resilient crop breeding. It critically evaluates the attributes of individual omics platforms and their integration strategies, emphasizing computational tools and challenges associated with multi-layer data integration. The contributions of machine learning and deep learning in phenotype prediction and genomic selection are examined, alongside the incorporation of high-throughput digital phenotyping and environmental data to tackle genotype-by-environment interactions. Additionally, this review addresses the technical, biological, and ethical challenges facing multi-omics-guided breeding and proposes practical strategies for building robust, scalable breeding pipelines. This comprehensive synthesis provides a valuable resource for researchers and breeders aiming to harness multi-omics and predictive frameworks to sustainably improve crop resilience in the context of global climate change.

2. The Multi-Omics Toolkit in Plant Stress Biology

Research on plant stress responses and adaptation has been greatly enhanced by the use of multi-omics technologies, which collectively provide a systems-level perspective on how plants perceive, respond, and adapt to environmental stressors. Each omics layer, from genomics to phenomics, elucidates distinct facets of plant biology, offering comple-

mentary insights that collectively construct a comprehensive picture of stress adaptation mechanisms [8]. Genomics elucidates the genetic basis and structural variation responsible for stress resilience, whereas transcriptomics uncovers the dynamic regulation of gene expression and alternative splicing patterns in response to stress conditions [16]. Proteomics sheds light on functional proteins and post-translational modifications that mediate rapid cellular responses, while metabolomics profiles the biochemical environment that reflects stress-induced metabolic reprogramming [17]. Phenomics quantifies observable traits using advanced imaging and sensor technologies, facilitating the identification of direct associations between molecular changes and phenotypic outcomes [18]. The integration of these layers informs the development of predictive models and guides the breeding of crops with augmented stress resilience.

2.1. Genomics and Pan-Genomics in Crop Stress Adaptation

Genomics forms the foundation for understanding plant stress adaptation through the identification of DNA-level variations such as single-nucleotide polymorphisms (SNPs), copy number variations, and larger structural rearrangements [19]. Traditional single-reference genome assemblies capture only a fraction of the genetic diversity present within a species, thereby constraining the identification of critical (and in many occasions genotype- or environment-specific) adaptive loci [20]. To overcome these shortcomings, the pan-genome concept integrates multiple genomes from diverse accessions to categorize genes into core genes shared by all individuals, dispensable genes present in some but not all genotypes, and unique genes specific to particular genotypes [21]. This approach unveils hidden genetic variation, particularly in dispensable and unique regions that often harbor stress-responsive genes not present in reference genomes [22]. These variants play crucial roles in conferring tolerance to environmental stresses such as drought, salinity, and temperature extremes [23]. Genome-wide association studies (GWASs) and quantitative trait locus (QTL) mapping, utilizing dense SNP and structural variant data, enable the identification of genomic regions associated with stress tolerance traits, accelerating marker-assisted breeding [24]. This approach expands the genetic toolkit available for crop improvement beyond conventional limits (Figure 1).

2.2. Transcriptomics and Alternative Splicing Analyses Under Stress

Transcriptomic analyses reveal how plants orchestrate gene expression changes to adapt to stress conditions [25]. RNA sequencing (RNA-seq) provides a comprehensive analysis of genome-wide expression, identifying differentially expressed genes involved in stress signaling, defense, and repair [26]. Alternative splicing contributes to transcriptome plasticity by generating multiple transcript isoforms from individual genes, diversifying the proteome, and modulating stress responses [27]. Regulatory non-coding RNAs, such as microRNAs and long non-coding RNAs, mediate additional control by fine-tuning gene expression at the post-transcriptional or epigenetic level [28]. Cutting-edge single-cell RNA sequencing (scRNA-seq) technologies have begun to unravel cell-type-specific transcriptional responses, uncovering heterogeneity in stress adaptation at the cellular level [29]. Extensive transcriptomic datasets are now available for numerous plant species and stress conditions, providing a rich resource for comparative genomics and functional studies. Table 1 compiles key plant transcriptomic datasets, detailing the platforms and stress types, offering researchers a valuable guide for selecting suitable datasets for integrative analyses.

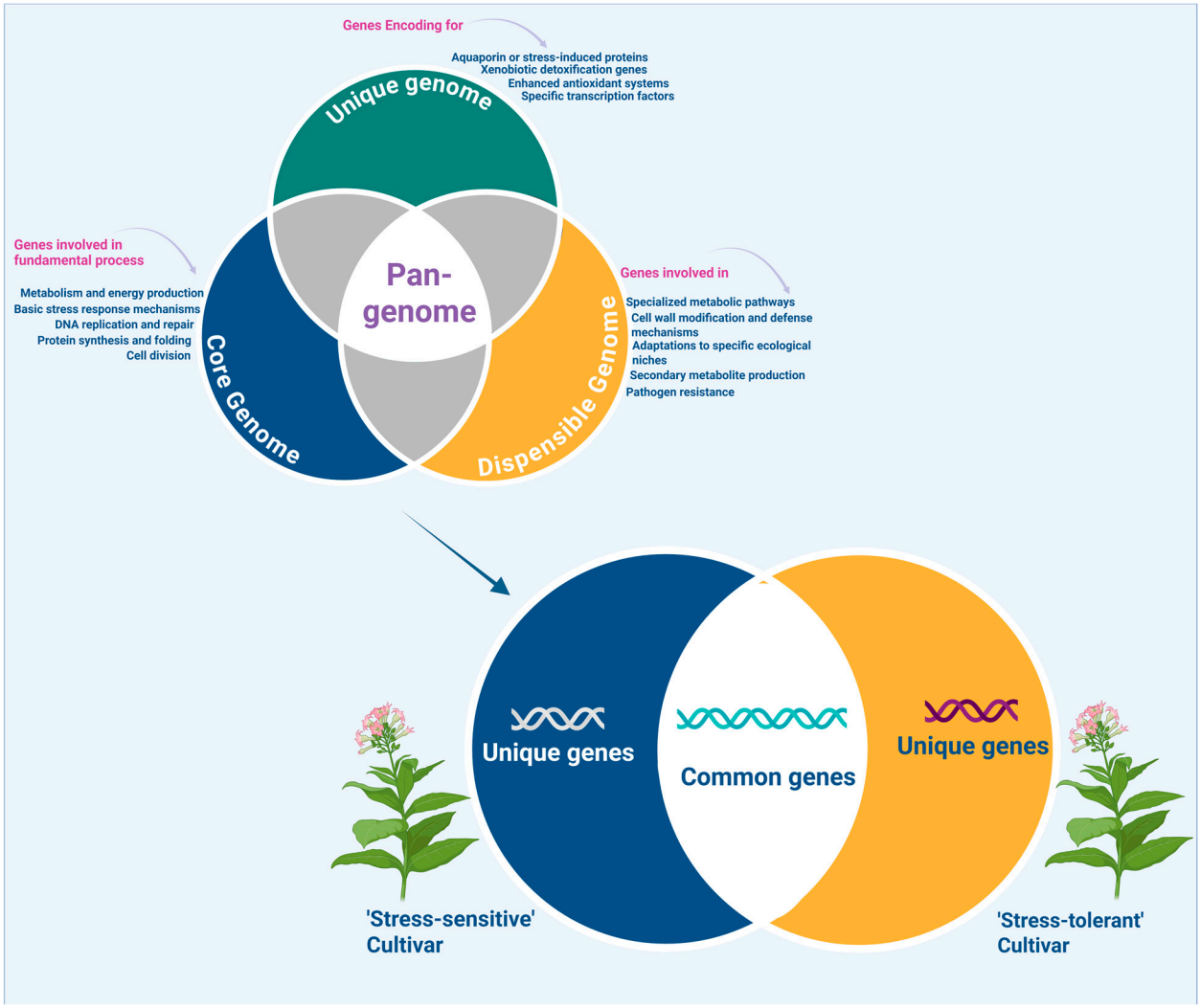

Figure 1. Schematic representation of pan-genome structure (core, dispensable, and unique genes) across 'stress-tolerant' and 'stress-sensitive' cultivars.

Table 1. Key plant transcriptomic datasets under abiotic/biotic stress with platforms (RNA-seq, scRNA-seq) and stress types.

Plant Species	Platform	Stress Type(s)	Key Findings	References
Arabidopsis thaliana	RNA-seq	Abiotic (heat and dehydration)	Stress increased full-length transcript variants via exon skipping in the *SR45a* gene	[29]
Arabidopsis thaliana	RNA-seq	Abiotic	42% of intron-containing genes were alternatively spliced; stress shifted isoform ratios	[30]
Populus trichocarpa	RNA-seq + Iso-Seq	Drought, salt, and temperature	Differential intron retention and isoform ratio switching across tissues	[31]
Cassava (Manihot esculenta)	Isoform-Seq, ssRNA-seq, Degradome-seq	Cold and drought	Intron retention dominant; cold stress altered splicing regulators and triggered transcript decay	[32]
Brassica napus	RNA-seq	Cold, salt, dehydration, and ABA	357 genes showed alternative splicing; hub genes linked to stress tolerance pathways	[33]

Table 1. *Cont.*

Plant Species	Platform	Stress Type(s)	Key Findings	References
Capsicum annuum (Pepper)	RNA-seq	Biotic (bacteria, virus, and oomycete)	4354 genes with stress-induced AS; 841.49 Gb data compiled from 425 samples	[34]
Zoysia japonica	RNA-seq	Cold	Dataset ZjRTD1.0 enables precise analysis of cold-induced splicing and co-regulation of isoforms	[35]
Schrenkiella parvula	Iso-Seq + RNA-seq	Salinity	Isoform diversity linked to salt tolerance; distinct isoform usage vs. *A. thaliana*	[36]
Glycine max (Soybean)	RNA-seq	Drought	Over 2000 genes alternatively spliced; splicing factors enriched under drought conditions	[37]
Camellia sinensis (Tea)	RNA-seq	Biotic (gray blight)	Splicing changes correlated with catechin biosynthesis; DASGs associated with disease defense	[38]

2.3. Proteomics and Post-Translational Modifications

Proteomics extends beyond transcriptional regulation to measure the abundance, modifications, and interactions of proteins carrying out cellular functions [39]. In response to stress, plants extensively remodel their proteomes, modifying enzyme activities, structural proteins, and regulatory components [40]. Post-translational modifications (PTMs) such as phosphorylation, ubiquitination, and acetylation act as molecular switches that rapidly adjust protein function, localization, and turnover, enabling plants to fine-tune their responses to environmental stressors [41,42]. Advances in mass spectrometry have enhanced the sensitivity and resolution of proteomic analyses, facilitating the comprehensive profiling of PTMs and protein networks. The 'protein–protein' interaction (PPI) analyses reveal clusters of proteins that act cooperatively in signaling pathways and stress response complexes [43]. The PPI network shown in Figure 2, generated under drought stress conditions in rice, exemplifies how proteomics uncovers key regulatory hubs that could be targeted for enhancing stress resilience through molecular breeding.

2.4. Metabolomics and Stress-Induced Biochemical Pathways

Metabolomics complements genomics and proteomics by profiling small molecules that reflect the physiological and biochemical state of plants under stress [8]. Both primary metabolites, such as sugars, amino acids, and organic acids, and secondary metabolites, including phenolics, alkaloids, and terpenoids, play vital roles in osmoprotection, antioxidant defense, signaling, and stress tolerance [44]. Advanced analytical platforms such as liquid chromatography–mass spectrometry (LC-MS), gas chromatography–mass spectrometry (GC-MS), and nuclear magnetic resonance (NMR) facilitate metabolite detection and quantification with high sensitivity. By integrating metabolomic data with transcriptomic and proteomic profiles, metabolic pathways perturbed during stress can be reconstructed, enabling the identification of key biomarkers predictive of tolerance [8]. Table 2 summarizes representative metabolite classes, their corresponding analytical techniques, and their relevance to various stress responses, serving as a valuable reference point for experimental design and interpretation.

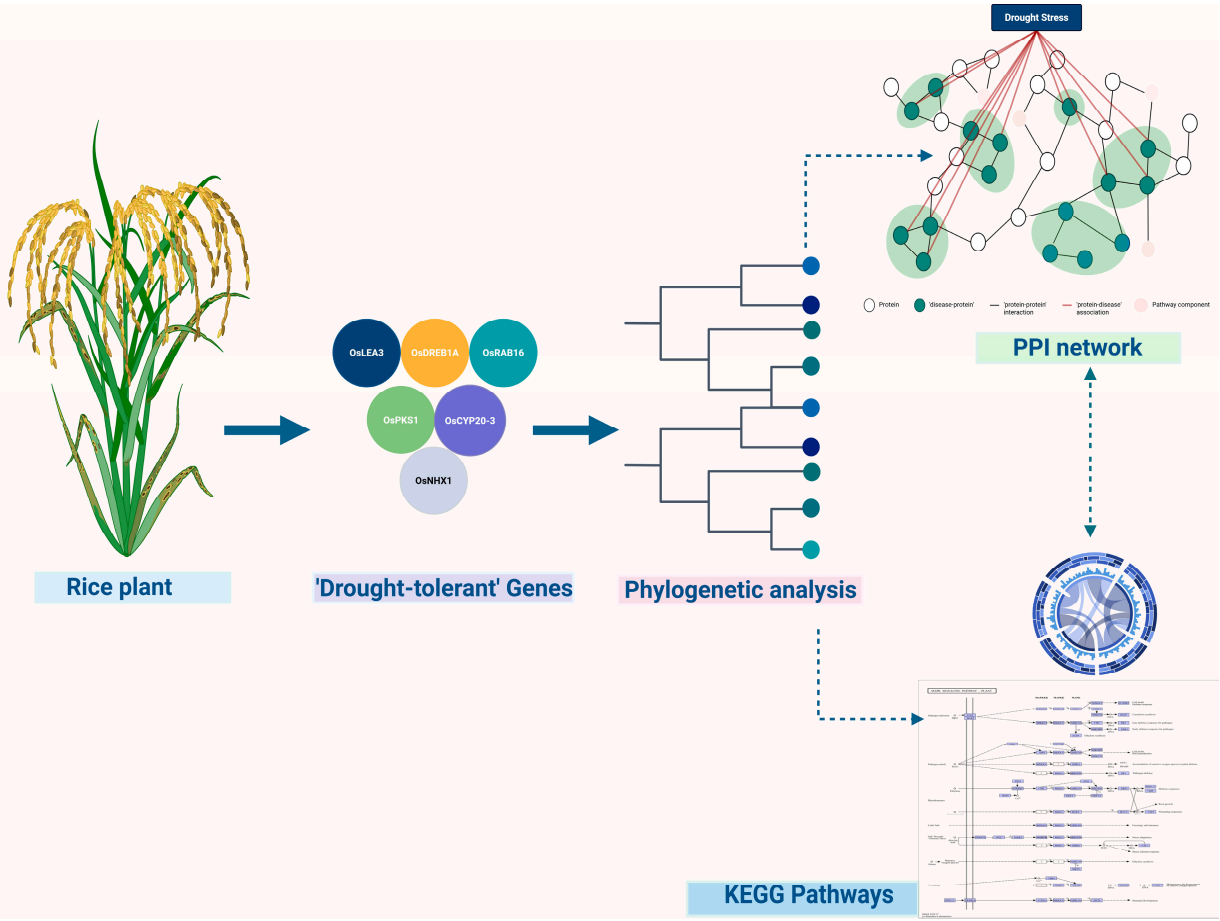

Figure 2. Protein–protein interaction (PPI) network and signaling pathway analysis under drought stress in rice.

Table 2. Representative metabolite classes, detection techniques, and their involvement in stress adaptation and tolerance.

Metabolite Class	Key Examples	Detection Techniques	Stress Type(s)	Physiological Functions	References
Sugars	Glucose, sucrose, trehalose, and raffinose	GC-MS, LC-MS, and NMR	Drought, salt, and cold	Osmoprotection, energy source, and stress signaling	[45]
Amino Acids	Proline, glutamate, GABA, and tryptophan	GC-MS, LC-MS, and NMR	Osmotic, oxidative, and heat	Osmotic balance, ROS scavenging, and precursors to secondary metabolites	[46,47]
Organic Acids	Malate, citrate, and fumarate	GC-MS and NMR	Drought, cold, and metal	Central in TCA cycle, energy production, and pH regulation	[48]
Phenolics	Flavonoids, phenolic acids, and lignans	LC-MS/MS and NMR	Oxidative, UV, and drought	Antioxidant activity, UV protection, and defense	[46,47]
Alkaloids	Serpentine, tabersonine, and vinblastine	Targeted LC-MS	Biotic (pathogen and herbivory)	Defense compounds and anti-insect and anti-fungal agents	[45]

Table 2. *Cont.*

Metabolite Class	Key Examples	Detection Techniques	Stress Type(s)	Physiological Functions	References
Terpenoids	Tanshinones and monoterpenes	LC-DAD-MS and GC-MS	Pathogen and drought	Antimicrobial properties, ROS modulation, and signaling	[46]
Fatty Acids	Linoleic acid and oleic acid	GC-MS and UPLC-QTRAP-MS	Heavy metal and cold	Membrane fluidity and precursors to signaling-related lipids	[49]
Sugar Alcohols	Mannitol and sorbitol	GC-MS and LC-MS	Salinity and drought	Osmoprotectants and ROS scavengers	[50]
Hormonal Intermediates	ABA precursors and auxin conjugates	LC-MS/MS	Drought, salinity, and temperature	Regulators of gene expression and stress adaptation	[51]
Shikimate Pathway Intermediates	Chorismate and phenylalanine	LC-MS and NMR	Drought and biotic stress	Links between primary and secondary metabolism; stress adaptation compound synthesis	[51]
Volatile Organic Compounds (VOCs)	Isoprene and linalool	GC-MS	Heat and biotic stress	Defense, communication, and thermotolerance	[52]

2.5. Phenomics for Stress Tolerance Trait Quantification

Phenomics harnesses advanced imaging and sensor technologies to capture complex phenotypic traits associated with stress tolerance [53]. High-throughput phenotyping (HTP) platforms utilize instruments and equipment including unmanned aerial vehicles (UAVs), multispectral and thermal cameras, fluorescence imaging, and root phenotyping chambers to measure traits such as the canopy temperature, leaf water content, stomatal conductance, chlorophyll fluorescence, leaf area, and root system architecture [54]. These non-destructive, high-resolution measurements facilitate the temporal and spatial monitoring of stress responses across many genotypes in both controlled and field environments [55]. Integrating phenotyping information with molecular data enhances the understanding of genotype–phenotype relationships and improves the accuracy of predictive breeding models [53]. Figure 3 illustrates a typical HTP setup combining UAVs, greenhouse imaging, and root phenotyping, highlighting the multifaceted nature of phenomic data acquisition.

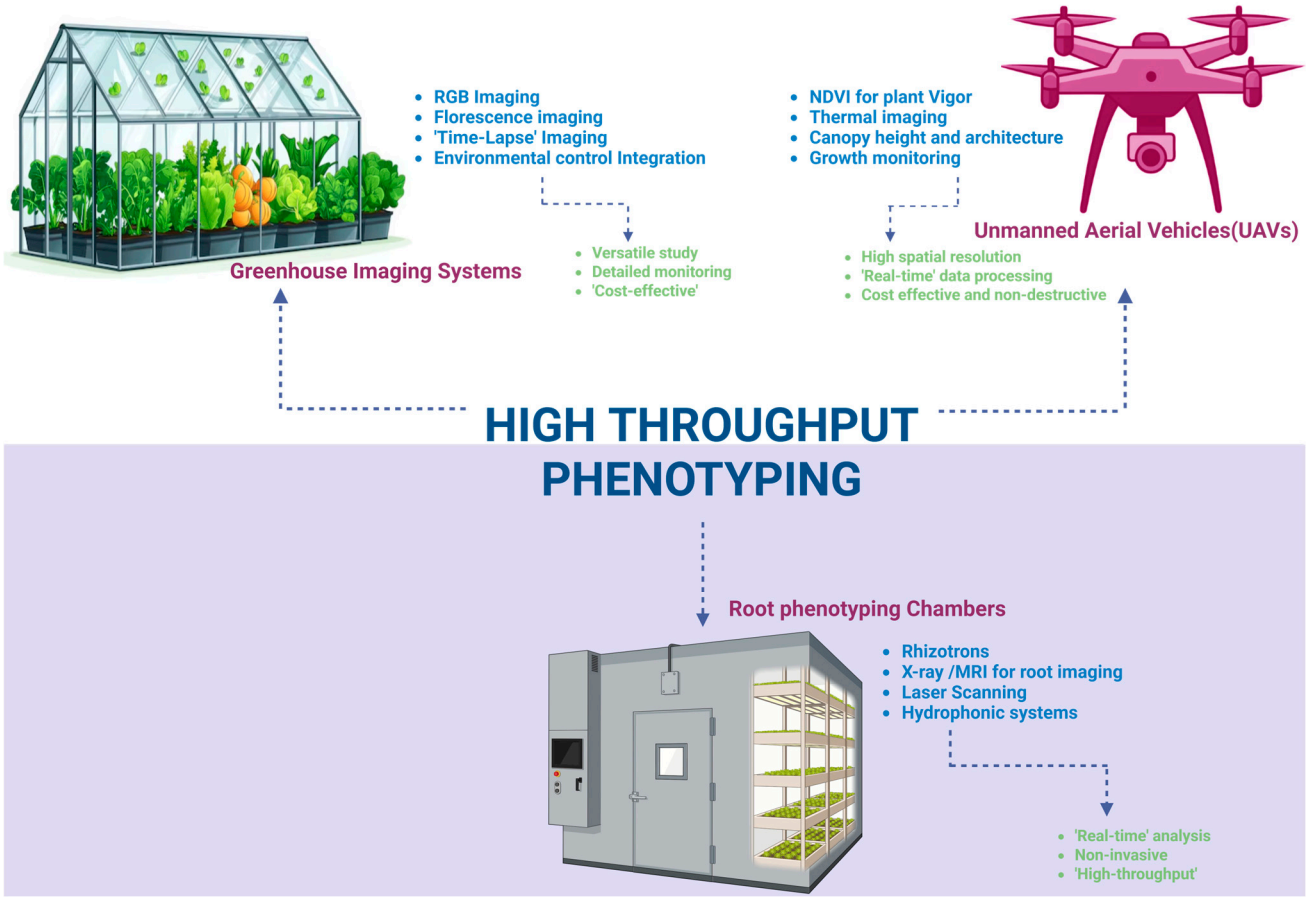

Figure 3. Schematic representation of a high-throughput phenotyping setup including UAVs, greenhouse imaging, and root phenotyping chambers with their key features and advantages.

3. Strategies for Multi-Omics Data Integration

Multi-omics data integration is crucial to elucidating complex traits like plant stress responses. Combining genomics, transcriptomics, proteomics, metabolomics, and phenomics outputs provides a multidimensional view of biological systems but requires sophisticated computational and statistical methods to effectively unify heterogeneous datasets [56]. The complexity arises from both the extensive volume and diversity of data types and from the biological intricacies of signaling pathways, gene regulation, and environmental interactions that these data represent [57]. Successfully integrating these layers uncovers hidden relationships and causal links, enabling improved predictive models of phenotype from genotype [58]. For example, the integration of transcriptomic and metabolomic data in rice under drought stress has revealed the coordinated regulation of abscisic acid biosynthesis and osmolyte accumulation, linking gene expression changes to metabolic adaptation [59]. Such integrative approaches are crucial for identifying robust biomarkers and key regulators for climate-resilient breeding.

3.1. Conceptual Frameworks for Omics Integration

Strategies for integrating multi-omics data can be broadly categorized into horizontal, vertical, and diagonal approaches [60]. Horizontal integration merges datasets of the same omics data type collected under different conditions or from diverse populations, thereby improving reproducibility and robustness [61]. For example, integrating transcriptome datasets from multiple drought stress experiments in maize enhances the reliability of the identified stress-responsive genes by emphasizing consistently regulated transcripts [62]. Vertical integration involves linking multiple omics layers measured on the same samples,

capturing biological cascades from DNA variation extending to RNA, a protein, and metabolite dynamics [63]. This approach can elucidate regulatory hierarchies; for instance, an integrative analysis in Arabidopsis linking SNPs, gene expression, protein abundance, and metabolic shifts under salt stress unraveled novel transcription factors governing tolerance [8,64]. Diagonal integration uses multi-omics data from distinct yet related samples connected by shared phenotypes or pathways, allowing inference when fully matched datasets are unavailable [65]. These conceptual models guide experimental design and computational workflows, with the choice depending on the study objectives and available data.

Beyond conceptual definitions, integration frameworks often employ network-based and multivariate statistical approaches. Horizontal and vertical models can be complemented by pathway and gene regulatory network reconstruction, facilitating biological interpretation [66]. A recent study integrated co-expression and protein interaction networks with metabolite correlation graphs to dissect drought tolerance in wheat, highlighting cross-omics modules predictive of yield stability [67]. The hierarchical structure of biological systems supports layered integration wherein outputs from one omics level inform analyses at another, thereby iteratively refining candidate genes and metabolite lists [68]. Visualizing these interactions through integrative frameworks, exemplified in Figure 4, elucidates the convergence of individual omics layers toward phenotype prediction.

3.2. Integration Tools and Data Fusion Platforms

A variety of computational tools have been developed to enable the robust integration of multi-omics datasets [69]. The mixOmics R package offers a comprehensive toolkit for multivariate analysis, encompassing canonical correlation analysis (CCA) and partial least squares (PLS), supporting the supervised and unsupervised integration of heterogeneous data [70]. For example, mixOmics facilitated the joint analysis of transcriptome and metabolome datasets in rice under cold stress, revealing molecular signatures associated with tolerance [71]. MOFA+ employs a Bayesian latent factor model that effectively manages missing data and detects both shared and dataset-specific variation across omics types [72]. It has been implemented in the integration of proteomics and metabolomics datasets in maize under nitrogen deficiency, revealing distinct, yet complementary, biological responses [73].

Network-based tools such as iOmicsPASS integrate multi-omics features with phenotype data to prioritize biomarkers and functional modules, useful in crop trait prediction [74]. Weighted Gene Co-expression Network Analysis (WGCNA) constructs modules based on correlated expression patterns and has been expanded to incorporate proteomics and metabolomics data, facilitating the identification of cross-omics trait-associated clusters [75]. DIABLO, part of the mixOmics suite, focuses on supervised integration for biomarker discovery and classification [76]. Table 3 delineates a comparison of these tools, highlighting their supported omics types, statistical methodologies, typical use cases, and limitations such as computational demand or sensitivity to data sparsity.

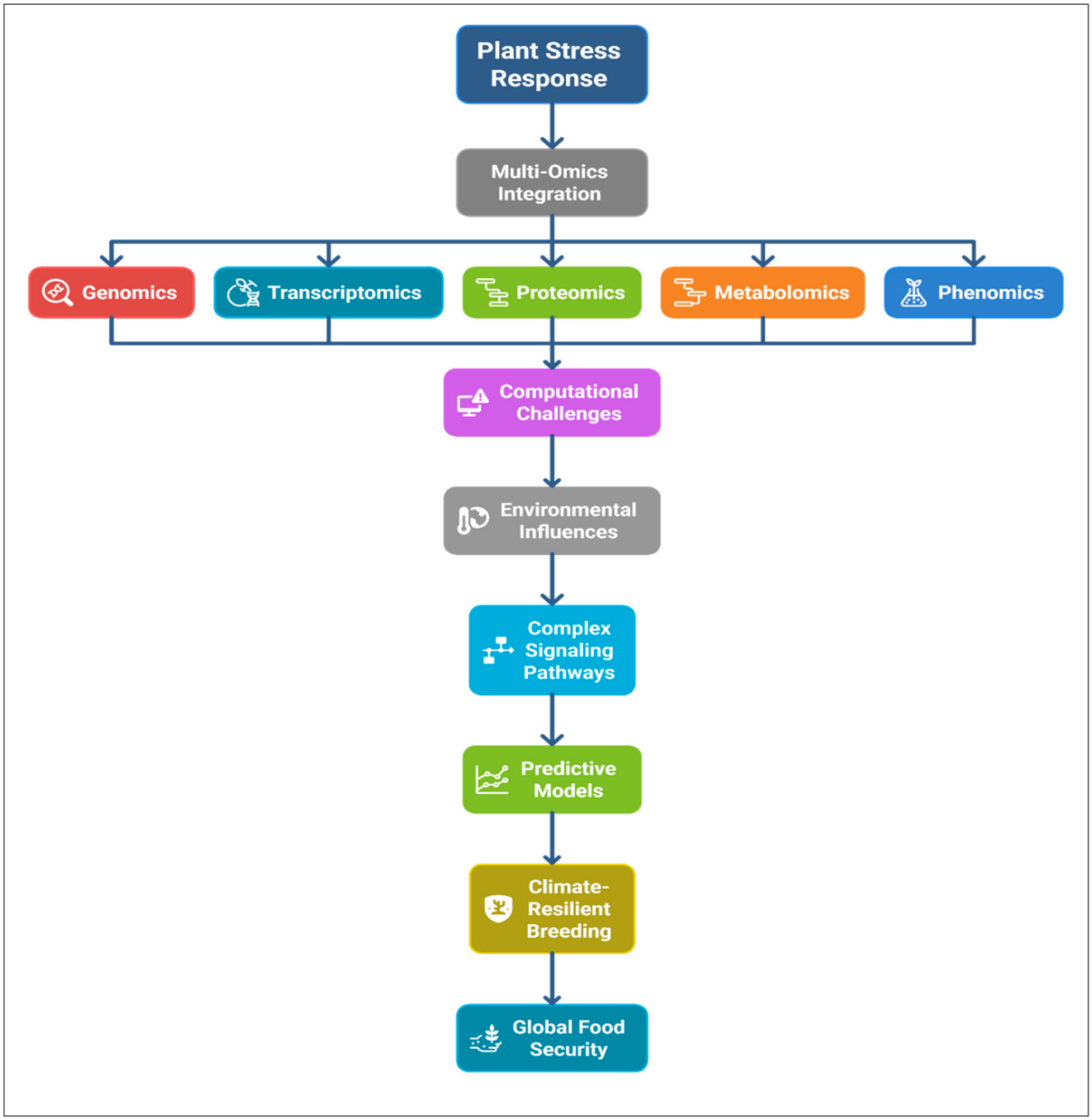

Figure 4. Layered integration framework illustrating genome–transcriptome–proteome–metabolome flow into phenotype prediction.

Table 3. Comparison of multi-omics integration tools: omics types, methods, and limitations.

Tool	Supported Omics Types	Statistical Methodology	Integration Approach	Data Handling	Visualization	Typical Use Case	Limitations	References
mixOmics	Transcriptomics, proteomics, metabolomics, microbiome, and epigenomics	PLS, CCA, sPLS, and DIABLO	Both (via different functions)	Handles missing data moderately well; requires scaling	Yes—heatmaps, networks, and correlation circles	Multi-omics classification, biomarker selection, and exploratory analysis	Sensitive to missing data; requires tuning; overfitting risk in small datasets	[70,77]
MOFA+	Any continuous omics: transcriptomics, epigenomics, proteomics, and metabolomics	Bayesian latent factor model	Unsupervised	Robust to missing values; normalizes data internally	Yes—factor plots and feature weights	Identification of shared and specific signals across -omics; sample stratification	Complex to interpret latent factors; high computational load for large datasets	[78]
iOmicsPASS	Transcriptomics, proteomics, metabolomics, and phenotypic data	Partial correlation networks + phenotype weighting	Supervised	Requires preprocessed complete matrices	Yes—modular network visualization	Disease/trait prediction; pathway–phenotype linkage	Phenotypic data needed; may struggle with very sparse networks	[79]

Table 3. *Cont.*

Tool	Supported Omics Types	Statistical Methodology	Integration Approach	Data Handling	Visualization	Typical Use Case	Limitations	References
WGCNA	Primarily transcriptomics, extendable to proteomics/metabolomics	Weighted correlation-based network analysis	Unsupervised	Requires complete data; sensitive to outliers	Yes—dendrograms and module–trait heatmaps	Identification of co-expressed modules, hub genes, and module–trait associations	Not natively multi-omics; manual integration required	[80]
DIABLO (in mixOmics)	Multi-block omics: transcriptomics, metabolomics, and proteomics	Supervised sparse PLS	Supervised	Performs variable selection; handles moderate sparsity	Yes—sample plots and relevance networks	Supervised feature extraction; class-based biomarker identification	Requires high-quality labels; less effective in unsupervised scenarios	[70]
Multi-Omics Factor Analysis (MOFA)	Transcriptomics, metabolomics, and epigenomics	Matrix factorization via variational inference	Unsupervised	Missing data allowed; scalable	Yes—dimensional reduction plots	Discovery of hidden factors driving variation across omics	Needs large sample sizes for meaningful factors; interpretability issues	[81]
IntLIM	Transcriptomics + metabolomics	Linear modeling with interaction terms	Supervised	Focused on two-omics comparisons	No (basic plots only)	Tests for phenotype-dependent omics interactions	Limited to two omics types; limited data scaling options	[82]
JIVE	Any omics (continuous)	Joint and individual variation explained	Unsupervised	Missing data imputation not supported	Limited (basic singular-value plots)	Decomposition of joint vs. specific signals across datasets	Requires manual feature interpretation; basic statistical output	[83]

3.3. Technical Challenges in Omics Fusion

Notwithstanding progress, the integration of multi-omics data continues to be hindered by technical and biological challenges. The heterogeneity of data types including discrete SNP genotypes, continuous expression values, and semi-quantitative metabolite abundances complicates normalization and scaling [84]. For instance, transcriptomics datasets frequently contain zero-inflated count data requiring specialized models, whereas metabolomics data may have batch effects and varying detection limits [85]. Missing data is prevalent due to incomplete sample profiling or technical failures; imputation methods exist but may introduce biases. The disparity in data dimensionality, characterized by one omics layer significantly exceeding the number of features in others, can result in dominance during integrative analyses, thereby obscuring weaker yet critical signals [69].

Biological complexity further complicates interpretation due to nonlinear relationships, pleiotropy, and epistasis, which hinder model construction [86]. Environmental effects introduce noise, especially in field-collected phenomics data [87]. Computational demands for processing high-dimensional multi-omics datasets require high-performance infrastructure, limiting accessibility. Emerging strategies such as transfer learning and federated data analysis seek to address certain limitations. The validation of integrative models using independent datasets or experimental assays remains essential but is often overlooked [88]. Addressing these challenges is critical for the implementation of multi-omics data integration as a routine tool in breeding programs.

4. Predictive Modeling for Trait Selection

The increasing availability of high-dimensional multi-omics datasets has revolutionized plant breeding by enabling the development of sophisticated predictive models that link genotypes to complex phenotypes [89,90]. These models harness machine learning (ML) and deep learning (DL) algorithms to identify nonlinear interactions and hidden patterns across genomic, transcriptomic, proteomic, metabolomic, and phenotypic layers [91]. The objective is to predict traits such as yield, stress tolerance, and disease resistance with higher accuracy, reducing reliance on laborious field trials and accelerating breeding

cycles [89]. Multi-omics integration enriches the input feature space by incorporating regulatory and metabolic information beyond genomic variants alone, thereby addressing the issue of missing heritability that constrains classical genomic selection [92]. In recent years, predictive modeling frameworks have evolved to incorporate environmental variables, longitudinal phenotyping data, and genotype-by-environment (G × E) interactions, enhancing the real-world applicability of predictions [93].

4.1. Machine Learning Algorithms for Trait Prediction

Machine learning algorithms, such as Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Support Vector Machines (SVMs), are extensively implemented in plant trait prediction due to their flexibility and ability to model complex nonlinear relationships [94]. RF constructs an ensemble of decision trees, each trained on bootstrapped samples, which collectively improve prediction robustness and handle high-dimensional multi-omics inputs with minimal parameter tuning [95]. To this end, Wu et al. (2024) applied RF models that integrated SNP genotypes and transcriptomic markers in maize subjected to drought conditions, achieving a predictive R^2 of 0.72 for grain yield, substantially outperforming single-omics models [96]. XGBoost, which sequentially optimizes decision trees to minimize error, can efficiently handle heterogeneous datasets and missing values [97] and was implemented to integrate genomic, metabolomic, and environmental data in rice, improving yield prediction accuracy. SVMs, relying on kernel methods, classify complex trait outcomes by mapping data into high-dimensional spaces. Liu et al. (2024) employed SVMs with multi-omics data to predict wheat disease resistance with higher accuracy [98]. These models typically necessitate extensive cross-validation and hyperparameter tuning to prevent overfitting and maximize generalizability. Their interpretability can be enhanced through the utilization of feature importance metrics, aiding breeders in prioritizing candidate loci or biomarkers. Figure 5 depicts a typical ML pipeline integrating multi-omics features to train models predicting phenotypes such as yield under drought.

4.2. Deep Learning for Omics-Guided Genotype-to-Phenotype Modeling

Deep learning enhances predictive capabilities by autonomously acquiring hierarchical representations from raw multi-omics data, excelling in capturing spatial, temporal, and nonlinear dependencies. Convolutional neural networks (CNNs), originally developed for image recognition, have been repurposed for genomic sequence analysis and spatial phenotyping data [99]. Liang et al. [100] employed CNNs to analyze soybean genomic and UAV-derived image data, thereby increasing stress tolerance prediction accuracy by 15% compared to traditional ML. Recurrent neural networks (RNNs), especially Long Short-Term Memory (LSTM) models, are well-suited for handling sequential data such as time-series transcriptomics and phenomics, capturing dynamic stress response trajectories [101]. Autoencoders execute unsupervised dimensionality reduction and denoising of noisy, high-dimensional omics datasets, facilitating the extraction of key features that improve downstream phenotype prediction [102]. Recent advances integrate multi-modal deep learning architectures that jointly process heterogeneous omics and imaging data, exemplified by maize yield prediction models integrating genomics, metabolomics, and multispectral images [103]. Table 4 summarizes performance metrics including the R^2, root mean square error (RMSE), and accuracy of various DL models applied to plant omics, highlighting their promise and current limitations related to model interpretability and data requirements.

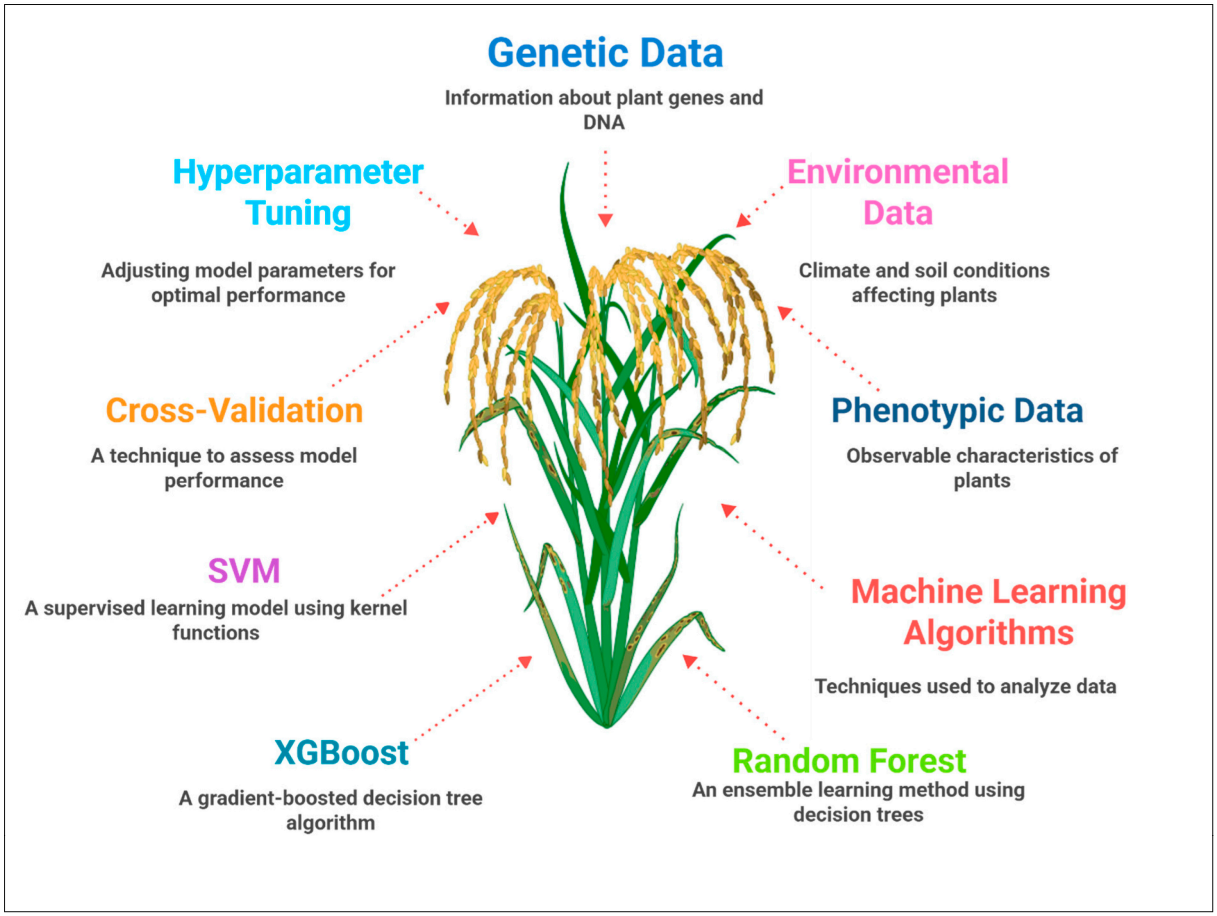

Figure 5. A detailed overview of different machine learning model training with multi-omics features for trait prediction.

Table 4. Performance metrics of DL models applied in plant omics prediction tasks (R^2, RMSE, and accuracy).

Model Type	Input Modalities	Task/Application	Performance Metrics	Key Strengths	Limitations	References
CNN (Convolutional Neural Network)	Genomic sequences + UAV-derived images	Soybean stress phenotype prediction	Accuracy increased 15% over baseline; R^2 ~0.71	Captures spatial dependencies; combines omics and imagery	Requires large, labeled image datasets	[104]
RNN/LSTM	Time-series transcriptome + stress phenotype data	Temporal modeling of stress responses	RMSE: 0.15–0.20; accuracy up to 89%	Effective for dynamic, sequential data	Sensitive to time-gap variation; needs careful tuning	[105]
Autoencoder	High-dimensional omics (e.g., transcriptome and metabolome)	Dimensionality reduction + phenotype prediction	RMSE: 0.12–0.18; comparable to LSTM; R^2 ~0.68	Denoises data; unsupervised feature extraction	Black-box interpretability; sensitive to latent dimension choice	[104]
Variational Autoencoder (VAE)	Genomics + imaging + environmental metadata	Multi-modal maize yield prediction	Accuracy: 85–90%; lower RMSE vs. linear models	Captures nonlinear joint distributions	Computational cost; sampling noise	[106]

Table 4. *Cont.*

Model Type	Input Modalities	Task/Application	Performance Metrics	Key Strengths	Limitations	References
Multi-modal DL Model	Genomics, metabolomics, and spectral imagery	Yield prediction under abiotic stress	R^2: 0.78; RMSE: 0.11; accuracy: ~88%	Integrates diverse-omics and environmental data	Requires harmonized, co-measured datasets	[107]
Hybrid CNN+LSTM	Genomic images + temporal gene expression	Combined spatial-temporal modeling	Accuracy: 90.2%; R^2 ~0.76; reduced overfitting compared to standalone models	Leverages strengths of both CNNs and RNNs	Higher model complexity	[108]

4.3. Genomic Prediction Enhancement with Multi-Omics Layers

Genomic selection (GS) predicts breeding values using genome-wide marker data and has emerged as a standard approach for complex trait improvement [109]. Nevertheless, GS models traditionally rely on genomic data alone, potentially overlooking intermediate molecular phenotypes that mediate gene-to-trait associations [110]. The integration of transcriptomic and metabolomic data improves GS efficacy by providing functional context and capturing regulatory variation influencing traits. For example, RNA-seq expression profiles were incorporated into GBLUP models for maize drought tolerance, resulting in a reported increase of over 5% in prediction accuracy compared to the use of SNP markers alone [111]. Similarly, it was observed that incorporating metabolite profiles into genomic data improved wheat grain protein content predictions, aiding selection for nutritional quality [112]. These models utilize Bayesian approaches like BayesA or GBLUP and increasingly adopt kernel-based and machine learning extensions to accommodate multi-omics inputs. Challenges include the need for large, matched multi-omics datasets, computational costs, and the modeling of genotype-by-environment interactions [113]. Nonetheless, multi-omics-assisted GS represents a powerful strategy for enhancing predictive breeding accuracy and accelerating the selection of climate-resilient crop varieties.

5. Digital Phenotyping and Environmental Interfacing

Digital phenotyping is rapidly transforming plant breeding by providing precise, high-throughput, and non-destructive measurements of plant traits across multiple scales, from individual organs to entire crop fields [54,114]. These technologies facilitate the continuous monitoring of plants under realistic environmental conditions, capturing the temporal dynamics of stress responses [115]. As climate change increases environmental variability, integrating digital phenotyping data with omics and environmental sensing has become crucial for the dissection of genotype-by-environment (G × E) interactions and the development of more predictive and robust breeding models [116]. Digital phenotyping platforms encompass aerial drones, ground sensors, and root imaging systems that collectively enable the comprehensive evaluation of physiological and morphological traits associated with stress tolerance [117], directly supporting precision breeding efforts.

5.1. High-Throughput Phenotyping Platforms

High-throughput phenotyping (HTP) platforms have evolved to combine multiple sensing modalities, enabling detailed trait measurements at unprecedented speed and scale. UAVs equipped with multispectral, thermal infrared, and RGB cameras can rapidly assess canopy temperature, leaf area index, vegetation indices such as NDVI, and indicators of

stress such as wilting or chlorosis symptoms across large field trials [118,119]. These aerial platforms can capture spatial heterogeneity and temporal changes in plant performance throughout the growing season. For example, thermal imaging has been successfully implemented to uncover heat- and drought-tolerant wheat genotypes by measuring canopy temperature depression, a proxy for transpiration efficiency [120].

In conjunction with UAVs, ground-based sensors such as chlorophyll fluorometers are employed to measure photosynthetic efficiency, providing insights into stress-induced damage at the leaf level [121]. Infrared gas analyzers can estimate stomatal conductance and transpiration rates, which are major indicators for extrapolating water use efficiency under drought [122]. Root phenotyping systems, including X-ray computed tomography and rhizotron imaging, can be implemented to accurately quantify belowground traits like root length, branching, and depth, which are critical for nutrient and water acquisition but traditionally difficult to measure [123]. These high-resolution datasets, collected non-invasively and repeatedly, enable dynamic phenotyping that captures the progression of stress responses, providing extensive temporal datasets essential for linking to underlying stress adaptation molecular mechanisms. Table 5 details key stress-related traits measurable by different HTP technologies and their corresponding sensing methods.

Table 5. Traits measurable by high-throughput phenotyping (HTP) platforms and associated technologies (e.g., canopy temperature via infrared sensor).

Trait Measured	HTP Platform/Modality	Sensor Type/Technique	Biological Relevance	References
Canopy temperature	UAV-based aerial thermal imaging	Thermal infrared cameras	Proxy for stomatal conductance and transpiration under heat/drought stress	[124]
NDVI (Normalized Difference Vegetation Index)	UAV multispectral imaging	Multispectral cameras (Red/NIR)	Indicator of photosynthetic activity and biomass	[125]
Chlorosis/leaf senescence	RGB + multispectral UAV	RGB and spectral indices (e.g., GNDVI and SAVI)	Visual cues of nutrient stress, senescence, and disease	[126]
Photosynthetic efficiency (ΦPSII)	Ground-based sensor	Pulse-amplitude-modulated (PAM) chlorophyll fluorometry	Captures efficiency of light reactions in photosynthesis under stress	[127]
Stomatal conductance (gs)	Portable leaf gas analyzers	Infrared gas analyzer (IRGA)	Measures gas exchange related to water loss and carbon uptake	[128]
Root system architecture	X-ray CT, rhizotrons, and Shovelomics	High-resolution 3D imaging or transparent interface	Essential for belowground trait monitoring under drought or nutrient stress	[129]
Canopy structure/LAI	UAV LiDAR + multispectral fusion	Light Detection and Ranging (LiDAR) + NDVI	Reflects total photosynthetic surface and canopy penetration	[130]
Water use efficiency (WUE)	UAV-based multispectral + ET modeling	NDVI-derived biomass + evapotranspiration estimates	Assesses yield relative to water use; key trait under water-limited conditions	[131]
Transpiration rate	Proximal thermal imaging in automated systems	Leaf surface temperature profiles over time	Indicates water loss dynamics and drought response	[132]
Plant height/growth rate	Time-lapse 3D LiDAR or stereo vision	Structure-from-motion (SfM) and laser range scanning	Non-invasive quantification of growth dynamics	[129]

5.2. Modeling G × E Interactions in Predictive Breeding

Genotype-by-environment (G × E) interactions profoundly affect plant performance, especially under fluctuating climate conditions [133]. Traditional statistical models like Additive Main Effects and Multiplicative Interaction (AMMI) and genotype plus genotype-by-environment interaction (GGE) biplots have been fundamental in assessing G × E by partitioning variance components and visualizing genotype stability and adaptability across environments [134,135]. These approaches discern genotypes exhibiting broad adaptability or specific niche suitability. Recent advancements have integrated spatial statistical models and machine learning algorithms to better capture environmental gradients and local site effects, thereby improving prediction accuracy. For instance, Tesfaye et al. (2016) applied spatial mixed models incorporating soil moisture and temperature sensors, increasing the accuracy of drought tolerance predictions in maize breeding [136].

Biplots generated by GGE (genotype + genotype × environment interaction) analysis serve as a powerful and intuitive tool for visualizing genotype performance across diverse environments and traits, offering critical insights into genotype–environment interactions [137,138]. Such biplots plot genotypes (typically different breeding lines or varieties) against multiple traits (such as yield, disease resistance, drought tolerance, etc.) under diverse environmental conditions, allowing for the identification of patterns in genotype performance. By assessing the performance of various genotypes across diverse environments, breeders can identify those that demonstrate stability and superior traits under specific environmental stresses [139]. For instance, some genotypes may perform exceptionally well in drought-prone areas but show poor performance in regions with higher rainfall, while others may demonstrate resilience to both drought and heat stress [140]. This visual representation facilitates the selection of genotypes with superior traits, allowing breeders to prioritize those that are best suited for specific target environments. Furthermore, GGE biplots can contribute to the delineation of genotype × environment interactions by separating the effects of the genotype from those of the environment, providing a clearer picture of how genetic factors and environmental conditions jointly influence trait expression [141]. Through such analyses, breeders can make more informed decisions, targeting genotypes with the greatest yield stability potential, stress resistance, and adaptability [142], thereby optimizing breeding programs aimed at improving crop performance under fluctuating environmental conditions. The biplot's intuitive design facilitates the communication of complex breeding data and results, thereby improving collaboration and decision-making in crop improvement strategies [143]. Moreover, integrating these models with omics-informed predictive frameworks establishes a holistic selection system that accounts for both genetic potential and environmental responsiveness.

5.3. Integrating Climate and Omics Data

The integration of high-resolution environmental data with omics and high-throughput phenotyping data represents a frontier in predictive breeding [116]. Satellite remote sensing provides spatially explicit data on climatic variables such as temperature, precipitation, soil moisture, and vegetation health at regional scales, whereas Internet of Things (IoT) sensors installed in experimental fields can record and transmit fine-scale microclimatic conditions including humidity, solar radiation, and wind speed [144,145]. Integrating these environmental data streams with genomic, transcriptomic, metabolomic, and phenomic profiles enables the formulation of environment-aware selection indices and predictive models that reflect real-world growing conditions.

To this end, Singh et al. [146] integrated hourly temperature and humidity data with multi-omics and phenomics datasets in wheat, improving the prediction of heat stress tolerance and facilitating the selection of genotypes tailored for specific climatic

niches. This approach supports dynamic breeding pipelines that adapt to spatial and temporal environmental variability, crucial for future-proofing crops against climate change. Challenges remain in the standardization of data formats, the management of extensive and heterogeneous datasets, and the development of computational frameworks that can concurrently model genotype, phenotype, and environmental features [147]. Nevertheless, this integrative approach paves the way for precision breeding strategies that maximize genetic gains while ensuring stability under variable climates.

6. Challenges and Opportunities

While multi-omics integration and predictive modeling have immense potential to accelerate climate-resilient crop breeding, several significant challenges limit their full realization and implementation. Such challenges encompass computational bottlenecks, biological complexities, data governance, and infrastructure gaps [148]. Addressing these obstacles requires interdisciplinary efforts involving computational scientists, plant biologists, breeders, and policymakers. Despite these challenges, innovative solutions and collaborative frameworks offer promising pathways to harness multi-omics for practical, scalable, and equitable crop improvement.

6.1. Computational and Data Integration Bottlenecks

One of the foremost challenges in multi-omics research lies in the computational infrastructure and methods required to process, store, and analyze vast heterogeneous datasets [149]. Scalability issues arise as omics datasets expand in both volume and complexity, demanding high-performance computing resources often inaccessible to many research groups. The heterogeneous nature of omics data ranging from discrete genomic variants to continuous transcript and metabolite abundance complicates normalization and harmonization [69]. Integrating RNA-seq counts, proteomic intensities, and metabolite concentrations requires careful statistical adjustments to ensure comparability and prevent biases in downstream analyses [63]. Moreover, batch effects, missing values, and data sparsity present substantial obstacles that can distort integrative models. Cloud computing platforms such as Terra, DNAnexus, and CyVerse have started to address these challenges by providing scalable environments, standardized workflows, and collaborative spaces [150]; however, adoption in plant sciences is sporadic.

Advanced algorithms and software tools tailored for multi-omics data integration also demonstrate limitations in handling large-scale datasets with missing or noisy data. The lack of standardized data formats and metadata annotation exacerbates interoperability issues, hindering data reuse and cross-study comparisons [151]. Table 6 outlines common computational bottlenecks associated with distinct omics layers—such as peak calling in metabolomics or variant calling in genomics—and highlights mitigation strategies including normalization methods, imputation techniques, and algorithmic advances. Addressing and resolving these computational challenges is critical to enabling reliable multi-omics integration and subsequent application in breeding programs.

Table 6. Omics-specific computational bottlenecks and mitigation strategies.

Omics Layer	Computational Bottleneck	Mitigation Strategy	Explanation	References
Genomics	Variant calling variability; batch-specific sequencing biases	Standardized pipelines (e.g., GATK); batch correction tools like reComBat	Ensures consistent variant detection across batches; tools like reComBat mitigate batch effects in heterogeneous datasets	[152]
Transcriptomics	Integration of RNA-seq datasets with missing samples and inconsistent coverage	Bayesian models (e.g., TiMEG); advanced imputation and joint modeling	Models like TiMEG handle partially missing transcript data without requiring full imputation, improving integration reliability	[153]
Proteomics	High frequency of missing values due to detection limits and instrumentation	Missing-value-tolerant frameworks (e.g., matrix dissection and no imputation)	Mitigates imputation biases by correcting only where data are present, maintaining statistical integrity	[154]
Metabolomics	Peak alignment errors; inconsistent quantification across batches	Transformation and normalization pipelines; batch-aware preprocessing	Careful preprocessing and normalization mitigate artifacts from instrumentation and sample prep variability	[155]
Epigenomics	Data sparsity; modality-specific biases (e.g., methylation depth vs. accessibility)	Cross-modal imputation; graph neural networks; contrastive learning	Emerging AI methods (e.g., SpaMosaic) reconstruct missing modalities and enable high-fidelity integration	[156]
Multi-omics (general)	Heterogeneous data types and missing modalities	Deep generative models (e.g., VAEs); multi-modal latent space integration	Variational autoencoders and other AI tools embed omics data into shared spaces, enabling joint analysis and imputation	[157,158]
All layers (meta-level)	Batch effects coupled with missing data (BEAMs: Batch-Effect-Associated Missing Values)	Hybrid methods accounting for batch and missingness (e.g., MultiBaC, BEAM-aware imputation)	MultiBaC exploits shared modalities across datasets to correct for lab-specific artifacts; BEAM-aware workflows prevent imputation bias	[159,160]
Single-cell omics	High sparsity and noise in individual cell profiles; multi-modal alignment	Product-of-Experts VAEs; latent space alignment and batch-aware learning	Enables robust integration across scRNA, ATAC, and protein modalities in sparse, high-dimensional data	[161]

6.2. Biological Interpretation and Trait Complexity

Multi-omics analyses are further complicated by the biological complexities inherent in plant stress responses. Numerous agronomically important traits exhibit pleiotropy, where a single gene affects multiple phenotypes, as well as epistasis, which entails interactions between genes that collectively influence traits [162]. Such nonlinear genetic architectures can obscure causal relationships when analyzed through traditional linear models. Moreover, a large portion of heritable variation, referred to as "missing heritability", remains unexplained even after incorporating genomic data, suggesting that regulatory mechanisms, epigenetics, and gene–environment interactions are likely major influential factors [163]. Environmental variability, including microclimatic differences and soil heterogeneity, adds noise and complexity, challenging the interpretation of omics–phenotype associations.

The complexity of integrating multiple molecular layers increases the difficulty of biological interpretation; for instance, changes in transcript levels do not always translate directly into changes in protein or metabolite levels due to post-transcriptional and post-translational regulation [164]. Therefore, careful experimental design, rigorous statistical modeling, and functional validation are essential for the precise interpretation of multi-omics results. Systems biology approaches, including network reconstruction and causal inference, can help disentangle complex interactions but remain computationally intensive and dependent on high-quality data [165]. Continued efforts to improve the biological

interpretability of integrated models are vital to translating omics insights into actionable breeding strategies.

6.3. FAIR Data Principles and Ethical Considerations

Ensuring that multi-omics data are Findable, Accessible, Interoperable, and Reusable (FAIR) is crucial for enhancing research impact and fostering reproducibility. Open data repositories such as NCBI GEO, EMBL-EBI, and MetaboLights facilitate broad access, but many datasets lack standardized metadata or consistent formatting, limiting integration [166]. The adoption of FAIR principles facilitates improved data sharing among institutions and bolsters global collaborative breeding initiatives [167]. However, challenges related to intellectual property rights (IPRs), especially in publicly funded research and private breeding programs, create tensions between open science and proprietary interests. Equitable benefit sharing with indigenous communities and countries of origin for genetic resources is also a critical ethical consideration, as mandated by international agreements like the Nagoya Protocol [168]. Figure 6 depicts the FAIR data life cycle in multi-omics projects, highlighting stages from data generation and annotation through deposition, discovery, reuse, and reanalysis. The integration of ethical frameworks and data governance policies with FAIR practices will be key to fostering trust and ensuring that multi-omics advances inclusively benefit global agriculture. Efforts to establish clear guidelines, data stewardship roles, and transparent licensing models are underway but require wider adoption and harmonization.

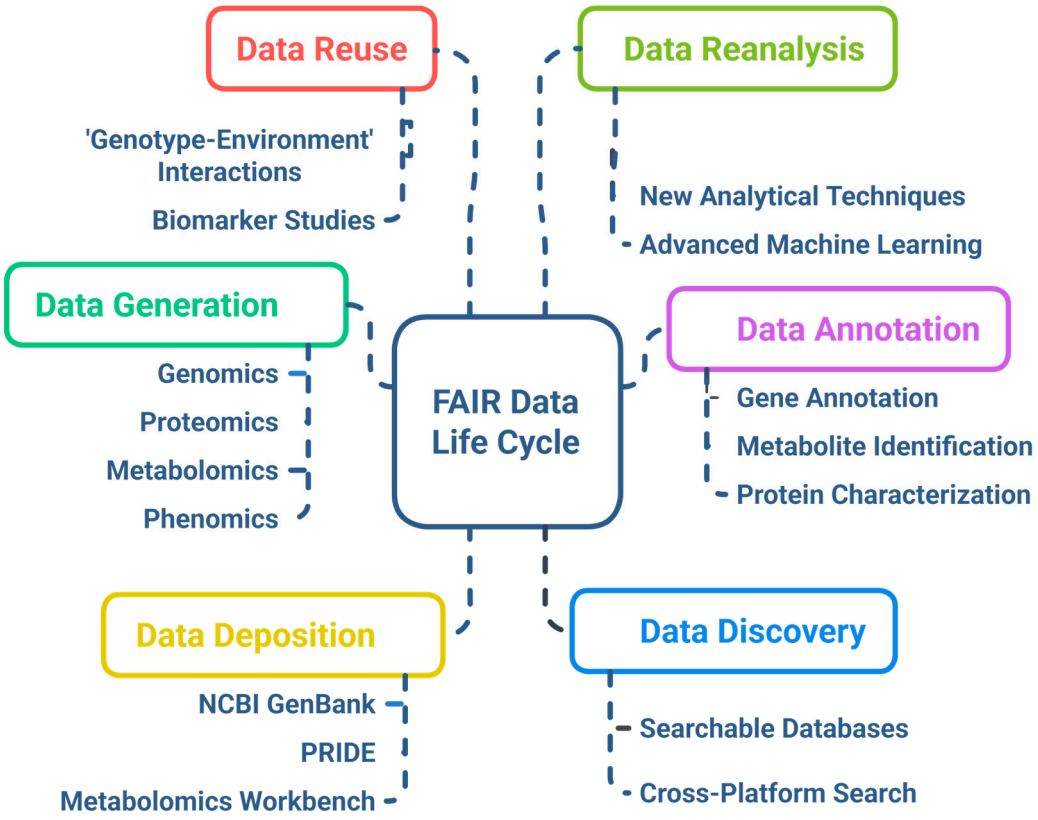

Figure 6. FAIR data life cycle implemented in multi-omics projects.

6.4. Roadmap for Future Integration

The future of multi-omics-guided crop breeding depends on the establishment of global platforms, consortia, and open-source tools that democratize access and foster collaboration. International initiatives such as the Crop Ontology Consortium and the

Wheat Initiative exemplify efforts to standardize trait definitions and data-sharing protocols, enhancing interoperability [169]. Cloud-based infrastructures that accommodate multi-omics datasets alongside scalable analysis pipelines reduce barriers for researchers and breeders, enabling real-time data integration and decision-making [170]. Additionally, open-source software frameworks promote transparency, community contributions, and reproducibility, thereby enabling continuous improvement and adaptation to emerging technologies [171].

Training the next generation of interdisciplinary scientists skilled in both computational and plant biology domains is crucial to sustain progress [172]. Integrative research models that connect molecular biologists, data scientists, breeders, and policy experts will accelerate translation from data to deployed improved varieties [12]. Establishing and implementing common standards for data formats, metadata annotation, and model reporting will further harmonize the field. Through coordinated efforts, multi-omics integration will evolve from an experimental approach to a routine, scalable component of climate-resilient breeding pipelines.

7. Conclusions

The integration of multi-omics technologies with advanced predictive modeling has revolutionized crop breeding, especially in the context of developing climate-resilient varieties. Through the integration of genomic, transcriptomic, proteomic, metabolomic, and phenomic data, unprecedented insights can be obtained regarding the molecular mechanisms and complex trait architectures underlying stress adaptation. Predictive models, including machine learning, deep learning, and multi-omics-informed genomic selection, leverage these integrated datasets to accurately predict phenotypic performance under variable environmental conditions. Such integrated approaches surpass traditional breeding methods by enabling the identification of robust biomarkers, the dissection of genotype-by-environment interactions, and the acceleration of selection cycles. Multi-omics and predictive analytics are collectively facilitating a paradigm shift from phenotype-based selection toward data-driven precision breeding, providing optimism for sustaining agricultural productivity amid increasing climate challenges.

Despite these advances, realizing the full potential of multi-omics-guided breeding requires the resolution of several key issues and obstacles. The standardization of data generation, processing, and integration protocols is essential to ensure reproducibility, comparability, and interoperability across studies and institutions. Establishing unified ontologies, metadata standards, and data-sharing frameworks will enable seamless aggregation and meta-analyses of heterogeneous datasets. The interpretability of predictive models is equally critical: breeders and biologists must be able to understand and trust the biological relevance of model predictions to make informed decisions. Advances in explainable AI and integrative network analyses are paving the way for more transparent models that link predictive features to functional biology. Furthermore, it is essential to prioritize scalability and practical applicability by developing computational tools and phenotyping platforms that are accessible, cost-effective, and deployable across various breeding programs worldwide, including in resource-limited settings.

Looking forward, the continued convergence of multi-omics technologies, environmental data integration, and predictive modeling will drive the next generation of climate-smart crop breeding. Collaborative international efforts to build open-source platforms, cloud-based infrastructures, and global consortia will democratize access to cutting-edge tools and data. Training interdisciplinary scientists capable of bridging computational biology, plant science, and breeding will be vital to sustain innovation. As these approaches mature, they are expected to deliver resilient crops tailored to specific agroecological niches, thereby

enhancing food security and agricultural sustainability. Ultimately, the integration of comprehensive biological data with robust predictive models marks a new era of precision agriculture that meets the challenges that arise from a changing climate.

Author Contributions: Conceptualization, writing—original draft preparation—resources, software, validation, and visualization, A.A. and W.Z.; writing—review and editing—and supervision, S.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the 2025 Yeungnam University Research Grant 225A380088.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Leisner, C. Review: Climate change impacts on food security- focus on perennial cropping systems and nutritional value. *Plant Sci. Int. J. Exp. Plant Biol.* **2020**, *293*, 110412.
2. Raza, A.; Safdar, M.; Adnan Shahid, M.; Shabir, G.; Khil, A.; Hussain, S.; Khan, M.; Aziz, S.U.R.; Azam, S.; Sattar, J. Climate Change Impacts on Crop Productivity and Food Security: An Overview. In *Transforming Agricultural Management for a Sustainable Future Climate Change and Machine Learning Perspectives*; Springer: Berlin/Heidelberg, Germany, 2024; pp. 163–186.
3. Chaudhry, S.; Sidhu, G.P.S. Climate change regulated abiotic stress mechanisms in plants: A comprehensive review. *Plant Cell Rep.* **2022**, *41*, 1–31. [PubMed]
4. Tester, M.; Langridge, P. Breeding technologies to increase crop production in a changing world. *Science* **2010**, *327*, 818–822. [PubMed]
5. Singh, B.D.; Singh, A.K. *Marker-Assisted Plant Breeding: Principles and Practices*; Springer: Berlin/Heidelberg, Germany, 2015.
6. Boopathi, N.M. *Genetic Mapping and Marker Assisted Selection*; Springer: New Delhi, India, 2013; Volume 10, pp. 978–981.
7. Jiang, G.-L. Molecular markers and marker-assisted breeding. In *Plant Breeding from Laboratories to Fields*; IntechOpen: London, UK, 2013; pp. 45–85.
8. Satrio, R.D.; Fendiyanto, M.H.; Miftahudin, M. Tools and techniques used at global scale through genomics, transcriptomics, proteomics, and metabolomics to investigate plant stress responses at the molecular level. In *Molecular Dynamics of Plant Stress and Its Management*; Springer: Berlin/Heidelberg, Germany, 2024; pp. 555–607.
9. Alexandrov, T. Spatial metabolomics and imaging mass spectrometry in the age of artificial intelligence. *Annu. Rev. Biomed. Data Sci.* **2020**, *3*, 61–87.
10. Li, Y.; Wu, F.-X.; Ngom, A. A review on machine learning principles for multi-view biological data integration. *Brief. Bioinform.* **2018**, *19*, 325–340.
11. Hesami, M.; Alizadeh, M.; Jones, A.M.P.; Torkamaneh, D. Machine learning: Its challenges and opportunities in plant system biology. *Appl. Microbiol. Biotechnol.* **2022**, *106*, 3507–3530.
12. Bhuiyan, M.M.R.; Rahaman, M.M.; Aziz, M.M.; Islam, M.R.; Das, K. Predictive analytics in plant biotechnology: Using data science to drive crop resilience and productivity. *J. Environ. Agric. Stud.* **2023**, *4*, 77–83.
13. Fan, B.-L.; Chen, L.-H.; Chen, L.-L.; Guo, H. Integrative Multi-Omics Approaches for Identifying and Characterizing Biological Elements in Crop Traits: Current Progress and Future Prospects. *Int. J. Mol. Sci.* **2025**, *26*, 1466.
14. Granier, C.; Vile, D. Phenotyping and beyond: Modelling the relationships between traits. *Curr. Opin. Plant Biol.* **2014**, *18*, 96–102.
15. Wörheide, M.A.; Krumsiek, J.; Kastenmüller, G.; Arnold, M. Multi-omics integration in biomedical research–A metabolomics-centric review. *Anal. Chim. Acta* **2021**, *1141*, 144–162.
16. Kamali, S.; Singh, A. Genomic and Transcriptomic Approaches to Developing Abiotic Stress-Resilient Crops. *Agronomy* **2023**, *13*, 2903. [CrossRef]
17. Mayr, M.; Mayr, U.; Chung, Y.L.; Yin, X.; Griffiths, J.R.; Xu, Q. Vascular proteomics: Linking proteomic and metabolomic changes. *Proteomics* **2004**, *4*, 3751–3761. [PubMed]
18. Upadhyay, V.R.; Ramesh, V.; Kumar, H.; Somagond, Y.M.; Priyadarsini, S.; Kuniyal, A.; Prakash, V.; Sahoo, A. Phenomics in livestock research: Bottlenecks and promises of digital phenotyping and other quantification techniques on a global scale. *OMICS A J. Integr. Biol.* **2024**, *28*, 380–393.
19. Silaiyiman, S.; Liu, J.; Wu, J.; Ouyang, L.; Cao, Z.; Shen, C. A Systematic Review of the Advances and New Insights into Copy Number Variations in Plant Genomes. *Plants* **2025**, *14*, 1399. [CrossRef]
20. Whibley, A.; Kelley, J.L.; Narum, S.R. *The Changing Face of Genome Assemblies: Guidance on Achieving High-Quality Reference Genomes*; Wiley Online Library: Hoboken, NJ, USA, 2021.
21. Sherman, R.M.; Salzberg, S.L. Pan-genomics in the human genome era. *Nat. Rev. Genet.* **2020**, *21*, 243–254.

22. Amir, R.; Maqsood, W.; Munir, F.; Fatima, N.; Siddiqa, A.; Ahmad, J. Pan-genomics of plant pathogens and its applications. In *Pan-Genomics: Applications, Challenges, and Future Prospects*; Elsevier: Amsterdam, The Netherlands, 2020; pp. 121–145.
23. Sarawad, A.; Hosagoudar, S.; Parvatikar, P. Pan-genomics: Insight into the Functional Genome, Applications, Advancements, and Challenges. *Curr. Genom.* **2025**, *26*, 2–14.
24. Tyagi, A.; Mir, Z.A.; Almalki, M.A.; Deshmukh, R.; Ali, S. Genomics-assisted breeding: A powerful breeding approach for improving plant growth and stress resilience. *Agronomy* **2024**, *14*, 1128. [CrossRef]
25. Pirona, R.; Frugis, G.; Locatelli, F.; Mattana, M.; Genga, A.; Baldoni, E. Transcriptomic analysis reveals the gene regulatory networks involved in leaf and root response to osmotic stress in tomato. *Front. Plant Sci.* **2023**, *14*, 1155797.
26. Javaid, M.H.; Khan, A.R.; Salam, A.; Neelam, A.; Azhar, W.; Ulhassan, Z.; Gan, Y. Exploring the adaptive responses of plants to abiotic stresses using transcriptome data. *Agriculture* **2022**, *12*, 211. [CrossRef]
27. Mastrangelo, A.M.; Marone, D.; Laidò, G.; De Leonardis, A.M.; De Vita, P. Alternative splicing: Enhancing ability to cope with stress via transcriptome plasticity. *Plant Sci.* **2012**, *185*, 40–49.
28. Hassan, M.Q.; Tye, C.E.; Stein, G.S.; Lian, J.B. Non-coding RNAs: Epigenetic regulators of bone development and homeostasis. *Bone* **2015**, *81*, 746–756. [PubMed]
29. Singh, S.; Praveen, A.; Dudha, N.; Sharma, V.K.; Bhadrecha, P. Single-cell transcriptomics: A new frontier in plant biotechnology research. *Plant Cell Rep.* **2024**, *43*, 294. [PubMed]
30. Filichkin, S.; Priest, H.; Givan, S.; Shen, R.; Bryant, D.; Fox, S.; Wong, W.-K.; Mockler, T. Genome-wide mapping of alternative splicing in Arabidopsis thaliana. *Genome Res.* **2010**, *20*, 45–58. [PubMed]
31. Filichkin, S.; Hamilton, M.; Dharmawardhana, P.; Singh, S.K.; Sullivan, C.; Ben-Hur, A.; Reddy, A.; Jaiswal, P. Abiotic Stresses Modulate Landscape of Poplar Transcriptome via Alternative Splicing, Differential Intron Retention, and Isoform Ratio Switching. *Front. Plant Sci.* **2018**, *9*, 5.
32. Li, S.; Yu, X.; Cheng, Z.; Zeng, C.; Li, W.; Zhang, L.; Peng, M. Large-scale analysis of the cassava transcriptome reveals the impact of cold stress on alternative splicing. *J. Exp. Bot.* **2019**, *71*, 422–434.
33. Yang, L.; Yang, L.; Zhao, C.; Liu, J.; Tong, C.; Zhang, Y.; Cheng, X.; Jiang, H.; Shen, J.; Xie, M.; et al. Differential alternative splicing genes and isoform co-expression networks of Brassica napus under multiple abiotic stresses. *Front. Plant Sci.* **2022**, *13*, 1009998.
34. Kim, N.; Lee, J.; Yeom, S.-I.; Kang, N.-J.; Kang, W.-H. The landscape of abiotic and biotic stress-responsive splice variants with deep RNA-seq datasets in hot pepper. *Sci. Data* **2024**, *11*, 3811.
35. Wu, Z.-H.; He, L.-L.; Wang, C.-C.; Liang, C.; Li, H.-Y.; Zhong, D.-W.; Dong, Z.-X.; Zhang, L.-J.; Zhang, X.-Q.; Ge, L.-F.; et al. Unveiling unique alternative splicing responses to low temperature in Zoysia japonica through ZjRTD1.0, a high-quality reference transcript dataset. *Physiol. Plant.* **2024**, *176*, e14280.
36. Wijesinghege, C.; Tran, K.-N.; Dassanayake, M. Alternative splicing preferentially increases transcript diversity associated with stress responses in the extremophyte Schrenkiella parvula. *bioRxiv* **2022**. [CrossRef]
37. Song, L.; Pan, Z.; Chen, L.; Dai, Y.; Wan, J.; Ye, H.; Nguyen, H.; Zhang, G.; Chen, H. Analysis of Whole Transcriptome RNA-seq Data Reveals Many Alternative Splicing Events in Soybean Roots under Drought Stress Conditions. *Genes* **2020**, *11*, 1520. [CrossRef]
38. Xu, L.; Deng, J.; Wang, H.; Zhang, L.; Mi, X.; Luo, L.; Xie, H.; Liu, S.; Wang, S.; Huang, S.; et al. Transcriptome analysis revealed alternative splicing landscape in response to gray-blight disease in tea plant. *Ind. Crops Prod.* **2025**, *225*, 120571.
39. Vogel, C.; Marcotte, E.M. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet.* **2012**, *13*, 227–232. [PubMed]
40. Kosová, K.; Vítámvás, P.; Urban, M.O.; Prášil, I.T.; Renaut, J. Plant abiotic stress proteomics: The major factors determining alterations in cellular proteome. *Front. Plant Sci.* **2018**, *9*, 122.
41. Samaržija, I. Post-translational modifications that drive prostate cancer progression. *Biomolecules* **2021**, *11*, 247. [CrossRef]
42. Lacoursiere, R.E.; Hadi, D.; Shaw, G.S. Acetylation, phosphorylation, ubiquitination (oh my!): Following post-translational modifications on the ubiquitin road. *Biomolecules* **2022**, *12*, 467. [CrossRef]
43. Basar, M.A.; Hosen, M.F.; Paul, B.K.; Hasan, M.R.; Shamim, S.M.; Bhuyian, T. Identification of drug and protein-protein interaction network among stress and depression: A bioinformatics approach. *Inform. Med. Unlocked* **2023**, *37*, 101174.
44. Salam, U.; Ullah, S.; Tang, Z.-H.; Elateeq, A.A.; Khan, Y.; Khan, J.; Khan, A.; Ali, S. Plant metabolomics: An overview of the role of primary and secondary metabolites against different environmental stress factors. *Life* **2023**, *13*, 706. [CrossRef]
45. Chen, Q.; Lu, X.; Guo, X.; Guo, Q.-X.; Li, D. Metabolomics Characterization of Two Apocynaceae Plants, *Catharanthus roseus* and Vinca minor, Using GC-MS and LC-MS Methods in Combination. *Molecules* **2017**, *22*, 997.
46. Dai, H.; Xiao, C.; Liu, H.; Tang, H. Combined NMR and LC-MS analysis reveals the metabonomic changes in Salvia miltiorrhiza Bunge induced by water depletion. *J. Proteome Res.* **2010**, *9*, 1460–1475.
47. Hamade, K.; Fliniaux, O.; Fontaine, J.-X.; Molinié, R.; N'Nang, E.O.; Bassard, S.; Guénin, S.; Gutierrez, L.; Lainé, É.; Hano, C.; et al. NMR and LC-MS-Based Metabolomics to Study Osmotic Stress in Lignan-Deficient Flax. *Molecules* **2021**, *26*, 767.

48. Raletsena, M.; Mdlalose, S.; Bodede, O.; Assress, H.; Woldesemayat, A.; Modise, D. 1H-NMR and LC-MS Based Metabolomics Analysis of Potato (*Solanum tuberosum* L.) Cultivars Irrigated with Fly Ash Treated Acid Mine Drainage. *Molecules* **2022**, *27*, 1187. [PubMed]
49. Jung, Y.; Ha, M.-K.; Lee, J.; Ahn, Y.; Kwak, J.; Ryu, D.; Hwang, G. Metabolite Profiling of the Response of Burdock Roots to Copper Stress. *J. Agric. Food Chem.* **2015**, *63*, 1309–1317. [PubMed]
50. Allwood, J.; Goodacre, R. An introduction to liquid chromatography-mass spectrometry instrumentation applied in plant metabolomic analyses. *Phytochem. Anal. PCA* **2010**, *21*, 33–47. [PubMed]
51. Tzin, V.; Malitsky, S.; Zvi, M.M.B.; Bedair, M.; Sumner, L.; Aharoni, A.; Galili, G. Expression of a bacterial feedback-insensitive 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase of the shikimate pathway in Arabidopsis elucidates potential metabolic bottlenecks between primary and secondary metabolism. *New Phytol.* **2012**, *194*, 430–439.
52. Kaiser, K. Metabolic Profiling of Primary and Secondary Biosynthetic Pathways in Angiosperms: Comparative Metabonomics and Applications of Hyphenated LC-NMR and LC-MS. Ph.D. Thesis, University of California, San Francisco, CA, USA, 2012.
53. Angidi, S.; Madankar, K.; Tehseen, M.M.; Bhatla, A. Advanced High-Throughput Phenotyping Techniques for Managing Abiotic Stress in Agricultural Crops—A Comprehensive Review. *Crops* **2025**, *5*, 8.
54. Zhang, Y.; Zhang, N. Imaging technologies for plant high-throughput phenotyping: A review. *Front. Agric. Sci. Eng.* **2018**, *5*, 406–419.
55. Li, D.; Quan, C.; Song, Z.; Li, X.; Yu, G.; Li, C.; Muhammad, A. High-throughput plant phenotyping platform (HT3P) as a novel tool for estimating agronomic traits from the lab to the field. *Front. Bioeng. Biotechnol.* **2021**, *8*, 623705.
56. Srivastava, U.; Kanchan, S.; Kesheri, M.; Gupta, M.K.; Singh, S. Types of omics data: Genomics, metagenomics, epigenomics, transcriptomics, proteomics, metabolomics, and phenomics. In *Integrative Omics*; Elsevier: Amsterdam, The Netherlands, 2024; pp. 13–34.
57. Knox, S.S. From'omics' to complex disease: A systems biology approach to gene-environment interactions in cancer. *Cancer Cell Int.* **2010**, *10*, 11.
58. Ritchie, M.D.; Holzinger, E.R.; Li, R.; Pendergrass, S.A.; Kim, D. Methods of integrating data to uncover genotype–phenotype interactions. *Nat. Rev. Genet.* **2015**, *16*, 85–97.
59. Ma, X.; Xia, H.; Liu, Y.; Wei, H.; Zheng, X.; Song, C.; Chen, L.; Liu, H.; Luo, L. Transcriptomic and metabolomic studies disclose key metabolism pathways contributing to well-maintained photosynthesis under the drought and the consequent drought-tolerance in rice. *Front. Plant Sci.* **2016**, *7*, 1886.
60. Vahabi, N.; Michailidis, G. Unsupervised multi-omics data integration methods: A comprehensive review. *Front. Genet.* **2022**, *13*, 854752.
61. Ulfenborg, B. Vertical and horizontal integration of multi-omics data with miodin. *BMC Bioinform.* **2019**, *20*, 649.
62. Wang, B.; Liu, C.; Zhang, D.; He, C.; Zhang, J.; Li, Z. Effects of maize organ-specific drought stress response on yields from transcriptome analysis. *BMC Plant Biol.* **2019**, *19*, 335.
63. Santiago-Rodriguez, T.M.; Hollister, E.B. Multi 'omic data integration: A review of concepts, considerations, and approaches. *Semin. Perinatol.* **2021**, *45*, 151456. [PubMed]
64. Ullah, M.A.; Abdullah-Zawawi, M.-R.; Zainal-Abidin, R.-A.; Sukiran, N.L.; Uddin, M.I.; Zainal, Z. A review of integrative omic approaches for understanding rice salt response mechanisms. *Plants* **2022**, *11*, 1430. [CrossRef]
65. Akdemir, D.; Knox, R.; Isidro y Sánchez, J. Combining partially overlapping multi-omics data in databases using relationship matrices. *Front. Plant Sci.* **2020**, *11*, 947.
66. Hecker, M.; Lambeck, S.; Toepfer, S.; Van Someren, E.; Guthke, R. Gene regulatory network inference: Data integration in dynamic models—A review. *Biosystems* **2009**, *96*, 86–103.
67. Lv, L.; Zhang, W.; Sun, L.; Zhao, A.; Zhang, Y.; Wang, L.; Liu, Y.; Li, Z.; Li, H.; Chen, X. Gene co-expression network analysis to identify critical modules and candidate genes of drought-resistance in wheat. *PLoS ONE* **2020**, *15*, e0236186.
68. Lee, B.; Zhang, S.; Poleksic, A.; Xie, L. Heterogeneous multi-layered network model for omics data integration and analysis. *Front. Genet.* **2020**, *10*, 1381.
69. Misra, B.B.; Langefeld, C.; Olivier, M.; Cox, L.A. Integrated omics: Tools, advances and future approaches. *J. Mol. Endocrinol.* **2019**, *62*, R21–R45.
70. Rohart, F.; Gautier, B.; Singh, A.; Lê Cao, K.-A. mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Comput. Biol.* **2017**, *13*, e1005752.
71. John Martin, J.J.; Song, Y.; Hou, M.; Zhou, L.; Liu, X.; Li, X.; Fu, D.; Li, Q.; Cao, H.; Li, R. Multi-Omics Approaches in Oil Palm Research: A Comprehensive Review of Metabolomics, Proteomics, and Transcriptomics Based on Low-Temperature Stress. *Int. J. Mol. Sci.* **2024**, *25*, 7695. [PubMed]
72. Calado, R.A. Statistical Methods for the Integrative Analysis of Single-Cell Multi-Omics Data. Ph.D. Thesis, University of Cambridge, Cambridge, UK, 2020.

73. Xue, Z. Integration of High-Throughput Phenotyping and Genomics Data to Explore Arabidopsis Natural Variation. Ph.D. Thesis, Université Paris-Saclay, Paris, France, 2020.
74. Roychowdhury, R.; Das, S.P.; Gupta, A.; Parihar, P.; Chandrasekhar, K.; Sarker, U.; Kumar, A.; Ramrao, D.P.; Sudhakar, C. Multi-omics pipeline and omics-integration approach to decipher plant's abiotic stress tolerance responses. *Genes* **2023**, *14*, 1281. [PubMed]
75. Pei, G.; Chen, L.; Zhang, W. WGCNA application to proteomic and metabolomic data analysis. In *Methods in Enzymology*; Elsevier: Amsterdam, The Netherlands, 2017; Volume 585, pp. 135–158.
76. Singh, A.; Gautier, B.; Shannon, C.P.; Vacher, M.; Rohart, F.; Tebbutt, S.J.; Cao, K.-A.L. DIABLO–an integrative, multi-omics, multivariate method for multi-group classification. *BioRxiv* **2016**. [CrossRef]
77. Chu, Y.; Kaushik, A.C.; Wang, X.; Wang, W.; Zhang, Y.; Shan, X.; Salahub, D.R.; Xiong, Y.; Wei, D.-Q. DTI-CDF: A cascade deep forest model towards the prediction of drug-target interactions based on hybrid features. *Brief. Bioinform.* **2019**, *22*, 451–462.
78. Cuba Samaniego, C.; Franco, E. Ultrasensitive molecular controllers for quasi-integral feedback. *Cell Syst.* **2021**, *12*, 272–288.e273.
79. Koh, H.W.L.; Fermin, D.; Vogel, C.; Choi, K.P.; Ewing, R.M.; Choi, H. iOmicsPASS: Network-based integration of multiomics data for predictive subnetwork discovery. *NPJ Syst. Biol. Appl.* **2019**, *5*, 22.
80. Langfelder, P.; Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinform.* **2008**, *9*, 559.
81. Argelaguet, R.; Velten, B.; Arnol, D.; Dietrich, S.; Zenz, T.; Marioni, J.C.; Buettner, F.; Huber, W.; Stegle, O. Multi-Omics Factor Analysis—A framework for unsupervised integration of multi-omics data sets. *Mol. Syst. Biol.* **2018**, *14*, e8124.
82. Guo, F.; Wang, D.; Wang, L. Progressive approach for SNP calling and haplotype assembly using single molecular sequencing data. *Bioinformatics* **2018**, *34*, 2012–2018.
83. Hill, J.; Su, Y.-S. Assessing lack of common support in causal inference using Bayesian nonparametrics: Implications for evaluating the effect of breastfeeding on children's cognitive outcomes. *Ann. Appl. Stat.* **2013**, *7*, 1386–1420.
84. Morris, J.S.; Baladandayuthapani, V. Statistical contributions to bioinformatics: Design, modelling, structure learning and integration. *Stat. Model.* **2017**, *17*, 245–289.
85. Huang, Z.; Wang, C. A review on differential abundance analysis methods for mass spectrometry-based metabolomic data. *Metabolites* **2022**, *12*, 305. [CrossRef] [PubMed]
86. Tyler, A.L.; Asselbergs, F.W.; Williams, S.M.; Moore, J.H. Shadows of complexity: What biological networks reveal about epistasis and pleiotropy. *Bioessays* **2009**, *31*, 220–227. [PubMed]
87. Fenu, G.; Malloci, F.M. Evaluating impacts between laboratory and field-collected datasets for plant disease classification. *Agronomy* **2022**, *12*, 2359. [CrossRef]
88. Kim, M.; Tagkopoulos, I. Data integration and predictive modeling methods for multi-omics datasets. *Mol. Omics* **2018**, *14*, 8–25.
89. Mahmood, U.; Li, X.; Fan, Y.; Chang, W.; Niu, Y.; Li, J.; Qu, C.; Lu, K. Multi-omics revolution to promote plant breeding efficiency. *Front. Plant Sci.* **2022**, *13*, 1062952.
90. Cembrowska-Lech, D.; Krzemińska, A.; Miller, T.; Nowakowska, A.; Adamski, C.; Radaczyńska, M.; Mikiciuk, G.; Mikiciuk, M. An integrated multi-omics and artificial intelligence framework for advance plant phenotyping in horticulture. *Biology* **2023**, *12*, 1298. [CrossRef]
91. Zafar, I.; Anwar, S.; Yousaf, W.; Nisa, F.U.; Kausar, T.; ul Ain, Q.; Unar, A.; Kamal, M.A.; Rashid, S.; Khan, K.A. Reviewing methods of deep learning for intelligent healthcare systems in genomics and biomedicine. *Biomed. Signal Process. Control.* **2023**, *86*, 105263.
92. Hayes, C.N.; Nakahara, H.; Ono, A.; Tsuge, M.; Oka, S. From Omics to Multi-Omics: A Review of Advantages and Tradeoffs. *Genes* **2024**, *15*, 1551. [CrossRef] [PubMed]
93. Bustos-Korts, D.; Malosetti, M.; Chapman, S.; van Eeuwijk, F. Modelling of genotype by environment interaction and prediction of complex traits across multiple environments as a synthesis of crop growth modelling, genetics and statistics. In *Crop Systems Biology: Narrowing the Gaps Between Crop Modelling and Genetics*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 55–82.
94. Sánchez, J.C.M.; Mesa, H.G.A.; Espinosa, A.T.; Castilla, S.R.; Lamont, F.G. Improving wheat yield prediction through variable selection using Support Vector Regression, Random Forest, and Extreme Gradient Boosting. *Smart Agric. Technol.* **2025**, *10*, 100791.
95. Tembhare, K.; Sharma, T.; Kasibhatla, S.M.; Achalere, A.; Joshi, R. Multi-ensemble machine learning framework for omics data integration: A case study using breast cancer samples. *Inform. Med. Unlocked* **2024**, *47*, 101507.
96. Wu, C.; Luo, J.; Xiao, Y. Multi-omics assists genomic prediction of maize yield with machine learning approaches. *Mol. Breed.* **2024**, *44*, 14. [PubMed]
97. Xu, Y.; Yang, W.; Qiu, J.; Zhou, K.; Yu, G.; Zhang, Y.; Wang, X.; Jiao, Y.; Wang, X.; Hu, S. Metabolic marker-assisted genomic prediction improves hybrid breeding. *Plant Commun.* **2025**, *6*, 101199. [PubMed]
98. Liu, Q.; Zuo, S.-m.; Peng, S.; Zhang, H.; Peng, Y.; Li, W.; Xiong, Y.; Lin, R.; Feng, Z.; Li, H. Development of machine learning methods for accurate prediction of plant disease resistance. *Engineering* **2024**, *40*, 100–110.

99. Pimpalkar, A.; Gandhewar, N.; Shelke, N.; Patil, S.; Chhabria, S. An Efficient Deep Convolutional Neural Networks Model for Genomic Sequence Classification. In *Genomics at the Nexus of AI, Computer Vision, and Machine Learning*; Wiely: Hoboken, NJ, USA, 2025; pp. 345–375.
100. Liang, H.; Zhou, Y.; Lu, Y.; Pei, S.; Xu, D.; Lu, Z.; Yao, W.; Liu, Q.; Yu, L.; Li, H. Evaluation of Soybean Drought Tolerance Using Multimodal Data from an Unmanned Aerial Vehicle and Machine Learning. *Remote Sens.* **2024**, *16*, 2043.
101. Lin, C. Analysis of Complex Dynamical Systems by Combining Recurrent Neural Networks and Mechanistic Models. Ph.D. Thesis, University of Ottawa, Ottawa, Canada, 2024.
102. Haji, A. Comparative Analysis of Autoencoder and PCA for Dimensionality Reduction in Gene Expression Data. 2024. Available online: https://his.diva-portal.org/smash/get/diva2:1883117/FULLTEXT02.pdf (accessed on 3 July 2025).
103. Montesinos-López, O.A.; Chavira-Flores, M.; Kismiantini; Crespo-Herrera, L.; Saint Piere, C.; Li, H.; Fritsche-Neto, R.; Al-Nowibet, K.; Montesinos-López, A.; Crossa, J. A review of multimodal deep learning methods for genomic-enabled prediction in plant breeding. *Genetics* **2024**, *228*, iyae161.
104. Wang, X.; Zeng, H.; Lin, L.; Huang, Y.; Lin, H.; Que, Y. Deep learning-empowered crop breeding: Intelligent, efficient and promising. *Front. Plant Sci.* **2023**, *14*, 1260089.
105. Nicolas, J. Artificial intelligence and bioinformatics. In *A Guided Tour of Artificial Intelligence Research: Volume III: Interfaces and Applications of Artificial Intelligence*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 209–264.
106. Lac, L.; Leung, C.K.; Hu, P. Computational frameworks integrating deep learning and statistical models in mining multimodal omics data. *J. Biomed. Inform.* **2024**, *152*, 104629.
107. Yue, T.; Wang, Y.; Zhang, L.; Gu, C.; Xue, H.; Wang, W.; Lyu, Q.; Dun, Y. Deep Learning for Genomics: From Early Neural Nets to Modern Large Language Models. *Int. J. Mol. Sci.* **2023**, *24*, 15858.
108. Chen, H.; Lu, D.; Xiao, Z.; Li, S.; Zhang, W.; Luan, X.; Zhang, W.; Zheng, G. Comprehensive applications of the artificial intelligence technology in new drug research and development. *Health Inf. Sci. Syst.* **2024**, *12*, 41. [PubMed]
109. Desta, Z.A.; Ortiz, R. Genomic selection: Genome-wide prediction in plant improvement. *Trends Plant Sci.* **2014**, *19*, 592–601. [PubMed]
110. Vogel, J.T.; Liu, W.; Olhoft, P.; Crafts-Brandner, S.J.; Pennycooke, J.C.; Christiansen, N. Soybean yield formation physiology—A foundation for precision breeding based improvement. *Front. Plant Sci.* **2021**, *12*, 719706.
111. Du, H.; Zhu, J.; Su, H.; Huang, M.; Wang, H.; Ding, S.; Zhang, B.; Luo, A.; Wei, S.; Tian, X. Bulked segregant RNA-seq reveals differential expression and SNPs of candidate genes associated with waterlogging tolerance in maize. *Front. Plant Sci.* **2017**, *8*, 1022.
112. Sandhu, K.S.; Mihalyov, P.D.; Lewien, M.J.; Pumphrey, M.O.; Carter, A.H. Combining genomic and phenomic information for predicting grain protein content and grain yield in spring wheat. *Front. Plant Sci.* **2021**, *12*, 613300.
113. Robert, P.; Goudemand, E.; Auzanneau, J.; Oury, F.-X.; Rolland, B.; Heumez, E.; Bouchet, S.; Caillebotte, A.; Mary-Huard, T.; Le Gouis, J. Phenomic selection in wheat breeding: Prediction of the genotype-by-environment interaction in multi-environment breeding trials. *Theor. Appl. Genet.* **2022**, *135*, 3337–3356.
114. Chawade, A.; Van Ham, J.; Blomquist, H.; Bagge, O.; Alexandersson, E.; Ortiz, R. High-throughput field-phenotyping tools for plant breeding and precision agriculture. *Agronomy* **2019**, *9*, 258.
115. Mansoor, S.; Chung, Y.S. Functional phenotyping: Understanding the dynamic response of plants to drought stress. *Curr. Plant Biol.* **2024**, *38*, 100331.
116. Xu, Y.; Zhang, X.; Li, H.; Zheng, H.; Zhang, J.; Olsen, M.S.; Varshney, R.K.; Prasanna, B.M.; Qian, Q. Smart breeding driven by big data, artificial intelligence, and integrated genomic-enviromic prediction. *Mol. Plant* **2022**, *15*, 1664–1695.
117. Gano, B.; Bhadra, S.; Vilbig, J.M.; Ahmed, N.; Sagan, V.; Shakoor, N. Drone-based imaging sensors, techniques, and applications in plant phenotyping for crop breeding: A comprehensive review. *Plant Phenome J.* **2024**, *7*, e20100.
118. Awais, M.; Li, W.; Cheema, M.J.M.; Zaman, Q.U.; Shaheen, A.; Aslam, B.; Zhu, W.; Ajmal, M.; Faheem, M.; Hussain, S. UAV-based remote sensing in plant stress imagine using high-resolution thermal sensor for digital agriculture practices: A meta-review. *Int. J. Environ. Sci. Technol.* **2022**, *20*, 1135–1152.
119. Messina, G.; Modica, G. Applications of UAV thermal imagery in precision agriculture: State of the art and future research outlook. *Remote Sens.* **2020**, *12*, 1491.
120. Ashfaq, W.; Brodie, G.; Fuentes, S.; Gupta, D. Infrared thermal imaging and morpho-physiological indices used for wheat genotypes screening under drought and heat stress. *Plants* **2022**, *11*, 3269. [CrossRef] [PubMed]
121. Wang, N. The Potential of UAV-Based Sun-Induced Chlorophyll Fluorescence in Understanding Crop Photosynthesis. Ph.D. Thesis, Wageningen University and Research, Wageningen, The Netherlands, 2022.
122. Driever, S.M.; Mossink, L.; Ocaña, D.N.; Kaiser, E. A simple system for phenotyping of plant transpiration and stomatal conductance response to drought. *Plant Sci.* **2023**, *329*, 111626. [PubMed]
123. McGrail, R.K.; Van Sanford, D.A.; McNear, D.H., Jr. Trait-based root phenotyping as a necessary tool for crop selection and improvement. *Agronomy* **2020**, *10*, 1328. [CrossRef]

124. Ludovisi, R.; Tauro, F.; Salvati, R.; Khoury, S.; Mugnozza Scarascia, G.; Harfouche, A. UAV-Based Thermal Imaging for High-Throughput Field Phenotyping of Black Poplar Response to Drought. *Front. Plant Sci.* **2017**, *8*, 1681.
125. Zhang, H.; Wang, L.; Jin, X.; Bian, L.; Ge, Y. High-throughput phenotyping of plant leaf morphological, physiological, and biochemical traits on multiple scales using optical sensing. *Crop J.* **2023**, *11*, 1303–1318.
126. Gill, T.; Gill, S.K.; Saini, D.K.; Chopra, Y.; de Koff, J.P.; Sandhu, K.S. A comprehensive review of high throughput phenotyping and machine learning for plant stress phenotyping. *Phenomics* **2022**, *2*, 156–183.
127. Wen, T.; Li, J.-H.; Wang, Q.; Gao, Y.-Y.; Hao, G.-F.; Song, B.-A. Thermal imaging: The digital eye facilitates high-throughput phenotyping traits of plant growth and stress responses. *Sci. Total Environ.* **2023**, *899*, 165626.
128. Jin, X.; Zarco-Tejada, P.J.; Schmidhalter, U.; Reynolds, M.P.; Hawkesford, M.J.; Varshney, R.K.; Yang, T.; Nie, C.; Li, Z.; Ming, B.; et al. High-Throughput Estimation of Crop Traits: A Review of Ground and Aerial Phenotyping Platforms. *IEEE Geosci. Remote Sens. Mag.* **2021**, *9*, 200–231.
129. Li, L.; Zhang, Q.; Huang, D. A Review of Imaging Techniques for Plant Phenotyping. *Sensors* **2014**, *14*, 20078–20111. [CrossRef] [PubMed]
130. Zhou, J.; Vong, C.N.; Zhou, J. Imaging Technology for High-Throughput Plant Phenotyping. In *Sensing, Data Managing, and Control Technologies for Agricultural Systems*; Ma, S., Lin, T., Mao, E., Song, Z., Ting, K.-C., Eds.; Springer International Publishing: Cham, Switzerland, 2022; pp. 75–99.
131. Thorp, K.R.; Thompson, A.L.; Harders, S.J.; French, A.N.; Ward, R.W. High-Throughput Phenotyping of Crop Water Use Efficiency via Multispectral Drone Imagery and a Daily Soil Water Balance Model. *Remote Sens.* **2018**, *10*, 1682.
132. Mertens, S.; Verbraeken, L.; Sprenger, H.; De Meyer, S.; Demuynck, K.; Cannoot, B.; Merchie, J.; De Block, J.; Vogel, J.T.; Bruce, W.; et al. Monitoring of drought stress and transpiration rate using proximal thermal and hyperspectral imaging in an indoor automated plant phenotyping platform. *Plant Methods* **2023**, *19*, 132. [PubMed]
133. Des Marais, D.L.; Hernandez, K.M.; Juenger, T.E. Genotype-by-environment interaction and plasticity: Exploring genomic responses of plants to the abiotic environment. *Annu. Rev. Ecol. Evol. Syst.* **2013**, *44*, 5–29.
134. Negash, A.W.; Mwambi, H.; Zewotir, T.; Taye, G. Additive main effects and multiplicative interactions model (AMMI) and genotype main effect and genotype by environment interaction (GGE) biplot analysis of multi-environmental wheat variety trials. *Afr. J. Agric. Res.* **2013**, *8*, 1033–1040.
135. Achenef, G. Advancement of Analytical models quantifying G× E interactions and Stability analysis in Multi-environment Trial. *Int. J. Res. Agric. Sci.* **2022**, *9*, 103–120.
136. Tesfaye, K.; Sonder, K.; Cairns, J.; Magorokosho, C.; Tarekegn, A.; Kassie, G.T.; Getaneh, F.; Abdoulaye, T.; Abate, T.; Erenstein, O. Targeting drought-tolerant maize varieties in southern Africa: A geospatial crop modeling approach using big data. *Int. Food Agribus. Manag. Rev.* **2016**, *19*, 75–92.
137. Yan, W.; Tinker, N.A. An integrated biplot analysis system for displaying, interpreting, and exploring genotype× environment interaction. *Crop Sci.* **2005**, *45*, 1004–1016.
138. Frutos, E.; Galindo, M.P.; Leiva, V. An interactive biplot implementation in R for modeling genotype-by-environment interaction. *Stoch. Environ. Res. Risk Assess.* **2014**, *28*, 1629–1641.
139. Al-Ashkar, I.; Sallam, M.; Almutairi, K.F.; Shady, M.; Ibrahim, A.; Alghamdi, S.S. Detection of high-performance wheat genotypes and genetic stability to determine complex interplay between genotypes and environments. *Agronomy* **2023**, *13*, 585. [CrossRef]
140. Blum, A.; Blum, A. Drought resistance and its improvement. In *Plant Breeding for Water-Limited Environments*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 53–152.
141. Rakshit, S.; Ganapathy, K.N.; Gomashe, S.S.; Rathore, A.; Ghorade, R.B.; Kumar, M.V.N.; Ganesmurthy, K.; Jain, S.K.; Kamtar, M.Y.; Sachan, J.S. GGE biplot analysis to evaluate genotype, environment and their interactions in sorghum multi-location data. *Euphytica* **2012**, *185*, 465–479.
142. Gonçalves, G.d.M.C.; Ferreira-Gomes, R.L.; Lopes, Â.C.d.A.; Vieira, P.F.d.M.J. Adaptability and yield stability of soybean genotypes by REML/BLUP and GGE Biplot. *Crop Breed. Appl. Biotechnol.* **2020**, *20*, e282920217.
143. Weraikat, D.; Šorič, K.; Žagar, M.; Sokač, M. Data Analytics in Agriculture: Enhancing Decision-Making for Crop Yield Optimization and Sustainable Practices. *Sustainability* **2024**, *16*, 7331. [CrossRef]
144. Ali, I.; Greifeneder, F.; Stamenkovic, J.; Neumann, M.; Notarnicola, C. Review of machine learning approaches for biomass and soil moisture retrievals from remote sensing data. *Remote Sens.* **2015**, *7*, 16398–16421.
145. Levizzani, V.; Cattani, E. Satellite remote sensing of precipitation and the terrestrial water cycle in a changing climate. *Remote Sens.* **2019**, *11*, 2301.
146. Singh, S.; Praveen, A.; Dudha, N.; Bhadrecha, P. Integrating physiological and multi-omics methods to elucidate heat stress tolerance for sustainable rice production. *Physiol. Mol. Biol. Plants* **2024**, *30*, 1185–1208.
147. Deng, C.H.; Naithani, S.; Kumari, S.; Cobo-Simón, I.; Quezada-Rodríguez, E.H.; Skrabisova, M.; Gladman, N.; Correll, M.J.; Sikiru, A.B.; Afuwape, O.O. Genotype and phenotype data standardization, utilization and integration in the big data era for agricultural sciences. *Database* **2023**, *2023*, baad088.

148. Mohr, A.E.; Ortega-Santos, C.P.; Whisner, C.M.; Klein-Seetharaman, J.; Jasbi, P. Navigating challenges and opportunities in multi-omics integration for personalized healthcare. *Biomedicines* **2024**, *12*, 1496. [CrossRef]
149. Mukherjee, A.; Abraham, S.; Singh, A.; Balaji, S.; Mukunthan, K.S. From data to cure: A comprehensive exploration of multi-omics data analysis for targeted therapies. *Mol. Biotechnol.* **2025**, *67*, 1269–1289. [PubMed]
150. Mathur, P. Cloud computing infrastructure, platforms, and software for scientific research. In *High Performance Computing in Biomimetics: Modeling, Architecture and Applications*; Springer: Berlin/Heidelberg, Germany, 2024; pp. 89–127.
151. Huang, Y.-N.; Munteanu, V.; Love, M.I.; Ronkowski, C.F.; Deshpande, D.; Wong-Beringer, A.; Corbett-Detig, R.; Dimian, M.; Moore, J.H.; Garmire, L.X. Perceptual and technical barriers in sharing and formatting metadata accompanying omics studies. *Cell Genom.* **2025**, *5*, 100845. [PubMed]
152. Adamer, M.F.; Brüningk, S.; Tejada-Arranz, A.; Estermann, F.; Basler, M.; Borgwardt, K.M. reComBat: Batch effect removal in large-scale, multi-source omics data integration. *bioRxiv* **2021**. [CrossRef]
153. Das, S.; Mukhopadhyay, I. TiMEG: An integrative statistical method for partially missing multi-omics data. *Sci. Rep.* **2021**, *11*, 24077.
154. Voß, H.; Schlumbohm, S.; Wurlitzer, M.; Dottermusch, M.; Neumann, P.; Barwikowski, P.; Schlüter, H.; Krisp, C.; Neumann, J. OTHR-07. A new framework for missing value tolerant data integration. *Neuro-Oncology* **2022**, *24*, i148.
155. Mertens, B. Transformation, Normalization, and Batch Effect in the Analysis of Mass Spectrometry Data for Omics Studies. *arXiv* **2016**, arXiv:1606.05360.
156. Yan, X.; Ang, K.S.; Van Olst, L.; Edwards, A.; Watson, T.; Zheng, R.; Fan, R.; Li, M.; Gate, D.; Chen, J. Mosaic integration of spatial multi-omics with SpaMosaic. *bioRxiv* **2024**. [CrossRef]
157. Baião, A.; Cai, Z.; Poulos, R.; Robinson, P.; Reddel, R.; Zhong, Q.; Vinga, S.; Gonccalves, E. A technical review of multi-omics data integration methods: From classical statistical to deep generative approaches. *arXiv* **2025**, arXiv:2501.17729.
158. Flores, J.; Claborne, D.; Weller, Z.; Webb-Robertson, B.; Waters, K.; Bramer, L. Missing data in multi-omics integration: Recent advances through artificial intelligence. *Front. Artif. Intell.* **2023**, *6*, 1098308.
159. Ugidos, M.; Tarazona, S.; Prats-Montalbán, J.; Ferrer, A.; Conesa, A. MultiBaC: A strategy to remove batch effects between different omic data types. *Stat. Methods Med. Res.* **2020**, *29*, 2851–2864.
160. Hui, H.; Goh, W. Uncovering the consequences of batch effect associated missing values in omics data analysis. *bioRxiv* **2023**. [CrossRef]
161. Wang, D.; Bodovitz, S. Single cell analysis: The new frontier in 'omics'. *Trends Biotechnol.* **2010**, *28*, 281–290. [PubMed]
162. Doust, A.N.; Lukens, L.; Olsen, K.M.; Mauro-Herrera, M.; Meyer, A.; Rogers, K. Beyond the single gene: How epistasis and gene-by-environment effects influence crop domestication. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 6178–6183. [PubMed]
163. Eichler, E.E.; Flint, J.; Gibson, G.; Kong, A.; Leal, S.M.; Moore, J.H.; Nadeau, J.H. Missing heritability and strategies for finding the underlying causes of complex disease. *Nat. Rev. Genet.* **2010**, *11*, 446–450. [PubMed]
164. Mata, J.; Marguerat, S.; Bähler, J. Post-transcriptional control of gene expression: A genome-wide perspective. *Trends Biochem. Sci.* **2005**, *30*, 506–514. [PubMed]
165. Joshi, A.; Rienks, M.; Theofilatos, K.; Mayr, M. Systems biology in cardiovascular disease: A multiomics approach. *Nat. Rev. Cardiol.* **2021**, *18*, 313–330. [PubMed]
166. Yurekten, O.; Payne, T.; Tejera, N.; Amaladoss, F.X.; Martin, C.; Williams, M.; O'Donovan, C. MetaboLights: Open data repository for metabolomics. *Nucleic Acids Res.* **2024**, *52*, D640–D646.
167. Ugochukwu, A.I.; Phillips, P.W.B. Open data ownership and sharing: Challenges and opportunities for application of FAIR principles and a checklist for data managers. *J. Agric. Food Res.* **2024**, *16*, 101157.
168. Buck, M.; Hamilton, C. The Nagoya Protocol on access to genetic resources and the fair and equitable sharing of benefits arising from their utilization to the Convention on Biological Diversity. *Rev. Eur. Community Int. Environ. Law* **2011**, *20*, 47–61.
169. Aydin, S.; Aydin, M.N. Semantic and syntactic interoperability for agricultural open-data platforms in the context of IoT using crop-specific trait ontologies. *Appl. Sci.* **2020**, *10*, 4460.
170. Adepoju, A.G.; Adepoju, D.A. Biomarker Discovery in Clinical Biology Enhances Early Disease Detection, Prognosis, and Personalized Treatment Strategies. *Int. J. Adv. Res. Publ. Rev.* **2025**, *2*, 229–252.
171. Rossi, B.; Russo, B.; Succi, G. Adoption of free/libre open source software in public organizations: Factors of impact. *Inf. Technol. People* **2012**, *25*, 156–187.
172. Committee on Key Challenge Areas for Convergence and Health; Board on Life Sciences; Division on Earth and Life Studies; National Research Council. *Convergence: Facilitating Transdisciplinary Integration of Life Sciences, Physical Sciences, Engineering, and Beyond*; National Academies Press (US): Washington, DC, USA, 2014.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.