GMR
The Original

DEVELOPMENT OF IMPROVED SEED VARIETIES USING GENOMICS, BIOINFORMATICS AND HIGH-THROUGHPUT PHENOTYPING

Pooja S.K [1*], Thota Joseph Raju [2*], Mohan Chavan[3], Pramila C.K[4], Umesh Hiremath[5], Geetha Govind [6], Sneha N[7], Sri Vidya M[8], T John Rathnakar[9], Dharini[10], Shaik Muskan Fathima[11], Manoj D[12], Nalini BS[13], Manu N[14], G. Avinash [15]

[1] Assistant Professor, Department of Agricultural Biotechnology, UAS, GKVK, Bangalore, Karnataka, India
[2] Senior Research Fellow, Advance Center on Bio energy Research, UAS, Bangalore, Karnataka, India
[3] Professor, Department of Plant Biotechnology, UAS, GKVK, Bangalore, Karnataka, India
[4]Assistant Professor, Agricultural Technology Information, Center, UAS, Bangalore, Karnataka, India
[5] Assistant Professor, Department of Seed Science and Technology, UAS, Raichur, Karnataka, India
[6] Assistant Professor, Department of Crop Physiology / Biotechnology, UAS, Mandya, Karnataka, India
[7] Assistant Professor, School of CSA, Reva University, Karnataka, India
[8] Research Scholar, Chronobiology Lab, DOS Zoology, University of Mysore, Karnataka, India
[9] Research Associate (Bio-chemistry), Dr. Y.S.R Horticultural University, Venkataramanappa, AP, India
[10] M.Sc. Scholar, Department of Seed Science and Technology, UAS, Raichur, Karnataka, India
[11] M.Sc. Scholar, Department of Plant Biotechnology, UAS, GKVK, Bangalore, Karnataka, India
[12] Research Associate, Department of Plant Breeding, ARS, Ragolu, ANGRAU, AP, India
[13] Consultant, Agri Export Cell, Karnataka State Agricultural Produce Processing and Export Corporation Limited, Karnataka, India
[14] M.Sc. Scholar, Department of Plant Biotechnology, UAS, GKVK, Bangalore, Karnataka, India
[15] Senior Executive (CEO), Avyagraha Research and Analytics, Kampli, India, India
*Corresponding author: poojaskbiotech17@gmail.com ; josephrajuthota@gmail.com;

ABSTRACT
The growing worldwide need for food security due to climate change, population expansion, and less land that can be farmed means that crop development tactics need to evolve from the Green Revolution to precise genetic changes made by genome editing. The period of enhanced seed varieties investigates the paradigm change characterizing the evolution from traditional agriculture to contemporary precision agriculture, driven by the synergistic integration of genomics, bioinformatics and high-throughput phenotyping (HTP). We stress how new genomic techniques like genomic selection (GS) and next-generation sequencing (NGS) have changed the way we find the best genetic markers, which has sped up the creation of strong agricultural varieties. Revolutionary genome editing technologies like CRISPR/Cas9 may now make specific changes that make plants more resistant to biotic and abiotic stress while using fewer chemicals. Concurrently, bioinformatics platforms serve as the critical computational infrastructure for managing massive, multidimensional "omics" datasets, enabling complex genome-wide association studies (GWAS), synteny analysis and predictive modeling. Furthermore, we examine state-of-the-art HTP technologies—spanning micro-computed tomography (micro-CT) for assessing invisible internal seed characteristics to unmanned aerial vehicles (UAVs) for large-scale field evaluations. By leveraging machine learning (ML) and deep learning algorithms, HTP effectively bridges the genotype-phenotype gap, allowing for rapid and non-destructive trait evaluation across diverse developmental stages. The important role of HTP in making it easier to measure a lot of different plant traits accurately and quickly, from yield to resistance to biotic and abiotic stress, shows how genomics, phenomics and advanced computing techniques work together to make things better. Despite these technological leaps, the translation of genomic innovations from laboratory settings to broad field applications remains constrained by complex genotype-environment (G×E) interactions, computational bottlenecks, stringent regulatory frameworks and public acceptance. Ultimately, this synthesis emphasizes that the combination of multi-omics, advanced data analytics, and precision phenotyping is essential for sustainable crop improvement. Overcoming current translational hurdles will require sustained multidisciplinary collaboration to deliver climate-smart agricultural solutions, ensuring resilient global food systems for the future.

KEYWORDS: Genotyping, Genomics, Bioinformatics, High Throughput Phenotyping, Sustainable Agriculture

1. INTRODUCTION

Rising demands for food security, climate resilience, and sustainable agriculture, driven by population growth and limited arable land, are transforming the search for improved seed varieties. Plant genomics, innovative phenotyping, and bioinformatics have revolutionized conventional breeding, enabling more productive and stress-

resistant crops and accelerating crop enhancement. Genomic research reveals the genetics of plant traits, allowing researchers to use vast crop DNA sequences to identify genes linked to biotic and abiotic stress tolerance and yield potential. Next-generation sequencing (NGS) has made crop genome sequencing fast and inexpensive. Genomic selection (GS) predicts the performance of untested genotypes using genetic markers and statistical models, helping breeders make informed decisions early and increasing the creation of novel types. Research shows that GS improves yield and disease resistance more than conventional breeding (Heffner et al., 2010). Genomic tools also give breeders access to genetic diversity, which is needed to grow crops that survive in varied environments and resist new pests and diseases. Exploiting diverse germplasm collections has helped identify novel beneficial alleles.

Bioinformatics, combining biology, computer science, and mathematics, analyzes and interprets the large datasets produced by high-throughput sequencing. Comprehensive genome sequences, along with their management and interpretation, are essential for capturing the genetic diversity of entire crop gene pools and identifying useful variants. By integrating genetic data with phenotypic data and genotype-environment interactions, bioinformatics helps researchers discover candidate genes, find correlations between markers and phenotypic expression, and understand plant responses to different environments. It also helps curate and visualize large datasets for valuable insights, supporting the development of robust crop types adapted to changing climates. Advanced phenotyping allows breeders to accurately link genotypic data with complex phenotypic traits. Traditional phenotyping is arduous and time-consuming, whereas HTP uses remote sensing, UAVs, and automated imaging systems to collect large volumes of plant data. Multispectral cameras capture detailed images of plant canopies, enabling indices such as the normalized difference vegetation index (NDVI), which measures plant vigor and biomass. This rapid evaluation lets breeders screen large populations and identify genotypes that perform well under specific conditions. HTP therefore strengthens breeding programs by providing real-time performance data across growth stages, which is especially useful for assessing stress tolerance and yield potential under varied environments. Together, genomic, bioinformatic, and HTP technologies have improved seed varieties, and GS using genome-wide marker data has proven successful for population improvement and genetic gain.

Combining genetics, bioinformatics, and HTP provides a solid foundation for crop enhancement aligned with global sustainability goals. Using genomic data to find markers associated with desirable traits, breeders can base decisions on evidence rather than intuition, boosting efficiency and innovation in climate-resilient breeding. These approaches have produced maize cultivars tolerant to biotic and abiotic stresses that maintain yield stability under water limitation, which is crucial amid growing climatic unpredictability. The integrated strategy also enables targeted biofortification to improve staple crop nutrition, reducing micronutrient deficiencies and improving human health. Such techniques have greatly improved understanding of plant genotype-phenotype relationships, enabling more accurate and focused breeding (Yoosefzadeh-Najafabadi et al., 2023). As populations grow, crops must produce more with fewer resources. Modern technologies let breeders develop high-yielding cultivars requiring less water and fertilizer, while region-adapted cultivars reduce chemical inputs and boost biodiversity, lowering agriculture's environmental impact. These improved seeds help farmers cope with pests and weather variability. As these fields advance, agriculture can become more efficient and resilient, meeting the needs of a growing population while preserving the planet's resources.

2. Era of improved seed varieties – Green : revolution to genome editing

Agricultural technologies have advanced significantly since the mid-20th century, with the Green Revolution and genome editing as the main stages, each addressing food security while raising ethical and environmental challenges. The 1960s Green Revolution transformed output in countries like India and Mexico through high-yielding varieties (HYVs) of staples such as wheat and rice, bred to respond better to fertilizers and irrigation. Indian wheat production rose from 12 million tons in 1965 to nearly 20 million tons by 1970, and national food production grew from 50.8 million tons in 1950-51 to 199.3 million tons in 1996-97, aided by chemical inputs and improved irrigation (Singh and Byerlee, 1990). Supportive policies, including minimum support prices, infrastructure investment, and farmer financing, contributed to its success. However, it was criticized for environmental degradation, widening agricultural disparity, and reliance on chemical inputs, which deteriorated soil, contaminated water, and disadvantaged smallholders. Climate change and population growth now demand more accurate and sustainable farming. CRISPR/Cas9 enables precise gene changes, producing pest-, disease-, drought-, and salt-resistant crops without foreign DNA (Bortesi and Fischer, 2015), speeding breeding and improving yield, nutrition, and resistance for a population expected to reach 10 billion by 2050 (Boserup, 1981). Regulation lags behind technology, and classification relative to conventional modification remains disputed, though some regions allow foreign-DNA-free edited crops. Genome editing supports sustainable agriculture by reducing chemical inputs while raising yields (Namo and Belachew, 2021), and can improve staple crop nutrition to combat malnutrition.

2.1 Genomic selection - the foundation of crop improvement

Genomic selection (GS) uses genome-wide markers to build a statistical model that produces Genome-Based Estimated Breeding Values (GEBVs) for choosing superior individuals. It requires a training population (TP) with genotype and phenotype data and a breeding population (BP) with genotype data only. The model, built from TP data, estimates marker effects to score BP individuals without phenotyping. GEBVs depend on allele similarity between TP loci and the target phenotype. Candidates with high GEBVs and favorable alleles can be used in

breeding to generate improved varieties (Xu et al., 2020) or released as new varieties after repeated site evaluations (Fig. 1).

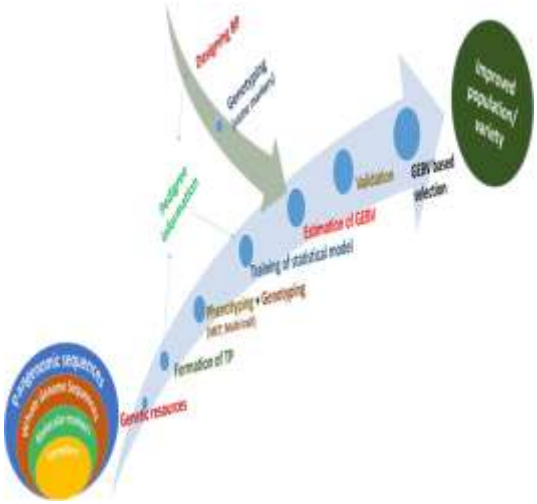


Fig 1: Overview of genomic selection process GEBV estimation and improved variety development (Bajaj et al., 2016)

2.2 Integrated plant breeding approaches

To add novel variations for desirable traits, marker-assisted selection, speed breeding, TILLING and ECO-TILLING, transgenic, cisgenics, intragenic techniques, and genome editing can be integrated to optimize GS.

2.2.1 Genotyping platforms – Marker system

Genomic selection (GS) models assume markers are in linkage disequilibrium (LD) with QTL or genes. Prediction accuracy rises with marker density but plateaus beyond an optimum, so uniformly dispersed genome-wide markers are preferable for maximizing genetic gain. Optimal density depends on LD, relatedness between training and breeding populations, trait heritability, and training-population genetic structure. SSR markers are more informative than SNPs but less available and cover less of the genome, so SNPs are used in most GS models. Marker density strongly affects accuracy, although further increases did not always help, density requirements follow LD and population structure. Self-pollinated and bi-parental populations, with high LD and strong structure, need fewer markers than wild populations (Hao et al., 2019; Smith et al., 2008). Biparental maize populations reached maximum accuracy with 1000 SNPs (Liu et al., 2018), and RIL and DH populations show substantial marker-QTL LD (Wang et al., 2018). In 363 Indica rice lines, more than 7500 markers (19 SNPs per Mb) did not improve accuracy (Spindel et al., 2015), whereas 13 SNPs per Mb in a synthetic rice population (Grenier et al., 2015) and 27 SNPs per Mb in 280 accessions (Bhandari et al., 2019) did. Lower optimal densities are associated with higher LD ($r^2 = 0.50$), and markers can be selected or excluded by their impact on accuracy (Ma et al., 2016).

Validating training-population models is essential, since most GS models overfit by estimating genetic variation with trait-irrelevant markers. Cross-validation prevents overfitting: the training model is combined with validation-set markers to estimate GEBVs, which are compared with phenotypes. Because true breeding values are unknown, predictive ability (correlation of GEBVs with observed phenotypes) is divided by the square root of heritability to estimate accuracy indirectly (Krishnappa et al., 2021). K-fold cross-validation splits data into k groups, validating one at a time; leave-one-out (k = n) suits small population. For example, 100 individuals can be split into 80 training and 20 validation, repeated to measure accuracy before finalizing the prediction equation.

2.2.2 Marker-Assisted Selection

Marker-assisted selection (MAS) overcomes the limitations of phenotype-based selection. Molecular markers have increased the precision of breeding crosses, though they have also created variation that is hard to integrate. Common approaches include MABC, MARS, and GS. MAS optimizes the assembly of target alleles, reducing unintended losses and selection cycles, which is valuable for gene pyramiding and stacking of disease resistance, productivity, and other traits. Current methods can develop high-heritability traits controlled by a single gene or a large QTL producing phenotypic variability. Many field programs use markers to improve selection and shorten varietal development time. In soybean, seed coat permeability affects storage, germination, and soy-food processing. The polygalacturonase gene PG031, limited to the cell wall and suggesting seed coat-specific activity, regulates this trait. Three H289Y haplotypes (PG031289H, PG031289Y, and PG031Hap3) cause permeability variation, and PG031 explained 15% of the phenotype in a QTL (Wang et al., 2022). Seed size is the key agronomic factor affecting soybean output, and its QTL genes, genetic architecture, molecular mechanisms, and regulatory networks have been uncovered (Zhang et al., 2024). Field selection in segregating populations demands extensive phenotyping and production resources. Once homozygous or genetically stable populations are produced, yield and economic traits can be assessed and breeding cycles shortened. Current methods include haploid breeding and speed breeding. Speed breeding manipulates environmental factors to promote early

flowering and seed set, producing 3-9 generations per year versus 1-2 in standard selection (Ghosh et al., 2018). It adjusts photoperiod, temperature, soil moisture, plant density, carbon dioxide, nutrition, hormones, and tissue culture, complementing MAS and high-throughput phenotyping (Wanga et al., 2021). It rapidly yields homozygous lines from complementary parental crosses, accelerating cultivar development and release (Table 1).

Table 1. Rapid advancement techniques for breeding cycles in different crops and selection strategies.

Crop	Techniques	50 % F	Generation /year	Selection method	Reference
Canola	Photoperiod, light intensity, temperature, immature seed germination and soil moisture	73	4	SSD	(Watson et al., 2018)
Chickpea	Photoperiod and immature seed germination	33	7	SPD	(Samineni et al., 2020)
Pea	Plant hormones, photoperiod and immature seed germination	33	5	-	(Mobini and Warkentin, 2016)
Pigeon pea	Photoperiod, temperature, immature seed germination	50-56	4	SPD	(Saxena et al., 2019)
Wheat	Photoperiod, temperature, soil fertility, immature seed germination and embryo rescue	28-41	7.6	SSD	(Zheng et al., 2013)

SSD: Single seed descent, SPD: single pod descent

2.2.3. Utilization of TILLING and ECO-TILLING for crop improvement

TILLING is a non-transgenic, reverse genetic method for most plants. Periodic chemical mutagen treatments have caused mutations for decades. Tilling causes high-density random point mutations that cause allelic missense mutations and truncation. Chemical mutagens can find natural population point mutations or polymorphisms. It employs chemicals to produce plant mutations and contemporary technology to detect and use favorable gene variants. EcoTILLING employs TILLING to find and apply advantageous gene variants (Mohapatra et al., 2023). Therefore, TIILING and EcoTIILING detect genetic variation, analyze gene function, and propose fresh genetic methods to improve plants. SNPs and PCR target gene regions for mutation detection. This unique global reverse genetic approach finds gene sequence point mutations. Mutations in TILLING are advantageous. It discovers every gene allelic change as a future reverse genetics tool, which is vital for functional genomic research in crops where genetic modification is difficult due to lack of contemporary transformation procedures. Eco-TILLING detected 1133 transcription factor seed weight-associated polymorphisms in 192 desi and kabuli chickpea accessions (Bajaj et al., 2016). In 25 wheat mutant lines, CEL-PAGE indicated EMS-induced alterations in Pinb, waxy, Agp2, and SSIIa-A, as well as 16 point mutations (nonsense mutations; C>T missense mutations). All frame shift, nonsense, and Pinb new alleles created gene allelic diversity. Wheat quality systems can employ mutant lines as intermediates. Although genome sequence access is needed for target gene selection, mutagenized strains in many species exhibit target gene allelic patterns. This method has drawbacks, but most species, including agricultural plants, have had their genomes sequenced, and better NGS technology may sequence the rest soon. Standardizing the mutagen dose can overcome genetic mutations that influence protein structure and phenotype in the TILLING population, a promising mutant population as asexually propagated plants take longer to create mutant populations. Equal high-quality DNA is needed for pooling. Reduced DNA concentration may reduce mutation efficiency. Testing several mutations with pooled samples saves time and money. Pooled samples with variances must be re-analyzed and sequenced to confirm mutations. The cost of sequencing all samples drops. Target a single-copy gene. Polyploid plants require more primer design and fewer samples due to many functional gene copies.

2.2.4 A new era of genome sequencing

Genomics studies biology-related gene function, regulation, and interconnection. Comparative functional genomics and transcriptomics in agriculture improve phenotypes by finding allelic variants. Condition-specific transcriptional activity is measured by microarrays, qPCR, MPSS, SAGE, and NGS-RNA sequencing (RNA-seq). Global transcriptome analysis of numerous tissues and developmental phases of diverse crop accessions has developed an atlas of variably regulated transcripts/genes, illustrating crop developmental complexity. Integrative genomics using global transcriptome resources can reveal gene functions (Azodi et al., 2020). Seed plasticity

RNA-Seq in two B. napus genotypes examined five co-expression modules and seed weight/quantity. Differentially expressed genes by S-S ratio. The network analysis identified numerous NAC transcription factors as hubs regulating seed weight and number adjustment (Canales et al., 2021), which could be a target gene for seed weight augmentation. MiRNA genes regulated Brassica juncea seed size, coat color, and oil. Using breeding, genetic engineering, or genome editing, DEGs are functionally identified for crop development (Jain et al., 2023).

2.2.5 Transgenics, cisgenics and intragenics

Genetically engineered crops have changed genes. Genetic engineering or transgene-based research can transfer genes from any source into plants, unlike breeding, which takes time and is confined to sexually compatible species. Maximum host species efficacy requires gene transfer and selection. A. tumefaciens-mediated genetic transformation stabilizes transgenic lines. Phenotypic changes from gene silencing and overexpression are the main transgenic methods for plant gene function detection. Enzymatic degradation of target mRNA by complementary antisense RNA or double-stranded RNA stretches inactivates the gene and stops translation. Strong promoters upstream of the gene coding region can generate gene overexpression, which is beneficial when the target gene is functionally redundant or gene silence is fatal. Bt cotton boosted India's economy, whereas Bt brinjal protected Bangladesh farmers. Transgenic crops have many benefits, but ecological risks, human safety, and policymaker antagonism limit acceptance. No research ties transgenic crops to health risks (de Vos and Swanenburg, 2018). Biotechnology is vital to agriculture's future, despite mounting hurdles. Intensive agriculture helps farmers and global grain and vegetable production. Agriculture is threatened by new diseases, pests, inappropriate rainfall, synthetic fertilizer issues, and monotony. High-vigor seeds are essential for seedling growth and crop yield under difficult conditions and they boost seeds germination, emergence, crop yield, and costs. Maize and wheat need precision seeding due to its extensive use. Viability and longevity may affect seed bank accessions' generation cycles. Storage at high temperatures and moisture may kill seeds. High-vigor seeds store well. Standard breeding prioritizes production above vigor. Physical priming (Araujo et al., 2016) and other invigorations improve commercial seed lot vigor. Most cost-efficient, sustainable, and effective technique to improve seed vigor is with GM technology. We know rice genotype and OsPIMT1's role in seed vigor and longevity. After 21 days of CDT, transgenic seeds overexpressing OsPIMT1 germination rose by 9-15 % (Wei et al., 2015), however OsPIMT1RNAi lines lost germination soon, as found in rice (Irshad et al., 2020). OsPIMTs replenish antioxidant enzymes and proteins to minimize ROS and lipid peroxidation, extending seed life. OsPIMT1 and 2 overexpression in transgenic rice increases seed vigor and lifetime (Petla et al., 2016). Future transgenics laws may be tougher. GM crops are designed for industrial farms, thus they may not help subsistence farmers in developing nations. Conservators and ecologists recommend examining transgenic crops' environmental consequences and delaying their introduction. Cisgenesis and intragenesis were created as transgenic approaches to address these difficulties while retaining efficient and ecologically friendly plant production. Both concepts use genetic material from the same or sexually compatible, hybridizing gene pools. In contrast, transgenesis transfers genes and DNA between species. Traditional breeding uses the same gene pool as cisgenesis and intragenesis. Intragenic or cisgenic transformants and their offspring should not have selection marker or vector-backbone genes (Vasudevan et al., 2023). Most nations control cisgenic and intragenic crops like transgenics. Europe, North America, Japan, Australia, and New Zealand use cisgenic plants most. Canada, the US, and Australia regulate cisgenic plants more rigorously than Europe, Japan, and New Zealand. Various public attitudes and factors impact consumer choices in different regions. Europeans prefer cisgenic plants and products for environmental concerns over quality and health. New Zealand disputes cisgenic plant effects on meat export and tourism (Daye et al., 2023).

2.2.6 Gene Editing Technologies

Genomic selection, genome editing, and modern breeding have shortened breeding cycles and enhanced genetic gain. Genetically modified crops like SDN1 and SDN2 are excluded under EPA, 1986 Rules 7-11 of 1989 because they lack exogenous genes/DNA. CRISPR and genomic selection increase genetic development at optimal chromosomal recombination sites (Bernardo, 2017). Many pulse crop genomes, including re-sequenced accessions, are publicly annotated (Varshney et al., 2019). Pulses can be converted with Agrobacterium tumefaciens in several ways, first-commercialized genome-edited high-oleic oil soybean shows how genome editing can improve pulse crops (Kim and Kim, 2019). CRISPR/Cas9 and heterosis usage were utilized to target the qSH1 gene, which causes rice seed shattering, to create hybrid rice cultivars with intermediate seed shattering, boosting stability and yields (Sheng et al., 2020). However, genetic selection, model development, and molecular technology integration across several crops are limited.

3. Bioinformatics

It includes database collecting and analysis and biological data (interpretation software and algorithms. Biological (genomic, transcriptomic, proteomic, and metabolomic) data nucleotides, amino acid sequences, protein domains, structures, and expression patterns) from numerous organisms must be stored and supplied.

3.1 Data Management and Integration

NGS and bioinformatics investigations have generated much plant genomic data. Submit data to free, focused databases online. PeanutBase collects genetic and genomic data on numerous Arachis species, including cultivated peanut. Oryzabase (rice), SoyBase (soybean), and Phytozome comprise 404 assembled and annotated genomes from 174 Archaeplastida species, including 54 Brachypodiumdistachyon lines from BrachyPan, 20 Brassicales

Map Alignment Project species, and 8 Co. cowpea genomes. NCBI (https://www. ncbi. nlm. nih. gov/) has 12,205 plant genomes through 2024. This database aids non-genus/species research. Researchers may readily access all genomic data in one database. Public plant genome databases for many genus and species are briefly described here. For GEO or SRA plant species genomic information, provide the scientific name. From each dataset's profiling platform, researchers can get gene symbols, Ensemble IDs, open reading frames, chromosomal locations, regulatory elements, etc. Data analysis tools include Gene Ontology, DAVID, and BLAST. Complete genome sequence databases relate gene sequence, function, and genetic map position across genomes. Software uses high-scoring genetic markers to anticipate hypotheses and add desirable plant traits for breeding reliability. Plant genomes are autonomously annotated by EnsemblPlants (https://plants.ensembl.org/index.html), correlated with other publicly available biological data, and stored in an open-access database for researchers. Other major plant databases are PlantGDB, Maize-DIG, soyBase, and Phytozome. UNITE (https://unite.ut.ee/) for fungal identification, SILVA for 16S rRNA (Quast et al., 2013), MGnify for microbiomemetagenomics data (Mitchell et al., 2020) and the Plant Promoter Database are also relevant. The platform provides genomic sequence data and other analytical information on relevant plant species, allowing plant bioinformatics specialists cut analysis time. Researchers may reevaluate their conclusions if their investigation is complete. Identifying phenotypic and function changes in genomics-proteomics-metabolomics linkages requires metabolite databases and genome sequence data. The key metabolomics datasets for flora and crops are Metlin (2,40,000 metabolites) and PlantCyc (72,000 high-resolution MS/MS spectra), which contains metabolic pathways, catalytic enzymes, and genes (Filho et al., 2018; Gomez-Casati et al., 2018). NGS lets RNAseq directly analyze mRNA profiles and identify SNPs (Tan et al., 2022). The NGS-bioinformatics plant breeding technique is shown in Figure 2.

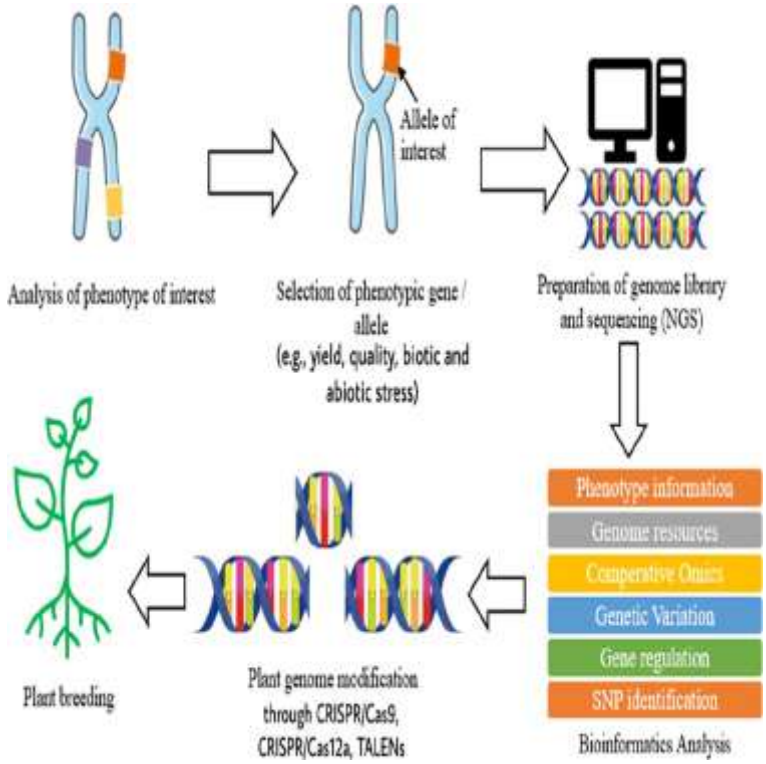


Fig 2. Overview of NGS and bioinformatics applications in the plant breeding process (Kim et al., 2019)

3.2 Analysis of genomic data

Understanding agronomic traits requires isolating complex variables into chromosomal loci and defining each quantitative trait locus. Sequence-based markers and QTL analysis, including single-nucleotide markers, provide precise and efficient genotyping through marker-based selection systems that closely influence complex genetic traits (Thomson, 2014). Rapid advances in sequencing and bioinformatics have improved gene mapping resolution, gene identification, and genome-wide association studies (GWAS) for crop development through NGS-based genotyping. Such studies have identified soybean loci and candidate genes linked to seed weight (Jing et al., 2018), protein and oil content (Li et al., 2019), and pod dehiscence (Hu et al., 2019). AhMITE1 markers, developed from whole-genome re-sequencing of 33 peanut genotypes, support genetic research and mapping (Gayathri et al., 2018). In rapeseed, novel QTLs were found for yield-related traits, notably ovule number per pod, and advances in single-locus and multi-locus GWAS were highlighted to improve efficacy and reliability for complex traits (Khan et al., 2021). The 60K Illumina Infinium SNP array in B. napus can be converted into a cost-effective, high-throughput, gene-based screening approach for gene mapping (Clarke et al., 2016).

3.3 Computational Tools for Genomic Analysis

Advanced references, pangenomes, and high-throughput genotyping and phenotyping enable precise genomic prediction and breeding, while comparative genomics reveals constrained deleterious variants and deep learning transforms protein design.

3.3.1 Synteny

Synteny (collinearity) is an important topic in comparative genomics. It involves conserving gene organization on chromosomes inherited from a common ancestor to identify conserved synteny blocks in highly diverse eukaryotes. MCscan, MCScanX, and JCVI use dynamic programming to build pairwise collinear gene chains. Synteny is studied using WGDI, GenoDup, WGDdetector, and WGD. Comparative genomics studies non-coding sequence evolution along with gene order and content. Modern methods can align large DNA sequences at the chromosomal or genome level, allowing single-nucleotide comparisons of two genomes. These include Minimap2, MUMmer, and AnchorWave.

3.3.2 Epigenetics and transcriptome profiling

Omics approaches analyze epigenetic marks, cistromes, and chromatin states to annotate non-coding genomic regions, as demonstrated in ENCODE's human and mouse genomes and some crops. DNA cytosine methylation, essential for gene expression and imprinting, is detected by bisulfite sequencing, which converts non-methylated cytosines to uracils, read as thymines after PCR. Aligners must handle natural variation and treatment-induced mismatches; popular options include BatMeth2, BSMAP, Bismark, BS-Seeker2, BWA-meth, and BISCUIT (Zhou et al., 2019). Differential methylation is assessed with tools like methylKit, DSS, and metilene, while ChIP-seq, DAP-Seq, CUT&Tag, and ATAC-Seq profile chromatin state.

3.3.3 mRNA profiling

RNA-seq underpins RNA biology, enabling transcriptome-wide analysis, splicing, single-cell, and spatial studies, while long-read methods, including nanopore direct RNA sequencing, increasingly identify alternative splicing isoforms and modifications. Reads are quality-trimmed with Trimmomatic or sickle and aligned using splice-aware mappers: HISAT2, which replaced TopHat2 with better efficiency and accuracy (Kim and Kim, 2019), and STAR, which detects splice-junction and fusion reads (Dobin et al., 2013). Benchmarking highlights the importance of tool choice and parameter optimization (Krizanoyic et al., 2018).

Differentially expressed genes are identified with DESeq2, Cuffdiff, and edgeR, which use negative binomial generalized linear models to handle few replicates, discreteness, wide dynamic range, and outliers (Love et al., 2014), the Wilcoxon rank-sum test better controls false discovery rate in large samples (Li et al., 2022). Co-expression and regulatory analyses include WGCNA, which finds modules, hubs, and network topologies, and GENIE3, which predicts target expression using random forests or extra trees. MEME Suite identifies motifs enriched in co-regulated genes.

3.3.4 Plant breeding by design

Genotyping still uses NGS short-read technology due to their reduced cost. Trim low-quality nucleotides, adaptors, and barcodes with Trimmomatic (Bolger et al., 2014) or Sickle (https://github.com/najoshi/sickle) after re-sequencing and evaluate quality with FastQC Mapping reads to a reference genome Bwa-mem and bowtie2 are top tools. Long-read sequencing has become popular because it makes functional mapping of structural variations (SVs) cheaper. SVs include CNVs, insertions, inversions, translocations, repetitive sequence expansions, and complex combinations beyond 50 bp. Longread sequencing finds many SVs by comparing reads or matching to a reference genome. Alignment-based SV callers include CuteSV, Duet, NanoVar, SVIM, Picky, and Sniffles2 (Smolka et al., 2024). To assign SVs, assembly-based SV calling methods link contigs to a reference genome. Examples include SVIM-asm, PAV, Assemblytics, and SyRI. Initial GWAS analyzed each variation with the null hypothesis of no genotype-group phenotypic difference and performed several modifications. This method has substantial false positive rates due to population stratification and relatedness. The GLM addresses this by include population structure as cofactors. Population structure is determined by major components or subpopulation membership. Fixed population structure and random kinship effects improve GLM in the mixed linear model (MLM). MLM equations are calculated faster using EMMA, P3D, EMMAX, Fast-LMM, and GEMMA. Recent fastGWA improved biobank-scale GWAS by minimizing population stratification by main components and relatedness with a sparse genetic relationship matrix (Jiang et al., 2019). Using a lower-rank kinship matrix, researchers produced compressed MLM (CMLM) and enriched CMLM (ECMLM) to increase GWAS power. Many markers in association models like MLMM, FarmCPU, and BLINK increase GWAS power. CNVs and PAVs are becoming more popular genotype data, but SNPs and short InDels remain the most common. Additionally, k-mer-based GWAS uncovered QTLs not in the reference genome (Voichek and Weigel, 2020). Despite widespread association and linkage studies, pinpointing the causative variation behind a QTL is difficult due to linkage disequilibrium between the variant and adjacent variables. Computing tools like VEP and SnpEff estimate variant impacts, followed by wet lab tests to determine chemical or physiological consequences. Two main plant molecular breeding strategies employ genes or QTLs. Genome-wide markers evaluate genomes for breeding in genomic selective breeding. This strategy usually requires little target genetic analysis. Instead, training populations acquire significant genetic and phenotypic data. Breeding values are predicted using rrBLUP, gBLUP, sommer, SeqBreed, BGLR, ASREML, and CropGS-Hub [81]. Genomic predictions employ DNNGP and DeepGS. Recent molecular breeding advances include gene editing. It can accurately modify weaker alleles

and create crop mutants. RiceNavi helps breeders detect inferior alleles. CRISPR-GE and CRISPRdirect generate geneediting primers, and Hi-Tom checks positive altered lines.

4. High-Throughput Phenotyping: Accelerating Trait Evaluation

Crop phenomics uses sensors, imaging, and analytics for high-throughput phenotyping, linking sequenced genes to agronomic traits, assessing stress and yield, and highlighting current limitations and future research directions.

4.1. Technologies Used in High-Throughput Phenotyping

Non-destructive sensing and imaging have greatly improved crop phenotyping in controlled and field settings, integrating robotics, computing, and aeronautics across microscopic, ground-based, and aerial platforms. Despite challenges like invasive pre-treatment, low microscale resolution, and limited automated analysis software, modern algorithms extract tissue and organ micro-phenotypes. Visible light and laser scanning measure external traits, supporting breeding programs and seed nutritional assessment [fig. 3].

4.1.1 Micro-computed tomography (micro-CT)

X-ray CT offers non-invasive, precise tissue measurement and has become increasingly popular in plant research. Seed functional and nutritional assessment is vital for breeding programs and global nutrition. Hughes' 2D morphological CT pipeline struggles to separate wheat grains from noisy backgrounds and adjacent tissues, whereas a newer 3D morphology pipeline automatically and non-destructively measures wheat spike grains and stem nodes, adding hard-to-measure 3D traits such as grain aspect ratio, porosity, grain-to-grain distance, and grain angle (Xiong et al., 2019). For rapeseed cell imaging, three methods were developed: trainable Weka Segmentation, Cellpose-Fiji (human-like segmentation of bright-field seed coat cells), and LabelsToROIs (CML), which performs well on single cells and streamlines micro-phenotype data processing (Jiao et al., 2022). Seed quality research aims to improve germination, growth, yield, and product quality. Radiography identifies morphological predictors of germination, sorts seeds, reveals internal structures, and tracks maturation, with internal free space, embryo abnormalities, and tissue defects informing predictions [91]. Micro-CT has linked internal morphometry to germination in tomato seeds (Gargiulo et al., 2020), enabled high-throughput phenotyping of wheat seed traits under drought and heat stress (Schmidt et al., 2020), distinguished soybean cultivars by density and physical properties (Hacisalihoglu and Armstrong, 2023;Yang et al., 2020), and classified maize hybrids by kernel hardness through volume, density, porosity, and cavity measurements (Guelpa et al., 2015). Seed traits divide into externally visible phenotypes (EVP), such as color, weight, shape, size, and damage from disease, pests, or injury, and internally invisible phenotypes (IIP), including biochemical composition (protein, water, starch, oil, fatty acids) and physiological structures (embryo, endosperm, cavity). Traditional IIP analysis is slow and potentially destructive, making it unsuitable for early-generation selection (Liu et al., 2024) [96]. Spectroscopy and spectral imaging digitize IIP traits; hyperspectral transmittance imaging classified hybrids and female parents for purity (Zhang et al., 2022). Time-lapse 4D synchrotron X-ray CT revealed heterogeneous root-soil deformation among phenotypes (Keyes et al., 2017).

4.1.2 Ground based phenotyping

Ground-based platforms assess individual and field-plot phenotypes, either portable or stationary. Portable devices, including smartphone apps like Leaf-GP, are low-cost and convenient but limited in scale and throughput. Stationary platforms combining RGB, multi/hyperspectral, thermal, fluorescent, and 3D sensors improve precision and throughput examples include CropDesign and Field Scanalyzer . Mobile platforms mount sensors on carts, tractors, or robots, such as LiDAR carts measuring canopy height and biomass. Sunlight variability is mitigated by enclosures like GPhenoVision.

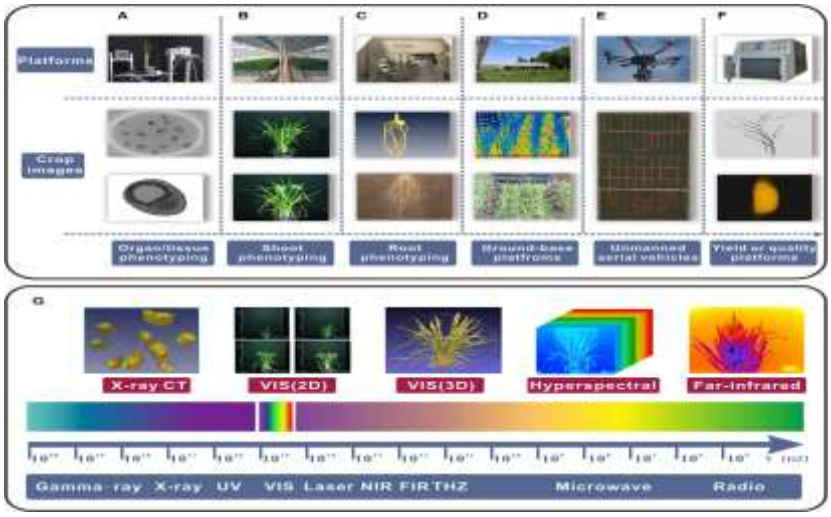


Fig 3. A schematic of phenotyping platforms at various scales. Organ/tissue (A), shoot (B), root (C), ground-base (D), and UAV (E) phenotyping for crop yield/quality (F) and illustrating several spectral methods (G) (Samineni et al., 2019).

4.1.3 Aerial based phenotyping

Satellites and UAVs increasingly support field phenotyping. Airborne systems, mainly UAVs, outperform ground platforms in large-scale breeding, assessing canopy temperature, structure, chlorophyll, nitrogen, height, and biomass across thousands of plots. Satellite imagery enables large-scale cultivar growth and genetic study (, though resolution limits monitoring [108]. UAVs are cheaper, fly lower, and are more accurate than helicopters, measuring maize and sorghum height via SfM (Malambo et al., 2018) and tomato yield (Chang et al., 2021), but face regulatory, payload, flight-time, weather, and data-processing limits (Jin et al., 2021). Satellites provide broad, high-throughput monitoring with lower precision, supporting yield prediction and crop health (Zhang et al., 2020; Prey et al., 2020]. Cost restricts use, though missions track drought and disease and microsatellites aid sustainability (Jain and Balwinder-singh, 2019).



Figure. 4. Dividing phenotypes into externally visible and internally invisible allows for efficient digitization approaches based on their properties, promoting the automation of seed testing (Sheng et al., 2020)**.**

4.2 Data Analysis in High-Throughput Phenotyping

High-throughput phenotyping enables efficient, non-destructive field phenotyping, but sensor-equipped platforms generate huge datasets requiring storage and processing. Machine learning and deep learning analyze them. Key DL models include GANs, RNNs, CNNs, and MLPs, deep CNNs achieve state-of-the-art image segmentation, detection, and classification.

Table 2. Different machine and deep learning methodologies for the treatment of abiotic stress in agricultural crops

Plant	Traits	Spectral range (nm)	Domain (field/ lab)	Sensing modality	Data analysis	References
Wheat	Water stress	841-1652	Landsat data	Multispectral imaging	Indices and regression	(Dangwal et al., 2016)
Cotton	Water stress	490-900 7500-13500	Field	Thermal imaging	Mapping and correlation	(Bian et al., 2019)
Chickpea	Salinity	380-760	Greenhouse	RGB imaging	PLS and correlation analysis	(Atieno et al., 2017)
Wheat	Nitrogen, phosphorus, potassium, sulphur	350-2500	Field	Hyperspectra l imaging	Regression	(Mahajan et al., 2014)
Sorghum	Grain quality	North Star Imaging X50 industrial 3D X-ray inspection system	Lab	X-ray CT	Restricted maximum likelihood (REML) and correlation	(Crozier et al., 2022)

4.3 Integration of High-Throughput Phenotyping and genomic data

HTP provides accurate quantitative data on large training populations, and combining it with genomic selection improves precision, efficiency, and labor savings in selecting superior cultivars.

4.3.1.RepeatabilityModel

Repeated measurements of the same genotype show comparable variation regardless of record time, with genetic

correlations of one across observations. Simple repeatability models generated BLUPs and BLUEs of longitudinal HTP variables for wheat genomic prediction (Sun et al., 2017).

4.3.2 Multiple-Trait Model

Each developmental time point is treated as a separate dependent variable (Sun et al., 2017). MANOVA resolves the covariance-structure problems of repeated-measures ANOVA.

4.3.3 Random Regression Model

Closer measurement intervals correlate more strongly. RRMs capture temporal genetic variability and give more accurate breeding values despite higher computing costs (Oliveira et al., 2019a). Genomic BLUPs are used for selection, though high-density genotyping may be impractical for all genotypes. CIMMYT's wheat program uses UAV-derived secondary traits, phenotyped before grain yield, to improve GS accuracy and enable early prediction. Modeling trait correlations improves multi-trait GS accuracy, despite abundant data, HTP has yielded few biological insights, though a GWAS identified the heat-related gene hsp18f (Adak et al., 2023).

5. Challenges and Considerations of advancement in genomics from lab to land

Plant genomics breakthroughs require substantial biology expertise and intuition, built through practice and exposure to varied challenges. Beginners must grasp data visualization, using browsers like IGV, JBrowse, and GBrowse. Installing academic software is often difficult because of missing dependencies and PATH issues; separate environments and Bioconda help. Software must also be benchmarked on suitable datasets with correct parameters, or findings may be misread.

Moving plant genomics from lab to field faces further challenges. Genotype-environment interactions make modifications hard to predict across soils, temperatures, and stresses. Breeding crops resilient to rapidly changing climates requires forecasting future conditions. Scaling from small laboratory production to large-scale agriculture is logistically and technically demanding, involving seed production, distribution, and new agronomic technologies. GMO regulations are restrictive, time-consuming, and expensive, while public skepticism demands clear communication and education. Genetically modified crops may also harm ecosystems through gene transfer to wild relatives and non-target effects. Collaboration among geneticists, breeders, agronomists, ecologists, legislators, and farmers is essential, given genomics' potential to improve yield, nutrition, and sustainability.

6. Conclusion

Integrating genomics, bioinformatics, and HTP to generate improved seed varieties improves our understanding of plant genetics and streamlines the breeding process, making it more efficient and precise. Public approval and other factors will influence enhanced seed varieties in the future. Public perception is one of the biggest obstacles facing genome editing technology, and each phase has helped address food security issues while raising new ethical and environmental issues that must be controlled. Gaining public trust requires transparency in risk communication and technology advantages. Developing regulatory frameworks that balance innovation and safety requires international collaboration among scientists, politicians, and stakeholders. Genomics helps identify agricultural productivity factors, whereas bioinformatics helps interpret complex genetic data. Breeders can focus on genetic indicators to improve yield, disease resistance, and environmental adaptability. HTP is essential for speedy and accurate phenotypic analysis in many circumstances, providing insights that standard approaches cannot. Numerous studies have shown that these technologies can address climate change and food security, and resilient agricultural types are needed owing to population growth and environmental changes. Genomics, bioinformatics, and high-throughput technology will improve breeding and seed production. These scientific advances must be implemented by researchers, agronomists, and policymakers for farmers and consumers. Technological advances in crop enhancement enable sustainable agriculture that ensures food security for future generations. Multidisciplinary approaches may help create a more resilient agricultural system that can adjust to changing conditions. Future research should emphasis sustainability by generating crops with low inputs and high yields using traditional breeding methods and genomic tools.

References

1. Adak, A., Kang, M., Anderson, S.L., Murray, S.C., Jarquin, D., Raymond, K.W., Wong, A. and Katzfuss, M (2023). Phenomic data-driven biological prediction of maize through field-based high-throughput phenotyping integration with genomic data. J Exp Bot 74(17): 5307–5326. https://doi.org/10.1093/jxb/erad216
2. Araujo, S.S., Paparella, S., Dondi, D., Bentivoglio, A., Carbonera, D. and Balestrazzi, A (2016). Physical methods for seed invigoration: advantages and challenges in seed technology. Front Plant Sci 7: 646. https://doi.org/10.3389/fpls.2016.00646
3. Atieno, J., Li, Y. and Langridge, P (2017). Exploring genetic variation for salinity tolerance in chickpea using image-based phenotyping. Sci Rep 7: 1300. https://doi.org/10.1038/s41598-017-01211-7
4. Azodi, C.B., Pardo, J., VanBuren, R., de Los Campos, G. and Shiu, S.H (2020). Transcriptome-based prediction of complex traits in maize. Plant Cell 32(1): 139–151. https://doi.org/10.1105/tpc.19.00332\
5. Bajaj D, Srivastava R, Nath M, Tripathi S, Bharadwaj C, Upadhyaya HD, Tyagi AK, Parida SK (2016) EcoTILLING-based association mapping efficiently delineates function-ally relevant natural allelic variants of candidate genes governing agronomic traits in chickpea. Front in Plant Science, 19(7): 450. doi: 10.3389/fpls.2016.00450

6. Bajaj, D., Srivastava, R., Nath, M., Tripathi, S., Bharadwaj, C., Upadhyaya, H.D., Tyagi, A.K. and Parida, S.K (2016). EcoTILLING-based association mapping efficiently delineates functionally relevant natural allelic variants of candidate genes governing agronomic traits in chickpea. Front Plant Sci 7: 450. https://doi.org/10.3389/fpls.2016.00450
7. Bian, J., Zhang, Z. and Chen, J (2019). Simplified evaluation of cotton water stress using high resolution unmanned aerial vehicle. Remote Sens 11(3): 267. https://doi.org/10.3390/rs1103267
8. Bolger AM, Lohse M, Usadel B (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30(15): 2114–2120. DOI:10.1093/bioinformatics/btu170
9. Bortesi, L. and Fischer, R (2015). The CRISPR/Cas9 system for plant genome editing and beyond. Biotechnol Adv 33(1): 41–52. https://doi.org/10.1016/j.biotechadv.2014.12.006
10. Boserup, E (1981). Population and Technological Change: A Study of Long-Term Trends. University of Chicago Press, Chicago.
11. Canales, J., Verdejo, J., Carrasco-Puga, G., Castillo, F.M., Arenas-M, A. and Calderini, D.F (2021). Transcriptome analysis of seed weight plasticity in Brassica napus. Int J Mol Sci 22(9): 4449. https://doi.org/10.3390/ijms22094449
12. Chang, A., Jung, J., Yeom, J., Maeda, M.M., Landivar, J.A. and Enciso, J.M (2021). Unmanned aircraft system- (UAS-) based high-throughput phenotyping (HTP) for tomato yield estimation. J Sens 2021: 8875606. https://doi.org/10.1155/2021/8875606
13. Clarke, W.E., Higgins, E.E., Plieske, J., Wieseke, R., Sidebottom, C., Khedikar, Y., Batley, J., Edwards, D., Meng, J. and Li, R (2016). High-density SNP genotyping array for Brassica napus and its ancestral diploid species based on optimised selection of single-locus markers in the allotetraploid genome. Theor Appl Genet 129(10): 1887–1899. https://doi.org/10.1007/s00122-016-2746-7
14. Crozier, D., Riera-Lizarazu, O. and Rooney, W.L (2022). Application of X-ray computed tomography to analyze the structure of sorghum grain. Plant Methods 18(1): 3. https://doi.org/10.1186/s13007-022-00837-7
15. Dangwal, N., Patel, N.R., Kumari, M. and Saha, S.K (2016). Monitoring of water stress in wheat using multispectral indices derived from Landsat-TM. Geocarto Int 31(6): 682–693. https://doi.org/10.1080/10106049.2015.1073369
16. Daye, C., Spök, A., Allan, A.C., Yamaguchi, T. and Sprink, T (2023). Social acceptability of cisgenic plants: public perception, consumer preferences, and legal regulation. In: Transgenic and Edited Plants. Springer, Cham, pp 43–75. https://doi.org/10.1007/978-3-031-10721-4_3
17. De Vos, C.J. and Swanenburg, M (2018). Health effects of feeding genetically modified (GM) crops to livestock animals: a review. Food Chem Toxicol 117: 3–12. https://doi.org/10.1016/j.fct.2017.08.031
18. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29(1): 15–21. doi: 10.1093/bioinformatics/bts635
19. Filho, H.A., Machicao, J. and Bruno, O.M (2018). A hierarchical model of metabolic machinery based on the core decomposition of plant metabolic networks. PLoS ONE 13(5): e0195843. https://doi.org/10.1371/journal.pone.0195843
20. Gargiulo, L., Leonarduzzi, C. and Mele, G (2020). Micro-CT imaging of tomato seeds: predictive potential of 3D morphometry on germination. Biosyst Eng 200: 112–122. https://doi.org/10.1016/j.biosystemseng.2020.09.003
21. Gayathri, M., Shirasawa, K., Varshney, R.K., Pandey, M.K. and Bhat, R.S (2018). Development of AhMITE1 markers through genome-wide analysis in peanut (Arachis hypogaea L.). BMC Res Notes 11: 10. https://doi.org/10.1186/s13104-017-3121-8
22. Ghosh, S., Watson, A., Gonzalez-Navarro, O.E., Ramirez-Gonzalez, R.H., Yanes, L., Mendoza-Suárez, M., Simmonds, J., Wells, R., Rayner, T., Green, P., Hafeez, A., Hayta, S., Melton, R.E., Steed, A., Sarkar, A., Carter, J., Perkins, L., Lord, J., Tester, M. and Hickey, L.T (2018). Speed breeding in growth chambers and glasshouses for crop breeding and model plant research. Nat Protoc 13(12): 2944–2963. https://doi.org/10.1038/s41596-018-0072-z
23. Gomez-Casati, D.F., Busi, M.V., Barchiesi, J., Peralta, D.A., Hedin, N. and Bhadauria, V (2018). Applications of bioinformatics to plant biotechnology. Curr Issues Mol Biol 27: 89–104. https://doi.org/10.21775/cimb.027.089
24. Grenier, C., Cao, T.V., Ospina, Y., Quintero, C., Chatel, M.H., Tohme, J., Courtois, B. and Ahmadi, N (2015). Accuracy of genomic selection in a rice synthetic population developed for recurrent selection breeding. PLoS ONE 10(8): e0136594. https://doi.org/10.1371/journal.pone.0136594
25. Guelpa, A., du Plessis, A., Kidd, M. and Manley, M (2015). Non-destructive estimation of maize (Zea mays L.) kernel hardness by means of an X-ray micro-computed tomography (µCT) density calibration. Food Bioprocess Technol 8(7): 1419–1429. https://doi.org/10.1007/s11947-015-1502-3
26. Hacisalihoglu, G. and Armstrong, P (2023). Crop seed phenomics: focus on non-destructive functional trait phenotyping methods and applications. Plants 12(5): 1177. https://doi.org/10.3390/plants12051177
27. Hao, Y., Wang, H., Yang, X., Zhang, H., He, C., Li, D., Li, H., Wang, G., Wang, J. and Fu, J (2019). Genomic prediction using existing historical data contributing to selection in biparental populations: a study of kernel oil in maize. Plant Genome 12(2): 180025. https://doi.org/10.3835/plantgenome2018.05.0025
28. Heffner, E.L., Sorrells, M.E. and Jannink, J.L (2010). Genomic selection for crop improvement. Crop Sci 50(1): 1–10. https://doi.org/10.2135/cropsci2009.11.0662

29. Hu, D., Kan, G., Hu, W., Li, Y., Hao, D., Li, X., Yang, H., Yang, Z., He, X. and Huang, F.J (2019). Identification of loci and candidate genes responsible for pod dehiscence in soybean via genome-wide association analysis across multiple environments. Front Plant Sci 10: 811. https://doi.org/10.3389/fpls.2019.00811
30. Irshad, A., Guo, H., Zhang, S. and Liu, L (2020). TILLING in cereal crops for allele expansion and mutation detection by using modern sequencing technologies. Agronomy 10(3): 405. https://doi.org/10.3390/agronomy10030405
31. Jain, R., Dhaka, N., Yadav, P., Sharma, M.K., Danish, M.D., Sharma, S., Kumari, S., Vashisht, I., Singh, R.K.B. and Sharma, R (2023). Integrated analysis of transcriptomic and small RNA sequencing data provides miRNA candidates for engineering agronomically important seed traits in Brassica juncea. Curr Plant Biol 35-36: 100306. https://doi.org/10.1016/j.cpb.2023.100306
32. Jiao, Y., Liang, B., Yang, G., Xin, Q. and Hong, D (2022). A simple and efficient method to quantify the cell parameters of the seed coat, embryo and silique wall in rapeseed. Plant Methods 18: 117. https://doi.org/10.1186/s13007-022-00948-1
33. Jin, X., Zarco-Tejada, P., Schmidhalter, U., Reynolds, M.P., Hawkesford, M.J., Varshney, R.K., Yang, T., Nie, C., Li, Z., Ming, B., Xiao, Y., Xie, Y. and Li, S (2021). High-throughput estimation of crop traits: a review of ground and aerial phenotyping platforms. IEEE Geosci Remote Sens Mag 9(1): 200–231. https://doi.org/10.1109/MGRS.2020.2998816
34. Keyes, S.D., Cooper, L., Duncan, S., Koebernick, N., Fletcher, D.M.M., Scotson, C.P. and Van Veelen, A (2017). Measurement of micro-scale soil deformation around roots using four-dimensional synchrotron tomography and image correlation. J R Soc Interface 14(136): 20170566. https://doi.org/10.1098/rsif.2017.0566
35. Khan, S.U., Saeed, S., Khan, M.H.U., Fan, C., Ahmar, S., Arriagada, O., Shahzad, R., Branca, F. and Mora-Poblete, F (2021). Advances and challenges for QTL analysis and GWAS in the plant-breeding of high-yielding: a focus on rapeseed. Biomolecules 11(10): 1516. https://doi.org/10.3390/biom11101516
36. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol 37(8): 907–915. doi: 10.1038/s41587-019-0201-
37. Kim, J. and Kim, J (2019). New era of precision plant breeding using genome editing. Plant Biotechnol Rep 13(5): 419–421. https://doi.org/10.1007/s11816-019-00581-w
38. Krishnappa, G., Savadi, S., Tyagi, B.S., Singh, S.K., Masthigowda, M.H., Kumar, S., Mishra, C.N., Khan, H., Krishnappa, G., Govindareddy, U. and Singh, G (2021). Integrated genomic selection for rapid improvement of crops. Genomics 113(3): 1070–1086. https://doi.org/10.1016/j.ygeno.2021.02.007
39. Krizanovic K, Echchiki A, Roux J, Sikic M (2018). Evaluation of tools for long read RNA-seq splice-aware alignment. Bioinformatics 34(5): 748–754. doi: 10.1093/bioinformatics/btx668.
40. Li, D., Zhao, X., Han, Y., Li, W. and Xie, F (2019). Genome-wide association mapping for seed protein and oil contents using a large panel of soybean accessions. Genomics 111(1): 90–95. https://doi.org/10.1016/j.ygeno.2018.01.004
41. Liu, F., Yang, R., Chen, R., Guindo, M.L., He, Y., Zhou, J., Lu, X., Chen, M., Yang, Y. and Kong, W (2024). Digital techniques and trends for seed phenotyping using optical sensors. J Adv Res 63: 1–16. https://doi.org/10.1016/j.jare.2023.11.010
42. Liu, F., Yang, R., Chen, R., Guindo, M.L., He, Y., Zhou, J., Lu, X., Chen, M., Yang, Y. and Kong, W (2024). Digital techniques and trends for seed phenotyping using optical sensors. J Adv Res 63: 1–16 https://doi.org/10.1016/j.jare.2023.11.010
43. Liu, X., Wang, H., Wang, H., Guo, Z., Xu, X., Liu, J., Wang, S., Li, W.X., Zou, C., Prasanna, B.M. and Olsen, M.S (2018). Factors affecting genomic selection revealed by empirical evidence in maize. Crop J 6(4): 341–352. https://doi.org/10.1016/j.cj.2018.03.005
44. Love MI. Huber W, Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15: 550.
45. Ma, Y., Reif, J.C., Jiang, Y., Wen, Z., Wang, D., Liu, Z., Guo, Y., Wei, S., Wang, S., Yang, C., Wang, H., Yang, C., Lu, W., Xu, R., Zhou, R., Wang, R., Sun, Z., Chen, H., Zhang, W., Wu, J., Hu, G., Liu, C., Luan, X., Fu, Y., Guo, T., Han, T., Zhang, M., Sun, B., Zhang, L., Chen, W., Wu, C., Sun, S., Yuan, B., Zhou, X., Han, D., Yan, H., Li, W. and Qiu, L (2016). Potential of marker selection to increase prediction accuracy of genomic selection in soybean (Glycine max L.). Mol Breed 36(8): 113. https://doi.org/10.1007/s11032-016-0504-9
46. Mahajan, G.R., Sahoo, R.N. and Pandey, R.N (2014). Using hyperspectral remote sensing techniques to monitor nitrogen, phosphorus, sulphur and potassium in wheat (Triticum aestivum L.). Precis Agric 15(5): 499–522. https://doi.org/10.1007/s11119-014-9348-7
47. Malambo, L., Popescu, S.C., Murray, S.C., Putman, E., Pugh, N.A. and Horne, D.W (2018). Multitemporal field-based plant height estimation using 3D point clouds generated from small unmanned aerial systems high-resolution imagery. Int J Appl Earth Obs Geoinf 64: 31–42. https://doi.org/10.1016/j.jag.2017.08.014
48. Mitchell, A.L., Almeida, A., Beracochea, M., Boland, M., Burgin, J., Cochrane, G. and et al (2020). MGnify: the microbiome analysis resource in 2020. Nucleic Acids Res 48(D1): D570–D578. https://doi.org/10.1093/nar/gkz1035
49. Mobini, S.H. and Warkentin, T.D (2016). A simple and efficient method of in vivo rapid generation technology in pea (Pisum sativum L.). In Vitro Cell Dev Biol Plant 52(5): 530–536. https://doi.org/10.1007/s11627-016-9772-7

50. Mohapatra, S.R., Majhi, P.K., Mondal, K. and Samantara, K (2023). TILLING and Eco-TILLING: Concept, Progress, and Their Role in Crop Improvement. In: Advances in Crop Improvement, vol 1. Springer, Cham, pp 349–377. https://doi.org/10.1007/978-3-031-28146-4_14
51. Namo, F.M. and Belachew, G.T (2021). Genome editing technologies for crop improvement: current status and future prospective. Plant Cell Biotechnol Mol Biol 22(15&16): 1–19.
52. Oliveira, H.R., Brito, L.F., Lourenco, D.A.L., Silva, F.F., Jamrozik, J. and Schaeffer, L.R (2019). Invited review: advances and applications of random regression models: from quantitative genetics to genomics. J Dairy Sci 102(9): 7664–7683. https://doi.org/10.3168/jds.2019-16265
53. Petla, B.P., Kamble, N.U., Kumar, M., Verma, P., Ghosh, S. and Singh, A (2016). Rice PROTEIN L-ISOASPARTYL METHYLTRANSFERASE isoforms differentially accumulate during seed maturation to restrict deleterious isoAsp and reactive oxygen species accumulation and are implicated in seed vigor and longevity. New Phytol 211(2): 627–645. https://doi.org/10.1111/nph.13923
54. Prey, L., Hu, Y. and Schmidhalter, U (2020). High-throughput field phenotyping traits of grain yield formation and nitrogen use efficiency: optimizing the selection of vegetation indices and growth stages. Front Plant Sci 10: 1672. https://doi.org/10.3389/fpls.2019.01672
55. Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. and Glöckner, F.O (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41(D1): D590–D596. https://doi.org/10.1093/nar/gks1219
56. Samineni S, Sen M, Sajja SB, Gaur PM (2019). Rapid generation advance (RGA) in chickpea to produce up to seven generations per year and enable speed breeding. Crop J8(1): 164–169.
https://doi.org/10.1016/j.cj.2019.08.003
57. Samineni, S., Sen, M., Sajja, S.B. and Gaur, P.M (2020). Rapid generation advance (RGA) in chickpea to produce up to seven generations per year and enable speed breeding. Crop J 8(1): 164–169.
https://doi.org/10.1016/j.cj.2019.08.003
58. Saxena, K.B., Saxena, R.K., Hickey, L.T. and Varshney, R.K (2019). Can a speed breeding approach accelerate genetic gain in pigeonpea? Euphytica 215(12): 202. https://doi.org/10.1007/s10681-019-2520-4
59. Sheng X, Sun Z, Wang X, Tan Y, Yu D, Yuan G, Yuan D and Duan M (2020) Improvement of the Rice "Easy-to-Shatter" Trait via CRISPR/Cas9-Mediated Mutagenesis of the qSH1 Gene. Front Plant Sci 11:619. doi: 10.3389/fpls.2020.00619.
60. Sheng, X., Sun, Z., Wang, X., Tan, Y., Yu, D., Yuan, G., Yuan, D. and Duan, M (2020). Improvement of the rice "easy-to-shatter" trait via CRISPR/Cas9-mediated mutagenesis of the qSH1 gene. Front Plant Sci 11: 619. https://doi.org/10.3389/fpls.2020.00619
61. Singh, R.P. and Byerlee, D (1990). Relative variability in wheat yields across countries and over time. In: Evenson RE, Pray CE (eds) Research and Technology Policies for Agricultural Development. Oxford & IBH Publishing Co., New Delhi, pp 75–90.
62. Smith, J.S., Hussain, T., Jones, E.S., Graham, G., Podlich, D., Wall, S. and Williams, M (2008). Use of doubled haploids in maize breeding: implications for intellectual property protection and genetic diversity in hybrid crops. Mol Breed 22(1): 51–59. https://doi.org/10.1007/s11032-007-9155-1
63. Smolka M, Paulin LF, Grochowski CM, Horner DW, Mahmoud M, Behera S, Kalef-Ezra E, Gandhi M, Hong K, Pehlivan D (2024). Detection of mosaic and population-level structural variants with Sniffles2. Nat Biotechnol42(10):1571-1580.doi: 10.1038/s41587-023-02024-y
64. Spindel, J., Begum, H., Akdemir, D., Virk, P., Collard, B., Redoña, E., Atlin, G., Jannink, J.L. and McCouch, S.R (2015). Genome-wide prediction models provide high marker-based prediction accuracies for agronomic traits in breeding programs for nutrient-deficient environments in rice (Oryza sativa L.). PLoS Genet 11(2): e1004982. https://doi.org/10.1371/journal.pgen.1004982
65. Sun, J., Rutkoski, J.E., Poland, J.A., Crossa, J., Jannink, J.L. and Sorrells, M.E (2017). Multitrait, random regression, or simple repeatability model in high-throughput phenotyping data improve genomic prediction for wheat grain yield. Plant Genome 10(2): 1–15. https://doi.org/10.3835/plantgenome2016.11.0111
66. Tan, Y.C., Kumar, A.U., Wong, Y.P. and Ling, A.P.K (2022). Bioinformatics approaches and applications in plant biotechnology. J Genet Eng Biotechnol 20: 106. https://doi.org/10.1186/s43141-022-00394-5
67. Thomson, M.J (2014). High-throughput SNP genotyping to accelerate crop improvement. Plant Breed Biotechnol 2(3): 195–212. https://doi.org/10.9787/PBB.2014.2.3.195
68. Varshney, R.K., Pandey, M.K., Bohra, A., Singh, V.K., Thudi, M. and Saxena, R.K (2019). Toward the sequence-based breeding in legumes in the post-genome sequencing era. Theor Appl Genet 132(3): 797–816. https://doi.org/10.1007/s00122-018-3252-x
69. Vasudevan, S.N., Pooja, S.K., Raju, T.J. and Damini, C.S (2023). Cisgenics and intragenics: boon or bane for crop improvement. Front Plant Sci 14: 1275145. https://doi.org/10.3389/fpls.2023.1275145
70. Voichek Y, Weigel D (2020). Identifying genetic variants underlying phenotypic variation in plants without complete genomes. Nat Genet 52(5):534-540.doi: 10.1038/s41588-020-0612.
71. Wang, F., Sun, X., Liu, B., Kong, F., Pan, X. and Zhang, H (2022). A polygalacturonase gene PG031 regulates seed coat permeability with a pleiotropic effect on seed weight in soybean. Theor Appl Genet 135(5): 1603–1618. https://doi.org/10.1007/s00122-022-04057-4

72. Wanga, M.A., Shimelis, H., Mashilo, J. and Laing, M.D (2021). Opportunities and challenges of speed breeding: a review. Plant Breed 140(2): 185–194. https://doi.org/10.1111/pbr.12909
73. Watson, A., Ghosh, S., Williams, M.J., Cuddy, W.S., Simmonds, J., Rey, M.D., Hatta, M.A.M., Hinchliffe, A., Steed, A., Reynolds, D., Adamski, N.M., Breakspear, A., Korolev, A., Rayner, T., Dixon, L.E., Riaz, A., Martin, W., Ryan, M., Edwards, D., Batley, J., Raman, H., Carter, J., Rogers, C., Domoney, C., Moore, G., Harwood, W., Nicholson, P., Dieters, M.J., DeLacy, I.H., Zhou, J., Uauy, C., Boden, S.A., Park, R.F., Wulff, B.B.H. and Hickey, L.T (2018). Speed breeding is a powerful tool to accelerate crop research and breeding. Nat Plants 4(1): 23–29. https://doi.org/10.1038/s41477-017-0083-8
74. Xiong, B., Wang, B., Xiong, S., Lin, C. and Yuan, X (2019). 3D morphological processing for wheat spike phenotypes using computed tomography images. Remote Sens 11(9): 1110. https://doi.org/10.3390/rs11091110
75. Xu, Y., Liu, X., Fu, J., Wang, H., Wang, J., Huang, C., Prasanna, B.M., Olsen, M.S., Wang, G. and Zhang, A (2020). Enhancing genetic gain through genomic selection: from livestock to plants. Plant Commun 1(1): 100005. https://doi.org/10.1016/j.xplc.2019.100005
76. Yang, W., Feng, H., Zhang, X., Zhang, J., Doonan, J.H., Batchelor, W.D., Xiong, L. and Yan, J (2020). Crop phenomics and high-throughput phenotyping: past decades, current challenges and future perspectives. Mol Plant 13(2): 187–214. https://doi.org/10.1016/j.molp.2020.01.008
77. Yoosefzadeh-Najafabadi, M., Hesami, M. and Eskandari, M (2023). Machine learning-assisted approaches in modernized plant breeding programs. Genes 14(4): 777. https://doi.org/10.3390/genes1404777
78. Zhang, C., Marzougui, A. and Sankaran, S (2020). High-resolution satellite imagery applications in crop phenotyping: an overview. Comput Electron Agric 175: 105584. https://doi.org/10.1016/j.compag.2020.105584
79. Zhang, C., Marzougui, A. and Sankaran, S (2020). High-resolution satellite imagery applications in crop phenotyping: an overview. Comput Electron Agric 175: 105584. https://doi.org/10.1016/j.compag.2020.105584
80. Zhang, Y., Bhat, J.A., Zhang, Y. and Yang, S (2024). Understanding the molecular regulatory networks of seed size in soybean. Int J Mol Sci 25(3): 1441. https://doi.org/10.3390/ijms25031441
81. Zheng, Z., Wang, H.B., Chen, G.D., Yan, G.J. and Liu, C.J (2013). A procedure allowing up to eight generations of wheat and nine generations of barley per annum. Euphytica 191(2): 311–316. https://doi.org/10.1007/s10681-013-0909-z
82. Zhou, J., Reynolds, D., Le Cornu, T., Websdale, D., Orford, S., Lister, C., González-Navarro, O., Laycock, S., Finlayson, G., Stitt, T., Clark, M.D., Bevan, M.W. and Griffiths, S (2017). CropQuant: an automated and scalable field phenotyping platform for crop monitoring and trait measurements to facilitate breeding and digital agriculture. bioRxiv: 161547. https://doi.org/10.1101/161547
83. Zhou, Q., Lim, J.Q., Sung, W.K. and Li, G (2019). An integrated package for bisulfite DNA methylation data analysis with Indel-sensitive mapping. BMC Bioinformatics 20: 47. https://doi.org/10.1186/s12859-018-2593-4