

Review

Next-Generation Genotyping: Innovations Driving Plant Genomic Improvement

Valeriya Kostyukova ^{1,2} , Roza Kenzhebekova ^{1,3} , Egor Protsenko ^{1,3} , Bakyt Dulat ^{1,2} ,
Marina Khusnitdinova ^{1,2} and Dilyara Gritsenko ^{1,2,3,*} 

¹ Laboratory of Molecular Biology, Institute of Plant Biology and Biotechnology, Almaty 050040, Kazakhstan; rozakenzhebekova344@gmail.com (R.K.) germironame@gmail.com (M.K.)

² Research Center AgriBioTech, Almaty 050040, Kazakhstan

³ Faculty of Biology and Biotechnology, Al-Farabi Kazakh National University, Almaty 050040, Kazakhstan

* Correspondence: d.kopytina@gmail.com

Abstract

In recent years, plant genotyping has been shifting from the accumulation of whole-genome data toward their effective use in breeding programs. This review examines key genotyping platforms, including single-nucleotide polymorphism (SNP) arrays, reduced-representation sequencing methods such as genotyping-by-sequencing (GBS) and restriction site-associated DNA sequencing (RAD-seq), targeted genotyping approaches, and whole-genome sequencing (WGS), analyzing their informativeness, cost, and computational limitations. The transition to pangenome-based genotyping and graph genomes is discussed, as these approaches reduce reference bias and increase sensitivity for detecting structural variants, introgressions, and rare alleles that are important for adaptation and breeding. The growing role of AI/ML is highlighted in modeling complex genotype–phenotype relationships, integrating genomic and phenotypic data, and improving the accuracy and interpretability of genomic predictions.

Keywords: GBS; genomic selection (GS); genotyping; graph genomes; RAD-seq; SNP arrays; whole-genome sequencing (WGS)

1. Introduction

Plant genotyping is one of the key tools of modern plant genomics and breeding. It enables the identification of polymorphisms associated with productivity, disease resistance, and adaptation to stresses, as well as the establishment of causal relationships between DNA variation and phenotype. Such an approach significantly accelerates the selection of agriculturally valuable forms. In the context of climate change and the need to enhance food security, rapid, accurate, and scalable assessment of genetic diversity is becoming an especially important task [1]. Advances in sequencing methods and the reduction in their cost have made it possible to move from the analysis of individual loci to the study of variation at the whole-genome level. This, in turn, has opened new opportunities for identifying genes controlling quantitative traits and complex adaptive mechanisms.

Traditional molecular methods such as RFLP, AFLP, and SSR played a key role in the early stages of plant genetics; however, they have a number of limitations. The main drawback of these approaches is limited genome coverage. In addition, RFLP and AFLP methods can be technically demanding and often demonstrate lower reproducibility compared with modern genotyping approaches [2]. Moreover, traditional approaches are poorly suited to polyploid or large genomes, where sequence homology interferes with accurate allele



Academic Editors: Yuri Shavrukov,
Serik Eliby and Nikolai Borisjuk

Received: 23 February 2026

Revised: 18 March 2026

Accepted: 20 March 2026

Published: 21 March 2026

Copyright: © 2026 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

identification [3]. These limitations make traditional methods insufficiently informative for modern breeding tasks, which has driven the transition to next-generation sequencing (NGS) technologies.

High-density SNP arrays, genotyping-by-sequencing, and whole-genome resequencing have enabled comprehensive detection of genetic variants across the genome in wheat [4] and maize [5], substantially increasing marker density and the accuracy of association mapping. Bread wheat (*Triticum aestivum*), a hexaploid with a large and repeat-rich genome, serves as an excellent model for studying polyploid genomics and complex trait architecture. The transition to SNP-based genotyping has made it possible to identify more than 260 million polymorphisms in modern cultivars, landraces, and wild relatives, providing a level of resolution in genomic diversity that is unattainable with classical markers [6]. In soybean [7] and rapeseed [8], these platforms have facilitated the identification of loci underlying yield components and stress resistance, whereas in vegetatively propagated and polyploid crops such as potato [9], they have improved the resolution of genetic diversity analyses and increased the accuracy of genomic prediction.

The transition to NGS-based methods has become an important stage in the development of genomics. Modern high-throughput sequencing technologies, such as Illumina, PacBio, and Oxford Nanopore Technologies (ONT), make it possible to obtain millions of single-nucleotide polymorphisms (SNPs) and structural variants in a single experiment. Methods such as GBS, RAD-seq, ddRAD-seq, amplicon sequencing, targeted sequencing, and whole-genome sequencing (including low-coverage approaches) make it possible to obtain a sufficiently comprehensive and representative picture of the genome at moderate cost [10]. One of the key advances in modern genotyping has been the introduction and gradual transition to genomic selection (GS). The essence of this approach lies in the use of genome-wide markers to predict the phenotypic potential of plants [11]. It is particularly effective for quantitative traits controlled by many genes with small individual effects, for example, yield, resistance to abiotic stresses, and fruit quality [12–14]. GS models make it possible to simultaneously account for the effects of thousands of SNPs and structural variants, thereby increasing selection accuracy and reducing the duration of the breeding cycle [15]. In combination with NGS technologies, this opens up opportunities for faster selection of optimal genotypes in large populations and for high-throughput screening.

The importance of implementing next-generation genotyping is also confirmed by large international projects aimed at studying the genetic diversity of crops. For example, in the 3000 Rice Genomes Project [16], more than three thousand rice accessions were sequenced, which made it possible to identify approximately 19 million SNPs and to determine genes associated with yield and stress resistance. Similar results have been obtained for *Arabidopsis thaliana* [17], where genomic analysis helped reconstruct evolutionary relationships and identify loci involved in adaptation. These examples clearly demonstrate that next-generation genotyping not only deepens the understanding of the genetic basis of traits but also becomes a practical tool for crop improvement.

An important aspect of applying these technologies is their adaptation to different types of plant genome architecture. In this regard, the review examines the methodological principles of NGS-based approaches and their applicability to complex genomes, including polyploid, large, and repeat-rich genomes that pose significant challenges for analysis. Particular attention is given to the role of modern genotyping methods in assessing genetic diversity, conducting genome-wide association studies (GWAS), and implementing genomic selection strategies, which in recent years have become key tools for accelerating the breeding process.

2. SNP Arrays as a Tool for High-Throughput Plant Genotyping and Genomic Selection

With the introduction of high-throughput plant genotyping, approaches to studying genomes have changed, substantially accelerating the application of genomics in breeding. As a result, there has been a need for reliable and standardized methods that provide reproducible, highly accurate data on polymorphisms across the genomes of different crops. One of the most widely used solutions has been platforms based on the detection of previously known variants. SNP arrays (SNP chips) are high-throughput genotyping platforms based on DNA hybridization to preselected oligonucleotide probes specific to particular SNP positions [18]. The ability to detect tens of thousands to millions of SNPs across the genome in parallel has driven the widespread use of this method [19] for constructing genetic maps [20], GWAS [21], and population structure analysis [22]. Since the first SNP chips were introduced, marker panels have expanded considerably, covering thousands of new loci. As a result, high-density panels have been developed for major crops: 660 K markers for hexaploid wheat [20], 618 K for soybean [23], 480 K for apple [24], 700 K for rice [25], and 600 K for maize [26]. These panels provide high resolution for studying population genetic structure and conducting genomic selection.

A number of specialized platforms are used to read SNP-array data and perform large-scale plant genotyping. Platforms based on DNA microarray technologies and sequencing are most widely used for genotyping, as they provide high reproducibility and throughput. For the analysis of polymorphisms in plant genomes, microarray scanning systems (e.g., iScan, NextScan, HiScan) as well as automated stations for microarray hybridization and signal detection (e.g., Axiom arrays using the GeneTitan station) have been successfully applied [27–31]. As an alternative for lower-density SNP genotyping, PCR-based systems are used, such as the QuantStudio 12 K Flex Real-Time PCR System from Thermo Fisher Scientific (USA) [32]. Although the density of such panels is lower than that of conventional chips, these approaches provide flexibility in studies where the focus needs to be placed on a limited set of functional or QTL-linked markers.

The widespread implementation of SNP arrays has become a foundation for the transition from traditional breeding to integrated genomic strategies. The use of these technologies makes it possible to develop crop-wide panels, accelerate genomic selection, improve the accuracy of estimating the genetic value of lines, control genetic purity, and track introgressions from wild relatives. In addition, SNP chips have become the basis for building GWAS resources and QTL maps, which substantially expand the possibilities for analyzing functional genome variation [33]. A major limitation of the method is ascertainment bias, associated with the fact that SNPs are selected from a limited panel of founder genomes, which reduces analytical accuracy for other genetic groups or closely related species [34]. Because chips contain predefined polymorphisms, they capture only the variants included during array design, limiting the discovery of new or rare alleles in populations with high genetic diversity [35]. Moreover, high-density chips are costly and are intended for processing large numbers of samples, making them economically unjustified for small-scale studies.

Currently, we observe that large standard panels are no longer available in the open catalogs of the main companies that work with SNP arrays. At present, predominantly custom designs are offered. This observation raises several questions: how widely commercial SNP panels are used, whether there is a shift toward user-defined solutions, and which genotyping platform is most in demand today. In addition, it remains unclear whether the reduction in open catalogs reflects the discontinuation of large fixed arrays or a transition to more flexible, sequencing-oriented genotyping methods. Table 1 presents the number of

original research publications for the major agricultural crops. An extended version of the table with the full list of sources is provided in Supplementary File S1.

Table 1. Commercial SNP panels for different agricultural crops, with publication activity indicated for each.

Family	Species	Platform	SNP-Array Name	Average Number of Informative Markers	Number of Articles
Rosaceae	Apple (<i>Malus domestica</i> Borkh.)	Illumina	RosBREEDApple	$\bar{x}$ = 2476	25
		Axiom	Axiom Apple	$\bar{x}$ = 268,776	14
		Axiom	Axiom Apple (Axiom JKI50kMd)	27,108	1
	Pear (<i>Pyrus communis</i> L.)	Axiom	Axiom Pear	NA	1
	Rosa (<i>Rosa</i> spp.)	Axiom	Axiom Rose	$\bar{x}$ = 20,048	19
	Peach (<i>Prunus persica</i> (L.) Batsch)	Illumina	RosBREEDPeach	$\bar{x}$ = 3246	4
	Cherry (<i>Prunus</i> spp.)	Illumina	RosBREEDCherry	$\bar{x}$ = 1855	19
	Strawberry (<i>Fragaria</i> spp.)	Axiom	Axiom Strawberry 50 K (Axiom Fana_SNP)	$\bar{x}$ = 24,145	3
		Axiom	Axiom Strawberry (i35)	$\bar{x}$ = 11,769	22
		Axiom	Axiom Strawberry (iStraw90K)	$\bar{x}$ = 10,416	21
Poaceae	Maize (<i>Zea mays</i> L.)	Illumina	MaizeSNP50	$\bar{x}$ = 28,652	120
		Axiom	Axiom Maize6H (60 K)	$\bar{x}$ = 30,979	7
		Axiom	Axiom Maize	$\bar{x}$ = 429,498	19
	Wheat (<i>T. aestivum</i>)	Axiom	Axiom TaNG1.1	12,490	1
		Axiom	Axiom Wheat HD (Bristol)	546,299	2
		Illumina	90 K iSelect	$\bar{x}$ = 16,484	137
	Barley (<i>Hordeum vulgare</i> L.), Wheat (<i>T. aestivum</i>)	Illumina	Wheat-Barley40K	$\bar{x}$ = 16,617	9
	Rice (<i>Oryza sativa</i> L.)	Illumina	RiceLD	$\bar{x}$ = 981	3
		Illumina	RiceSNP50	$\bar{x}$ = 35,597	8
		Axiom	Axiom Rice	NA	0
Solanaceae	Rye (<i>Secale cereale</i> L.)	Axiom	Axiom Rye	NA	1
	Barley (<i>H. vulgare</i>)	Illumina	Barley50K Consortium Array (barley 50 K iSelect)	$\bar{x}$ = 26,869	41
	Tomato (<i>Solanum lycopersicum</i> L.)	Illumina	SolCAP Tomato 2013	$\bar{x}$ = 5967	26
	Pepper (<i>Capsicum annuum</i> L.)	Axiom	Axiom Tomato	$\bar{x}$ = 29,457	5
	Potato (<i>Solanum tuberosum</i> L.)	Illumina	TraitGenetics Pepper Consort.	NA	0
Fabaceae	Soybean (<i>Glycine max</i> (L.) Merr.)	Illumina	GGP Potato-24	$\bar{x}$ = 16,737	4
		Illumina	SolCAP 8303	$\bar{x}$ = 4333	28
		Illumina	BARCSoySNP6k	$\bar{x}$ = 3003	56
	Pea (<i>Pisum sativum</i> L.)	Axiom	Axiom Soybean	$\bar{x}$ = 51,758	47
	Lentil (<i>Lens culinaris</i> Medik.), Pea (<i>P. sativum</i>), Chickpea (<i>Cicer arietinum</i> L.), Lupin (<i>Lupinus</i> L.)	Illumina	GenoPea INRA 13.2 K	$\bar{x}$ = 9480	7
Malvaceae	Cotton (<i>Gossypium</i> spp.)	Illumina	Pulses Array	$\bar{x}$ = 13,216	3
		Axiom	Axiom Peanut (Arachis1, Arachis2)	$\bar{x}$ = 9089	39
		Illumina	CottonSNP63K	$\bar{x}$ = 13,273	33
		Illumina	CottonSNP80K	$\bar{x}$ = 43,901	21
		Axiom	Axiom Cotton	NA	0

NA—not available. $\bar{x}$ —mean value.

The absolute leader is the 90 K iSelect panel (141 publications), followed by MaizeSNP50 (123 publications) and BARCSoySNP6k (58 publications). Publication activity does not show a direct dependence on marker density. The most widely used panels are not necessarily the highest-density ones, but rather those that were standardized early and supported by major research consortia (e.g., 90 K iSelect, MaizeSNP50, BARCSoySNP6k). This confirms that the level of technology adoption is determined not only by technical characteristics, but also by the degree of integration into the scientific community.

Table 2 presents the number of custom SNP designs for major crops and the most common marker densities used in the panels. Supplementary File S2 provides a table of references.

Table 2. Custom SNP panel designs for main agricultural crops.

Species/Number of Markers	1–12 K	Average Number of Informative Markers	13 K–45 K	Average Number of Informative Markers	46 K–95 K	Average Number of Informative Markers	96 K+	Average Number of Informative Markers
Maize (<i>Z. mays</i>)	31	$\bar{x}$ = 978	3	$\bar{x}$ = 19,000	3	1653	9	$\bar{x}$ = 184,800
Rice (<i>O. sativa</i>)	19	$\bar{x}$ = 2051	3	NA	1	NA	0	NA
Wheat (<i>T. aestivum</i>)	2	NA	4	$\bar{x}$ = 16,888	1	NA	2	92,166
Apple (<i>M. domestica</i>)	2	$\bar{x}$ = 2832	2	$\bar{x}$ = 13,793	0	NA	0	NA
Cotton (<i>Gossypium</i> spp.)	0	NA	2	$\bar{x}$ = 21,898	0	NA	0	NA
Pepper (<i>C. annuum</i>)	1	27	3	$\bar{x}$ = 7313	0	NA	0	NA
Pear (<i>P. communis</i>)	4	$\bar{x}$ = 807	0	NA	1	66,616	2	166,335
Peach (<i>P. persica</i>)	3	$\bar{x}$ = 3015	0	NA	0	NA	0	NA
Barley (<i>H. vulgare</i>)	5	$\bar{x}$ = 1868	0	NA	0	NA	0	NA
Tomato (<i>S. lycopersicum</i>)	6	$\bar{x}$ = 2836	1	NA	0	NA	0	NA
Soybean (<i>G. max</i>)	4	$\bar{x}$ = 730	0	NA	1	47,337	1	128
Pea (<i>P. sativum</i>)	3	$\bar{x}$ = 606	0	NA	0	NA	0	NA

NA—not available. $\bar{x}$ —mean value.

An analysis of publications devoted to custom SNP panels demonstrates a pronounced predominance of low- and medium-density panels. The largest number of studies is associated with panels containing 1–12 thousand markers (94 publications) and 13–45 thousand SNPs (83 publications), whereas ultra-high-density panels (>95 thousand) appear in only 14 studies. This pattern of publication activity indicates that in practical breeding and applied genetics, moderate-density panels are predominantly used, as they provide an optimal balance between genotyping cost and informativeness. Unlike large consortium-based arrays (e.g., 90 K iSelect or MaizeSNP50), custom panels are more often developed for specific research objectives and do not always achieve broad international dissemination.

The dynamics of publication activity over time show a gradual increase in the number of studies starting in the late 2000s, peaking around 2018–2019, followed by fluctuations. The growth in publications during 2012–2019 reflects the active implementation of GWAS and genomic selection in breeding programs. The subsequent decline or instability may be associated with a gradual shift in some studies toward sequencing-based approaches (GBS, WGS), which provide a more flexible and scalable framework for variant discoveries.

However, it is important to understand that the transition to next-generation sequencing methods has not displaced SNP arrays; on the contrary, it has substantially simplified and accelerated their development. Massively parallel sequencing has enabled the discovery of millions of SNPs across a wide range of genotypes. As a result, SNP arrays have become more informative, more representative, and better tailored to specific crops and breeding objectives. The logic of these technologies lies in their complementarity: NGS is used at the stage of variant discovery and construction of reference panels, whereas SNP arrays provide highly reproducible, standardized, and cost-effective genotyping of large populations. This is particularly important for genomic selection, assessment of line purity, and long-term monitoring programs.

3. Reduced-Representation and Targeted Sequencing in Plant Genotyping

There has been significant interest in SNP markers, and the development of next-generation sequencing techniques has driven the rise of new high-throughput SNP-typing methods. The principle of RAD-seq is that before sequencing, the genome is digested by a single restriction enzyme at specific sites, producing a stable set of fragments for sequencing [36]. RAD-seq provides more predictable genome coverage, but requires more careful library preparation and often a higher per-sample cost [37]. RAD-seq involves digestion of the genome at specific sites using a single restriction enzyme. After restriction,

DNA is additionally fragmented (shearing) by mechanical methods, yielding fragments of different lengths around restriction sites. This is followed by size selection, generating a controlled library [36]. The presence of shearing and size-selection steps fundamentally distinguishes RAD-seq from other approaches [38]. A key limitation of RAD-seq applicability to large plant genomes is the loss of loci under selective pressure. Despite relatively even coverage around restriction sites, RAD-seq is more suitable for sequencing relatively simple genomes—with fewer repeats and a more uniform distribution of restriction sites—or for cases where fragment length control is important [39]. Standardized bioinformatic processing can also lead to the loss of important polymorphisms [40].

It should be noted that the initial RAD-seq protocols served as a starting point for the development of a number of derivative methods aimed at increasing data reproducibility, simplifying library preparation, and reducing the cost of analysis. In particular, the 2b-RAD approach was developed, based on the use of type IIB restriction enzymes, which excise short fragments of fixed length around restriction sites [41], allowing the formation of highly standardized libraries and reducing variability in fragment length between samples. The method was successfully tested on an F2 rice population (*O. sativa* L.), where an average of 2,000,332 high-quality reads per sample was obtained with high coverage uniformity [42]. In total, 3598 markers containing 3804 SNPs were identified, with a missing data rate of 18.9%.

Another example is ezRAD, which uses standard library preparation protocols applied on most high-throughput sequencing platforms [43]. The use of universal library preparation kits makes it possible to avoid specialized adapters and complex laboratory procedures characteristic of classical RAD-seq protocols. The ezRAD method was specifically designed for work with organisms that lack complete genomic resources.

Several modifications and derivatives of RAD-based approaches have been developed to improve reproducibility, scalability, and applicability across different plant systems; key variants are briefly outlined below. In addition, other modifications of RAD approaches have also become widespread. For example, double-digest RAD-seq (ddRAD) uses a combination of two restriction enzymes [44], which allows more precise control over the number of fragments obtained and increases the reproducibility of libraries between experiments [45]. The ddRAD method is characterized by high flexibility and adaptability to various objects and tasks. Jordon-Thaden et al. (2020) showed that the basic protocol with two restriction enzymes works effectively even with herbarium specimens and silica-gel-dried tissues in representatives of four plant genera [46].

Alongside RAD-derived methods, alternative reduced-representation sequencing approaches have been developed, among which genotyping-by-sequencing (GBS) represents a conceptually related but distinct strategy. The emergence and development of alternative genome complexity-reduction methods have been driven by the aim to improve the reproducibility and accuracy of genotyping. Genotyping-by-sequencing (GBS) is one such widely used method. The major benefit of GBS is the simultaneous use of reduced genome representation, NGS, and multiplexing that allows for its broad application to very large and complex genomes in practice at an affordable cost [47]. The approach relies on initial genome complexity reduction by one or more restriction enzymes [48] and subsequent sequencing of the generated fragments employing NGS platforms [49].

Since the first protocol published by Elshire et al. [49], GBS has become one of the most widely applied approaches for studying genetic diversity, mapping quantitative traits, and implementing genomic selection in crop plants [47,50]. One example of large-scale application of GBS is maize (*Z. mays*) research, where this method was used to genotype national and international collections of inbred lines comprising 2815 samples [51]. In that study, GBS was applied to analyze the genetic diversity of the U.S. collection, resulting in

the identification of 681,257 SNP markers. These data subsequently formed the basis for GWASs and genomic selection programs, confirming the suitability of GBS for working with large and genetically diverse panels [52,53].

The application of GBS in wheat (*T. aestivum*) is also distinctive when considering its hexaploid nature. Under the CIMMYT program, GBS has been employed for genotyping > 130,000 lines of bread wheat, resulting in 24,069 high-quality SNP markers distributed evenly across the 22 chromosomes [54]. A study of wheat and Barley (*H. vulgare*) genomes was among the initial demonstrations of the GBS approach in plant genomics. The genetic map was generated based on these data, and 34,000 SNPs for barley and over 30,000 SNPs for wheat were mapped; de novo maps without a reference genome were demonstrated [55]. Likewise, GBS has been utilized in a set of crop species such as sorghum [56], maize [57], and legumes [58,59], allowing the identification of polymorphisms associated with agriculturally important traits.

However, GBS is readily affected by errors caused by low read coverage and limited ability to detect true homozygotes [38], often resulting in a large number of missing data points and complicating the identification of true heterozygous genotypes [60]. Because of these challenges, proper genotype imputation and error correction during bioinformatic processing and downstream analyses are important [60]. Moreover, the efficiency of GBS largely depends on the quality of the reference genome [61]. The large genome size (16 Gb) of hexaploid wheat, which contains more than 85% duplicated sequences [62], also increases the likelihood of genotyping errors, because the genome includes a large number of paralogous and homeologous sequences [63]. Despite limitations associated with genome size, GBS remains one of the most cost-effective and scalable approaches for breeding programs.

Targeted next-generation sequencing occupies an intermediate position between whole-genome sequencing and reduced-representation methods and has become an important tool in applied plant genomics. Unlike GBS and RAD-seq, which target random subsets of the genome, targeted approaches focus from the outset on predefined DNA regions of greatest interest for breeding and functional genomics. This can substantially increase genotyping accuracy, simplify data interpretation, and make the method more applicable in practical breeding programs. Among targeted approaches, a fundamental distinction is made between amplicon-oriented methods, based on PCR amplification of preselected loci, and capture-based sequencing methods, which use sets of synthetic oligonucleotide probes for the selective enrichment of target genomic fragments. These strategies differ substantially both in their technical implementation and in the types of tasks for which they are optimal.

Amplicon panels are typically used for high-throughput genotyping of previously known SNP markers in breeding populations, where the analysis of a large number of samples at minimal cost is required [64,65]. Such studies show that amplicon-based approaches are particularly effective for monitoring previously identified functional variants associated with agronomically important traits.

In contrast, hybridization capture methods, such as Hyb-Seq, make it possible not only to genotype known markers but also to identify new polymorphisms within selected genomic regions, which makes them especially useful for studies of genetically diverse populations, phylogenomics, and analyses of gene structural variation [66–68]. An analysis of the literature shows that capture-based methods are actively used for sequencing hundreds and thousands of nuclear genes across a wide range of plants, including both model and wild species, and allow the recovery of comparable sets of orthologous sequences even at considerable phylogenetic distances between taxa [69–71].

Amplicon-based sequencing is a method in which predefined genomic regions are amplified by PCR, and the resulting amplicons are sequenced [72]. The method can use multiplex panels, enabling simultaneous analysis of dozens to hundreds of loci [73,74]. This approach differs from WGS in its focus: it requires less sequencing, reduces background higher depth, specifically in the regions of interest [75]. In plant breeding and genetic studies, this approach is used to target QTL regions [2,76], genes involved in resistance to biotic and abiotic stresses [64,75,77], and functionally important SNP and InDel markers [78]. This focus enables direct work with loci that have a proven association with traits of agronomic value, which is especially important for marker-assisted selection and validation of candidate genes. A key advantage of amplicon-based approaches is high genotyping accuracy, achieved through deep coverage of each target region and a minimal number of missing data points [64,79]. Unlike GBS, where low coverage can lead to errors in distinguishing homozygous and heterozygous genotypes, amplicon sequencing provides stable and reproducible detection of allelic variants even in complex populations [78]. This makes the method particularly valuable for routine analysis of breeding lines and cultivars.

Amplicon-based genotyping has proven to be a highly accurate and practically oriented approach in real crop improvement programs. For example, Yang et al. demonstrated in grapevine that GBS markers can be converted into multiplex amplicon panels for marker-assisted selection: with coverage depths on the order of hundreds of reads per locus, low missingness and high genotype reproducibility were achieved, enabling effective tracking of QTLs associated with resistance and quality traits even in highly heterozygous populations [64]. The development of this approach, in the form of rhAmpSeq, was demonstrated in another study [65], in which a panel of ~2000 amplicons designed to conserved “core genome” regions demonstrated high marker transferability across *Vitis* species and cost efficiency at high multiplexing. In soybean (*G. max*), the efficiency of highly multiplexed amplicon sequencing was demonstrated for targeted analysis of genes associated with phenological traits. High sequencing depth ensured low missingness and high genotype reproducibility [80]. In wheat, multiplex genotyping of markers linked to agronomically important traits has been widely applied using high-throughput sequencing-based approaches [81]. Amplicon-based approaches have also been used to identify induced mutations in predefined genes in the flax genome [82]. In crop improvement practice, amplicon sequencing should be viewed as an optimal tool for targeted and reproducible monitoring of known functionally important loci, but not as a universal replacement for broader genomic approaches.

The Hyb-Seq method combines hybridization capture of target regions with high-throughput sequencing and enables the simultaneous analysis of hundreds to thousands of genes [66]. The method includes designing probes (oligonucleotides) complementary to predefined genomic regions (e.g., genes of interest, conserved regions, or QTL-associated loci) [71]. DNA is then fragmented, and probe-bound DNA is identified and isolated. The resulting target fragments are subjected to massively parallel sequencing (Illumina) [83]. The method is applicable in phylogenomics, evolutionary and population studies, and plant breeding, where it is important to examine many genes across species and cultivars [66,71,84]. Another simplified approach is genome skimming. Genome skimming is an independent, low-depth sequencing method; however, within the Hyb-Seq framework, it is used as an additional data source, enabling the simultaneous acquisition of organellar genomes and ribosomal DNA sequences alongside targeted capture of nuclear genes [66]. It involves shallow sequencing of the genome to obtain information on high-copy sequences (chloroplast and mitochondrial genomes, ribosomal repeats) [85]. The method includes DNA library preparation, shotgun sequencing, and bioinformatic analysis. Shotgun sequencing breaks the genome into many random fragments that are sequenced in parallel.

This method does not require prior knowledge of genome structure. Using bioinformatic algorithms, reads are aligned and assembled into complete sequences [86]. Given the high copy number of target genes (thousands of copies per cell), a coverage depth of $1\text{--}5\times$ is sufficient [87–89].

In research devoted to genotyping and the analysis of genetic diversity, terminological confusion between different sequencing approaches is often observed. In particular, GBS and RAD-seq methods, which share a common conceptual basis (restriction-based genome reduction), are frequently used as interchangeable terms, despite substantial differences in protocols, library types, and data reproducibility [90]. Such blurring of terminology complicates comparisons between studies, assessment of method reproducibility, and correct interpretation of the resulting genomic data.

Taken together, the approaches discussed reflect the evolution of plant genotyping methods from large-scale screening of genomic variation to more precise and applied analytical strategies. The rational choice between GBS, RAD-seq, and targeted sequencing is determined by the balance between study scale, genome complexity, and specific breeding objectives, while their combined use allows genomic data to be integrated most effectively into modern crop improvement programs. In Figure 1, we present practical decision pathways for selecting appropriate genotyping technologies depending on the type of research objective, genome complexity, population characteristics, and available resources.

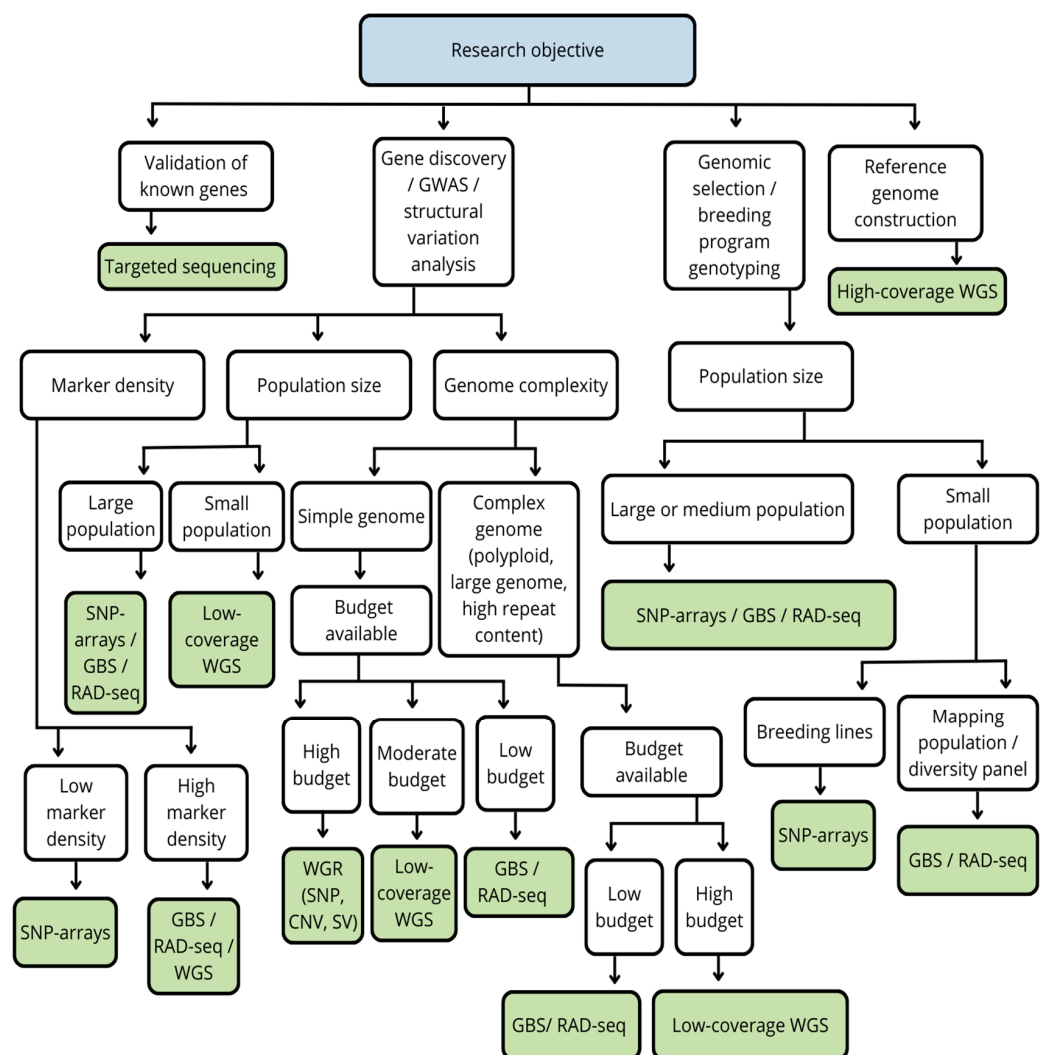


Figure 1. Practical decision tree for selecting appropriate genotyping platforms in plant genomics and breeding. This scheme integrates key experimental factors.

4. Whole-Genome Genotyping Strategies in Plant Breeding

Reduced-representation genome methods were developed as a cheaper alternative to whole genome sequencing (WGS) that enable the genotyping of hundreds of samples at orders of magnitude lower costs for sequencing and data analysis.

Whole-genome sequencing, owing to its complete genome coverage, enables analysis of various types of genetic variation, including SNPs [91,92], indels (InDels) [93], structural variants [94], and copy number variation (CNV) [95,96]. WGS can also be used for genomic mapping [97], evolutionary and population genomics [98], as well as functional genomics focusing on genes involved in stress resistance, productivity and other traits [91,99,100]. The benefits of WGS are the ability to capture the whole genome without certainty bias [101]. However, such an approach is severely restricted. Aside from the high cost, whole-genome sequencing generates a large set of data that needs a large amount of computational resources for storage and analysis [102]. In addition, accurate detection of low- and rare-variant heterozygotes is challenging and requires high sequencing depth, thereby increasing costs [103].

Whole-genome resequencing (WGR) is one of the most informative and versatile approaches to genotyping, allowing the analysis of genetic variability with maximum accuracy [98]. In applied genomics, WGR plays a key role, as it serves as the basis for the development of high-density marker panels, including SNP arrays and reference panels for subsequent genotype imputation [18,20,104]. WGR data are used to select informative and evenly distributed markers, which reduces systematic biases and ensures the comparability of genotypic data across different populations and breeding materials [10,77]. At the same time, despite its clear advantages, the widespread use of WGR in breeding programs is still limited by the high cost of sequencing, substantial computational resource requirements, and the complexity of storing and processing large volumes of data, especially for species with large and polyploid genomes.

In response to the limitations of classical WGR, a strong trend has emerged in recent years toward using low-coverage whole-genome sequencing (low-coverage WGS). In this approach, each sample is sequenced at a depth of approximately $0.5\text{--}5\times$, which substantially reduces analysis costs while retaining the ability to obtain genome-wide data [105]. Missing genotype information is reconstructed using high-quality reference panels generated from high-coverage sequencing reads [106]. This approach can be particularly effective for genomic selection tasks, where the key objective is not precise determination of every locus but accurate reconstruction of haplotype structure for phenotype prediction [107]. Low-coverage WGS with imputation demonstrates high predictive ability for quantitative traits [108] and can be considered a compromise between the marker density of WGR and the cost efficiency of reduced-representation methods. In addition, low-coverage WGR can be used to refine genotype identification and sample origin [109].

The development of long-read technologies, represented by the PacBio HiFi and Oxford Nanopore platforms, has opened a new stage in plant genotyping associated with systematic analysis of structural variation. Long-read sequencing enables direct detection of large-scale genome rearrangements, including long deletions and duplications, inversions, complex rearrangements, and copy number variation, which often remain inaccessible to standard NGS approaches [110–113]. These technologies are particularly important for haplotype genotyping, as read length allows reconstruction of contiguous genomic sequences without the need for complex statistical models [114]. They also facilitate analysis of epigenetic modifications (e.g., methylated bases) without bisulfite conversion. ONT can “read” methylation directly from native sequencing signals [115,116].

PacBio is a technology that generates long reads (typically $\sim 10\text{--}25$ kb) with very high accuracy ($\sim 99.8\%$) [117]. It produces so-called HiFi reads through repeated sequencing

of the same molecule (circular consensus sequencing, CCS). The method is based on preparing an SMRTbell library, followed by multiple passes over the same molecule to increase accuracy [118]. ONT is a platform capable of reading very long DNA fragments (sometimes hundreds of kilobases or more), i.e., ultra-long reads [119]. Limitations of these methods include increased requirements for the quality and quantity of input DNA: it must be high-molecular-weight and minimally damaged [85,90]. In addition, the need for substantial computational resources (storage, assembly, long-read alignment, handling large fast5/BAM files) makes these technologies expensive [120]. A brief comparison of the approaches discussed is presented in Figure 2.

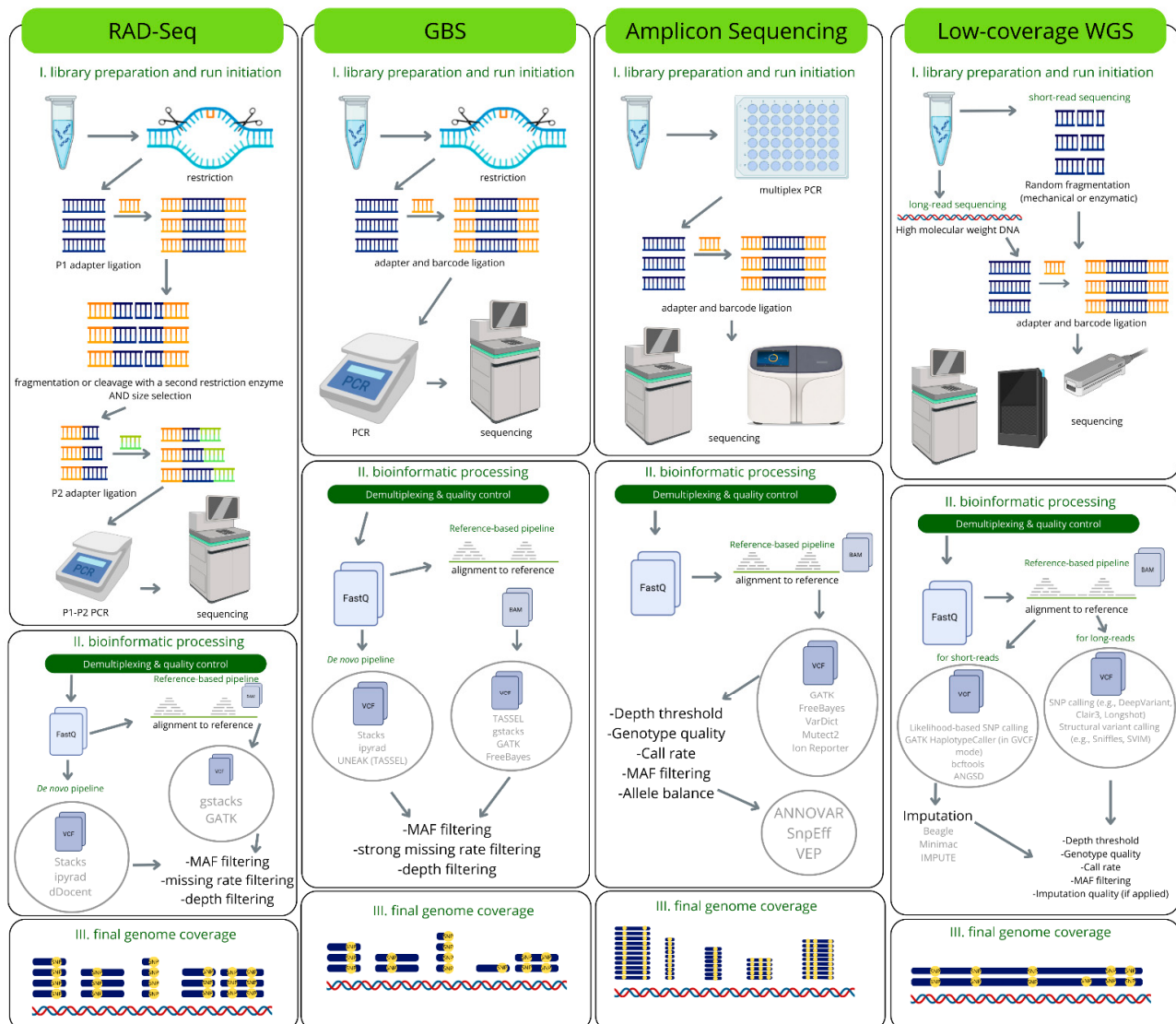


Figure 2. Differences in library preparation and analysis among the methods used in genomic selection.

5. Innovations in Plant Genotyping

In recent years, plant genotyping has been undergoing changes associated with a rethinking of the role of whole-genome sequencing. Although WGS has become the gold standard for detecting all types of genetic variation, its use in large-scale breeding programs remains limited in practice due to high cost and substantial computational burden. As a result, the focus is shifting from simply accumulating whole-genome data to the efficient use of genetic diversity, including structural variants, introgressions, and rare alleles.

The transition to pangenome-based genotyping reflects a fundamental shift in plant genomics—from the use of a single reference genome to representing species-wide genetic diversity as a pangenome that includes both “core” genes shared by all genotypes and “dispensable” or variable regions present only in a subset of lines [100,121–123]. Classical genotyping approaches based on a single reference are limited by reference bias: variants absent from the reference assembly are either not detected at all or are interpreted incorrectly [124]. The use of pangenomes substantially reduces this bias and enables a more complete description of the true genetic diversity of crop plants [125].

A key technical outcome of the pangenome approach has been graph genomes, in which alternative alleles, structural variants, and insertions are represented as branching paths in a graph rather than as a linear sequence [126]. Read alignment and variant calling in graph space increase sensitivity for detecting SNPs, InDels, and especially structural variation [127–129]. For crops with an intensive breeding history (wheat, maize, rice), this enables proper accounting for variants that have been lost in certain reference lines but retained in other cultivars or local populations [130–132]. For wheat, an interactive pangenome visualization system based on a graph model was developed, which allows the presence and absence of genes to be clearly displayed across sixteen bread wheat genomes [133]. Such visualization facilitates the comparison of cultivars, helping researchers more quickly identify structural variation, differences in gene content, and regions potentially important for adaptation and breeding.

Pangenome-based genotyping is of particular importance for wild relatives of crops, which often serve as sources of resistance to diseases, stresses, and extreme environmental conditions. When using a single cultivated reference, introgressed segments from wild species may be aligned incorrectly or not detected at all, which complicates their use in breeding [100,131,134]. Pangenome and graph-based approaches enable accurate identification of such regions.

Finally, methods of artificial intelligence and machine learning are playing an increasingly prominent role in data analysis. AI/ML algorithms are used to model complex nonlinear relationships between genotype and phenotype, integrate genomic, transcriptomic, and phenotypic data, and improve the accuracy of genomic predictions [135–137]. One of the key directions is phenotype prediction from genotype—models trained on marker data make it possible to predict both quantitative and qualitative plant traits, including yield, stress resistance, flowering time, and others [138–140]. In one study, artificial intelligence methods were applied to almond data; Random Forest models achieved a correlation of approximately 0.73 in predicting shelling fraction [141]. Another aspect is comparing the performance of different types of ML models: linear and regularized regression models, ensemble methods (Random Forest, Gradient Boosting), and deep learning models. Lourenço et al. evaluated models on real maize data and showed that in some cases simple regularized methods can compete with more complex ones, especially with a moderate number of markers and limited sample size, as complex models often require large datasets and substantial computational resources [142]. ML is applied in the context of “orphan crops”—plants for which genomic data are scarce—to transfer knowledge from well-studied crops. For example, MacNish et al. [143] showed that machine learning enables the transfer of knowledge from well-studied crops to understudied species, improving trait prediction and accelerating breeding under conditions of limited genomic data.

Despite the growing popularity of machine learning methods in genomic selection, their application is accompanied by a number of methodological limitations. One of the key problems is the so-called “curse of dimensionality,” which is typical for breeding problems where the number of markers (p) greatly exceeds the number of available samples (n) [144,145]. Under such conditions, complex machine learning models can easily overfit

and demonstrate high accuracy on training data but limited transferability to independent datasets [146]. Traditional statistical methods, including Bayesian genomic selection models (e.g., BayesA, BayesB, BayesC) and regularized regressions (ridge regression, LASSO), were originally developed to operate under conditions of $p \gg n$ and often show comparable or even more stable predictive performance [145,147].

In addition, many machine learning algorithms belong to the category of “black-box models,” which complicates the biological interpretation of results [148]. In the context of genetic research, it is important not only to predict the phenotype but also to understand which genetic factors underlie the predictions. Therefore, in recent years there has been growing interest in interpretable machine learning methods, such as SHAP (Shapley Additive Explanations), which allow the quantitative assessment of the contribution of individual markers to the model [149]. For example, in the study by Novielli et al., in addition to accurate prediction, SHAP methods were used to identify significant SNPs and genomic regions associated with the trait [141]. Nevertheless, even with the use of such tools, the question remains as to how well the obtained explanations reflect real biological mechanisms.

The speed of implementing genotyping into the breeding cycle and the speed of breeding decision-making depend more on the organization of the end-to-end pipeline [150] than on the choice of a specific genotyping platform per se: how cheaply and at what scale early generations can be genotyped [151], how quickly predictive models can be updated [152], and how genomic signals can be matched with high-frequency phenotypes [153]. An ultra-low-cost genotyping approach obtains extremely sparse coverage per sample—on the order of hundredths of the genome—and then reconstructs dense genotypes via imputation [154]. Its practical value for breeding is that genotyping becomes sufficiently inexpensive to enable continuous, high-throughput screening of thousands to tens of thousands of candidates per cycle. Using intermediate wheatgrass (*Thinopyrum intermedium* (Host) Barkworth & Dewey) as an example, a pipeline has been presented that is directly oriented toward breeding-program operation: ultra-low coverage skim-seq (approximately $0.01 \times$ – $0.05 \times$), imputation using STITCH, and subsequent application of these data in genomic selection to predict breeding values in the crop [155].

The emergence of practical haplotype graphs has simplified the reuse of haplotype information from a reference panel within a breeding program, primarily for imputation from low-coverage data. For sorghum, a Practical Haplotype Graph (PHG) has been proposed, demonstrating that even at extremely low coverage of about $0.01 \times$, imputed genotypes suitable for breeding applications can be obtained [156]. In work on creating the Wheat PHG database, it was shown that imputation accuracy can remain high even at very low coverage of about $0.01 \times$ (~92% accuracy) [157]. As a result, ultra-low-cost genotyping ceases to be merely an auxiliary tool: genotypes are updated rapidly and at scale, GS models are rebuilt more frequently, and expensive phenotyping can be concentrated on a smaller number of candidates.

Single-cell DNA genotyping in plants remains a niche and methodologically challenging endeavor, whereas single-cell/single-nucleus transcriptomics and other single-cell “omics” are already being applied to dissect complex traits by cell type/trajectory and to build mechanistic links between genetic variation and phenotype [158,159]. An example of the move toward reproducible protocol solutions is a detailed single-cell ATAC-seq protocol for nuclei from maize (*Z. mays*) seedlings, which enables the generation of cell-type-specific chromatin accessibility maps in an agricultural crop [160]. A time-resolved single-cell transcriptomic analysis of Arabidopsis embryo germination has been presented, showing how cell states are formed, and regulatory programs shift as growth is initiated [161]. Single-cell data enable analysis of how genetic variants alter gene regulation and expression in specific

cell types and states, helping to explain and predict complex traits that depend on limited tissue and cellular contexts [162,163]. In practical breeding, the near-term role of single-cell approaches is to refine which tissues, stages, and regulatory nodes should be measured and prioritized to improve the transferability of predictions [164].

Integration with phenomics and multi-omics is a direction that makes genomic models more accurate and transferable across environments. For most agronomically important traits, strong dependence on growing conditions and temporal dynamics is typical; therefore, one-time measurements at a single developmental stage often do not reflect the true dynamics of phenotype formation [165]. High-throughput phenomics (HTP) addresses this issue through repeated, standardized observations throughout the growing season [166,167]. Additional gains in informativeness are achieved through multi-omics approaches (transcriptomics, epigenomics/chromatin accessibility, metabolomics), which serve as intermediate layers between genotype and phenotype [168,169]. Such data can act as additional predictors and simultaneously explain why the same genetic background manifests differently under changing environments. Taken together, combining large-scale genotyping with omics approaches forms the basis for predictive breeding of dynamic traits, where models are regularly updated as more data accumulate.

6. Genotyping of Complex Genomes and Its Importance in GWAS and Genomic Selection

Genotyping crops with polyploid and/or highly repetitive genomes remains one of the most technically challenging areas of applied plant genomics. For species such as hexaploid wheat (*T. aestivum*), tetraploid cultivated potato (*S. tuberosum*), and allopolyploid rapeseed (*Brassica napus* L.), key constraints are associated with multiple homeologous copies of loci, a high proportion of repeats and paralogs [63,170–172]. This complicates both the design of marker panels and the interpretation of NGS data, increasing the risk of erroneous variant calling [173]. Unlike diploids, where genotypes are typically encoded as AA/AB/BB, in tetraploids and hexaploids it is necessary to estimate allele dosage (e.g., for a tetraploid, classes from AAAA to BBBB) [9]. Dosage errors lead to systematic bias in marker-effect estimates, reduced GWAS power, and decreased genomic selection accuracy [174].

The most common strategy for polyploids remains SNP arrays with pre-optimized marker sets that minimize cross-hybridization and ambiguous assignment to homeologs. For example, the Affymetrix Wheat660K array has been used to construct a high-density genetic map in hexaploid wheat, where the large SNP panel provided sufficient genome coverage for high-resolution mapping and subsequent genetic studies [20]. For rapeseed (*B. napus*), similar roles have been played by ~60 K arrays designed to select predominantly single-locus markers in an allotetraploid genome [27]. In particular, the Brassica 60K Infinium SNP array has been described as a universal platform for linkage mapping, association analysis, diversity assessment, and introgression detection, with particular attention to which SNPs are truly “locus-specific” and do not amplify multiple homeologous/paralogous regions simultaneously.

Allele dosage determination in polyploids is usually based on the analysis of allele signal ratios (for example, SNP array intensities or the proportion of reads for the alternative allele in NGS data). Different software packages use different statistical approaches to address this task. For example, polyRAD implements a Bayesian model that estimates posterior genotype probabilities taking into account sequencing depth, sequencing errors, and population structure [175].

The packages updog [176] and fitPoly [177] use models of allele signal distributions to estimate dosage based on the probabilistic distribution of intensities or allele frequencies, which makes it possible to account for genotyping uncertainty. Other tools, such as

SuperMASSA [178] or MAPpoly [179], apply probabilistic and EM algorithms for the joint estimation of dosage and segregation parameters in mapping populations.

The transition from rigid discrete genotype calls to probabilistic representations makes it possible to account more appropriately for measurement uncertainty and reduces the likelihood of systematic errors in the analysis of polyploid genomes [180].

Classical single-locus GWAS often underestimates traits with a polygenic architecture and many small-effect loci [181,182]. Multi-locus GWAS (ML-GWAS) models simultaneously account for multiple loci and thereby increase power to detect signals, especially in breeding panels of moderate size [183]. In applied genetics, GWAS results are increasingly interpreted not as single “signals,” but as sets of associated genomic regions, which are then used to refine candidate genes and for functional validation [184]. The key methodological feature of genomic selection (GS) is the use of the entire marker set to predict genomic estimated breeding values (GEBVs), rather than relying on a small number of significant markers. This enables work with traits controlled by many small-effect loci and shifts the emphasis from causal interpretation to predictive accuracy and robustness [185,186]. Therefore, the use of probabilistic genotyping methods and the explicit consideration of genotype uncertainty are regarded as an important step toward improving the reliability of association studies and breeding predictions in polyploid crops.

The genotyping methods discussed in this work demonstrate a consistent evolution in plant genomics—from standardized, high-throughput screening of predefined polymorphisms to more flexible and informative strategies for analyzing genomic variation. SNP arrays remain a key tool for high-throughput genotyping due to their high reproducibility, data standardization, and cost efficiency when working with large populations. Their importance is especially high in genomic selection, monitoring the genetic purity of lines, and long-term breeding programs, where data consistency across experiments and generations is critical (Figure 3).

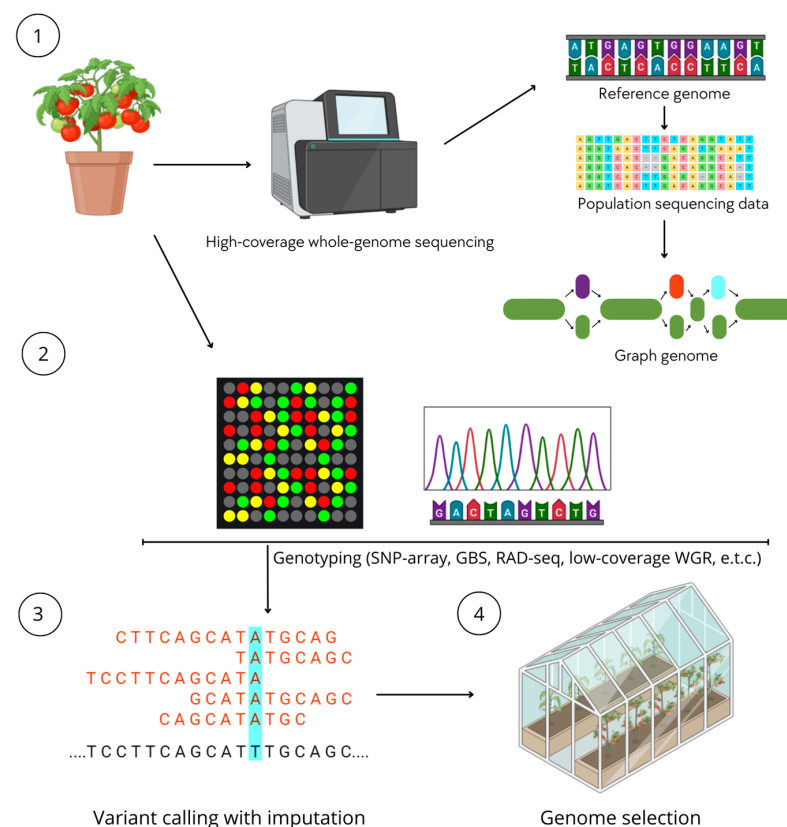


Figure 3. Conceptual diagram of genotyping and its use in genomic selection.

At the same time, reduced-representation and targeted sequencing methods (GBS, RAD-seq, targeted sequencing) have substantially expanded the possibilities for analyzing genetic diversity by enabling work with novel, rare, and population-specific variants that are not accessible with SNP chips. Whole-genome strategies, including WGR, low-coverage WGS, and long-read sequencing, offer the highest resolution for analyzing genetic structure, structural variants, and haplotypes, but are constrained by cost and computational requirements. Their use is particularly justified at the stages of variant discovery, QTL mapping, and reference panel construction.

7. What Is Next?

Marker-assisted selection (MAS), which dominated the early stages of molecular plant breeding, focused on using a limited number of loci with large effects. However, these approaches proved to be limited when dealing with quantitative traits controlled by thousands of genetic factors with small effects [185]. Modern plant genotyping is entering a phase in which the key driver of progress is no longer further increases in marker density or refinement of individual platforms, but rather the integration of genomic data into a continuous breeding process. In this context, the concept of genomic estimated breeding value is becoming central as an operational metric. The development of ultra-low-cost approaches enables the analysis of tens of thousands of candidates in each breeding cycle. Under these conditions, GEBV ceases to be a one-time estimate and becomes a dynamic value that is regularly updated as new genotypic and phenotypic data accumulates.

Another important direction of development is the shift from SNP-centric representations of the genome toward haplotype-based and pangenomic models. Pangenomic and graph-based genome representations reduce reference bias and enable proper consideration of structural variants, introgressions from wild relatives, and presence/absence variations, which play a significant role in adaptive traits. In the longer term, this is expected to lead to the development of haplotype-oriented GEBVs that are more robust to transfer across populations and breeding programs.

The limitation of classical GEBV models remains the strong dependence of many agronomically important traits on environmental conditions and plant developmental dynamics [15]. In this context, a key direction for future research is the integration of genomic predictions with high-throughput phenomics and multi-omics data [187,188]. Incorporating phenotypic, transcriptomic, and epigenetic data makes it possible to interpret GEBV as a prediction of a genotype's response to specific growing conditions. Such models are particularly promising for breeding for abiotic stress tolerance under changing climatic conditions [189–191].

Machine learning and artificial intelligence methods are increasingly prominent in GEBV construction, especially for analyzing complex nonlinear relationships and genotype $\times$ environment interactions. It has been shown that joint modeling of multiple traits and environments improves predictive accuracy compared with single-trait approaches, particularly for quantitative traits with strong environmental dependence [191,192]. Using maize and wheat as examples, both classical Bayesian multi-trait models and deep learning-based approaches have been shown to effectively exploit correlations among traits and environments to improve estimation of genomic breeding values [193]. At the same time, neural network models achieve comparable accuracy with lower computational costs, whereas statistical models retain an advantage when explicitly modeling genotype $\times$ environment interactions. Similar results have been reported in other studies on cereal crops [194], indicating a general trend toward a transition from simple GEBVs to complex multivariate predictive models. However, accumulated experience shows that the advantages of highly complex, poorly interpretable models become evident only when large, well-balanced training datasets are

available, which are limited for polyploid, cross-pollinating species [195]. Consequently, there is growing interest in interpretable ML approaches that improve prediction accuracy while identifying gene regions and haplotypes that contribute most to trait variation, thereby bridging predictive breeding and functional genomics.

8. Conclusions

Modern approaches to plant genotyping are not competing technologies but form a complementary ecosystem of tools optimized for different stages of the breeding process. SNP arrays provide a reliable foundation for large-scale, standardized analyses; reduced-representation and targeted sequencing methods provide flexibility and access to novel variation, and whole-genome and pangenome strategies provide depth and completeness in representing genetic diversity. A rational combination of these approaches enables effective integration of genomic data into practical breeding.

The future of plant genotyping is not so much about the universal adoption of a single method as about the development of integrated strategies that combine high-quality genotyping data, pangenome models, and artificial intelligence algorithms. This approach opens opportunities for more accurate phenotype prediction, acceleration of breeding cycles, and fuller utilization of the genetic potential of both cultivated plants and their wild relatives, which is of key importance under climate change and increasing demands for agroecosystem resilience.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/life16030521/s1>. Table S1—Table listing publications using selected commercial SNP panels. Table S2—Table listing publications using selected custom SNP panels.

Author Contributions: Conceptualization, D.G.; writing—original draft preparation, M.K., V.K., R.K. and E.P.; writing—review and editing, D.G.; visualization, V.K. and B.D.; supervision, D.G.; project administration, D.G.; funding acquisition, D.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Agriculture of the Republic of Kazakhstan within the framework of a targeted funding program BR22887230.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

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

References

- Ćeran, M.; Miladinović, D.; Đorđević, V.; Trkulja, D.; Radanović, A.; Glogovac, S.; Kondić-Špika, A. Genomics-Assisted Speed Breeding for Crop Improvement: Present and Future. *Front. Sustain. Food Syst.* **2024**, *8*, 1383302. [CrossRef]
- Kumar, R.; Das, S.P.; Choudhury, B.U.; Kumar, A.; Prakash, N.R.; Verma, R.; Chakraborti, M.; Devi, A.G.; Bhattacharjee, B.; Das, R.; et al. Advances in Genomic Tools for Plant Breeding: Harnessing DNA Molecular Markers, Genomic Selection, and Genome Editing. *Biol. Res.* **2024**, *57*, 80. [CrossRef] [PubMed]
- Jeon, D.; Kim, C. Polyploids of Brassicaceae: Genomic Insights and Assembly Strategies. *Plants* **2024**, *13*, 2087. [CrossRef] [PubMed]
- Winfield, M.O.; Allen, A.M.; Burridge, A.J.; Barker, G.L.A.; Benbow, H.R.; Wilkinson, P.A.; Coghill, J.; Waterfall, C.; Davassi, A.; Scopes, G.; et al. High-Density SNP Genotyping Array for Hexaploid Wheat and Its Secondary and Tertiary Gene Pool. *Plant Biotechnol. J.* **2016**, *14*, 1195–1206. [CrossRef]
- Ganal, M.W.; Durstewitz, G.; Polley, A.; Bérard, A.; Buckler, E.S.; Charcosset, A.; Clarke, J.D.; Graner, E.-M.; Hansen, M.; Joets, J.; et al. A Large Maize (*Zea mays* L.) SNP Genotyping Array: Development and Germplasm Genotyping, and Genetic Mapping to Compare with the B73 Reference Genome. *PLoS ONE* **2011**, *6*, e28334. [CrossRef]

6. Sun, C.; Dong, Z.; Zhao, L.; Ren, Y.; Zhang, N.; Chen, F. The Wheat 660K SNP Array Demonstrates Great Potential for Marker-Assisted Selection in Polyploid Wheat. *Plant Biotechnol. J.* **2020**, *18*, 1354–1360. [[CrossRef](#)]
7. Doszhanova, B.; Zatybekov, A.; Didorenko, S.; Fang, C.; Abugalieva, S.; Turuspekoy, Y. Genome-Wide Association Study of Seed Quality and Yield Traits in a Soybean Collection from Southeast Kazakhstan. *Agronomy* **2024**, *14*, 2746. [[CrossRef](#)]
8. Lu, K.; Wei, L.; Li, X.; Wang, Y.; Wu, J.; Liu, M.; Zhang, C.; Chen, Z.; Xiao, Z.; Jian, H.; et al. Whole-Genome Resequencing Reveals Brassica Napus Origin and Genetic Loci Involved in Its Improvement. *Nat. Commun.* **2019**, *10*, 1154. [[CrossRef](#)]
9. Uitdewilligen, J.G.A.M.L.; Wolters, A.-M.A.; D'hoop, B.B.; Borm, T.J.A.; Visser, R.G.F.; van Eck, H.J. A Next-Generation Sequencing Method for Genotyping-by-Sequencing of Highly Heterozygous Autotetraploid Potato. *PLoS ONE* **2013**, *8*, e62355. [[CrossRef](#)]
10. Bhat, J.A.; Ali, S.; Salgotra, R.K.; Mir, Z.A.; Dutta, S.; Jadon, V.; Tyagi, A.; Mushtaq, M.; Jain, N.; Singh, P.K.; et al. Genomic Selection in the Era of Next Generation Sequencing for Complex Traits in Plant Breeding. *Front. Genet.* **2016**, *7*, 221. [[CrossRef](#)]
11. Jannink, J.-L.; Lorenz, A.J.; Iwata, H. Genomic Selection in Plant Breeding: From Theory to Practice. *Brief. Funct. Genom.* **2010**, *9*, 166–177. [[CrossRef](#)] [[PubMed](#)]
12. Alemu, A.; Åstrand, J.; Montesinos-López, O.A.; Isidro y Sánchez, J.; Fernández-González, J.; Tadesse, W.; Vetukuri, R.R.; Carlsson, A.S.; Ceplitis, A.; Crossa, J.; et al. Genomic Selection in Plant Breeding: Key Factors Shaping Two Decades of Progress. *Mol. Plant* **2024**, *17*, 552–578. [[CrossRef](#)]
13. Stewart-Brown, B.B.; Song, Q.; Vaughn, J.N.; Li, Z. Genomic Selection for Yield and Seed Composition Traits Within an Applied Soybean Breeding Program. *G3 Genes | Genomes | Genet.* **2019**, *9*, 2253–2265. [[CrossRef](#)]
14. Yamamoto, E.; Matsunaga, H.; Onogi, A.; Ohshima, A.; Miyatake, K.; Yamaguchi, H.; Nunome, T.; Iwata, H.; Fukuoka, H. Efficiency of Genomic Selection for Breeding Population Design and Phenotype Prediction in Tomato. *Heredity* **2017**, *118*, 202–209. [[CrossRef](#)] [[PubMed](#)]
15. Crossa, J.; Pérez-Rodríguez, P.; Cuevas, J.; Montesinos-López, O.; Jarquín, D.; de los Campos, G.; Burgueño, J.; González-Camacho, J.M.; Pérez-Elizalde, S.; Beyene, Y.; et al. Genomic Selection in Plant Breeding: Methods, Models, and Perspectives. *Trends Plant Sci.* **2017**, *22*, 961–975. [[CrossRef](#)]
16. The 3000 Rice Genomes Project. The 3000 Rice Genomes Project. *GigaSci* **2014**, *3*, 2047–217X-3-7. [[CrossRef](#)]
17. The 1001 Genomes Consortium. 1135 Genomes Reveal the Global Pattern of Polymorphism in *Arabidopsis thaliana*. *Cell* **2016**, *166*, 481–491. [[CrossRef](#)]
18. Gunderson, K.L.; Steemers, F.J.; Lee, G.; Mendoza, L.G.; Chee, M.S. A Genome-Wide Scalable SNP Genotyping Assay Using Microarray Technology. *Nat. Genet.* **2005**, *37*, 549–554. [[CrossRef](#)] [[PubMed](#)]
19. You, Q.; Yang, X.; Peng, Z.; Xu, L.; Wang, J. Development and Applications of a High Throughput Genotyping Tool for Polyploid Crops: Single Nucleotide Polymorphism (SNP) Array. *Front. Plant Sci.* **2018**, *9*, 104. [[CrossRef](#)]
20. Cui, F.; Zhang, N.; Fan, X.; Zhang, W.; Zhao, C.; Yang, L.; Pan, R.; Chen, M.; Han, J.; Zhao, X.; et al. Utilization of a Wheat660K SNP Array-Derived High-Density Genetic Map for High-Resolution Mapping of a Major QTL for Kernel Number. *Sci. Rep.* **2017**, *7*, 3788. [[CrossRef](#)]
21. Kim, K.-W.; Nawade, B.; Nam, J.; Chu, S.-H.; Ha, J.; Park, Y.-J. Development of an Inclusive 580K SNP Array and Its Application for Genomic Selection and Genome-Wide Association Studies in Rice. *Front. Plant Sci.* **2022**, *13*, 1036177. [[CrossRef](#)]
22. Wu, X.; Li, Y.; Shi, Y.; Song, Y.; Wang, T.; Huang, Y.; Li, Y. Fine Genetic Characterization of Elite Maize Germplasm Using High-Throughput SNP Genotyping. *Theor. Appl. Genet.* **2014**, *127*, 621–631. [[CrossRef](#)] [[PubMed](#)]
23. Li, Y.-F.; Li, Y.-H.; Su, S.-S.; Reif, J.C.; Qi, Z.-M.; Wang, X.-B.; Wang, X.; Tian, Y.; Li, D.-L.; Sun, R.-J.; et al. SoySNP618K Array: A High-Resolution Single Nucleotide Polymorphism Platform as a Valuable Genomic Resource for Soybean Genetics and Breeding. *J. Integr. Plant Biol.* **2022**, *64*, 632–648. [[CrossRef](#)]
24. Bianco, L.; Cestaro, A.; Linsmith, G.; Muranty, H.; Denancé, C.; Théron, A.; Poncet, C.; Micheletti, D.; Kerschbamer, E.; Di Pierro, E.A.; et al. Development and Validation of the Axiom® Apple480K SNP Genotyping Array. *Plant J.* **2016**, *86*, 62–74. [[CrossRef](#)] [[PubMed](#)]
25. McCouch, S.R.; Wright, M.H.; Tung, C.-W.; Maron, L.G.; McNally, K.L.; Fitzgerald, M.; Singh, N.; DeClerck, G.; Agosto-Perez, F.; Korniliev, P.; et al. Open Access Resources for Genome-Wide Association Mapping in Rice. *Nat. Commun.* **2016**, *7*, 10532. [[CrossRef](#)]
26. Unterseer, S.; Bauer, E.; Haberer, G.; Seidel, M.; Knaak, C.; Ouzunova, M.; Meitinger, T.; Strom, T.M.; Fries, R.; Pausch, H.; et al. A Powerful Tool for Genome Analysis in Maize: Development and Evaluation of the High Density 600 k SNP Genotyping Array. *BMC Genom.* **2014**, *15*, 823. [[CrossRef](#)]
27. Clarke, W.E.; Higgins, E.E.; Plieske, J.; Wieseke, R.; Sidebottom, C.; Khedikar, Y.; Batley, J.; Edwards, D.; Meng, J.; Li, R.; et al. A 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.* **2016**, *129*, 1887–1899. [[CrossRef](#)]
28. Marcotuli, I.; Gadaleta, A.; Mangini, G.; Signorile, A.M.; Zacheo, S.A.; Blanco, A.; Simeone, R.; Colasuonno, P. Development of a High-Density SNP-Based Linkage Map and Detection of QTL for β -Glucans, Protein Content, Grain Yield per Spike and Heading Time in Durum Wheat. *Int. J. Mol. Sci.* **2017**, *18*, 1329. [[CrossRef](#)]

29. Seo, J.; Lee, S.-M.; Han, J.-H.; Shin, N.-H.; Lee, Y.K.; Kim, B.; Chin, J.H.; Koh, H.-J. Characterization of the Common Japonica-Originated Genomic Regions in the High-Yielding Varieties Developed from Inter-Subspecific Crosses in Temperate Rice (*Oryza sativa* L.). *Genes* **2020**, *11*, 562. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Hiraoka, Y.; Ferrante, S.P.; Wu, G.A.; Federici, C.T.; Roose, M.L. Development and Assessment of SNP Genotyping Arrays for Citrus and Its Close Relatives. *Plants* **2024**, *13*, 691. [\[CrossRef\]](#)
31. Singh, S.; Mahato, A.K.; Jayaswal, P.K.; Singh, N.; Dheer, M.; Goel, P.; Raje, R.S.; Yasin, J.K.; Sreevathsa, R.; Rai, V.; et al. A 62K Genic-SNP Chip Array for Genetic Studies and Breeding Applications in Pigeonpea (*Cajanus cajan* L. Millsp.). *Sci. Rep.* **2020**, *10*, 4960. [\[CrossRef\]](#)
32. Broccanello, C.; Chiodi, C.; Funk, A.; McGrath, J.M.; Panella, L.; Stevanato, P. Comparison of Three PCR-Based Assays for SNP Genotyping in Plants. *Plant Methods* **2018**, *14*, 28. [\[CrossRef\]](#)
33. Negro, S.S.; Millet, E.J.; Madur, D.; Bauland, C.; Combes, V.; Welcker, C.; Tardieu, F.; Charcosset, A.; Nicolas, S.D. Genotyping-by-Sequencing and SNP-Arrays Are Complementary for Detecting Quantitative Trait Loci by Tagging Different Haplotypes in Association Studies. *BMC Plant Biol.* **2019**, *19*, 318. [\[CrossRef\]](#)
34. Lachance, J.; Tishkoff, S.A. SNP Ascertainment Bias in Population Genetic Analyses: Why It Is Important, and How to Correct It. *BioEssays* **2013**, *35*, 780–786. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Geibel, J.; Reimer, C.; Weigend, S.; Weigend, A.; Pook, T.; Simianer, H. How Array Design Creates SNP Ascertainment Bias. *PLoS ONE* **2021**, *16*, e0245178. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Davey, J.W.; Blaxter, M.L. RADSeq: Next-Generation Population Genetics. *Brief. Funct. Genom.* **2010**, *9*, 416–423. [\[CrossRef\]](#)
37. Wickland, D.P.; Battu, G.; Hudson, K.A.; Diers, B.W.; Hudson, M.E. A Comparison of Genotyping-by-Sequencing Analysis Methods on Low-Coverage Crop Datasets Shows Advantages of a New Workflow, GB-eaSy. *BMC Bioinform.* **2017**, *18*, 586. [\[CrossRef\]](#)
38. Reyes, V.P.; Kitony, J.K.; Nishiuchi, S.; Makihara, D.; Doi, K. Utilization of Genotyping-by-Sequencing (GBS) for Rice Pre-Breeding and Improvement: A Review. *Life* **2022**, *12*, 1752. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Lowry, D.B.; Hoban, S.; Kelley, J.L.; Lotterhos, K.E.; Reed, L.K.; Antolin, M.F.; Storfer, A. Breaking RAD: An Evaluation of the Utility of Restriction Site-Associated DNA Sequencing for Genome Scans of Adaptation. *Mol. Ecol. Resour.* **2017**, *17*, 142–152. [\[CrossRef\]](#)
40. Díaz-Arce, N.; Rodríguez-Ezpeleta, N. Selecting RAD-Seq Data Analysis Parameters for Population Genetics: The More the Better? *Front. Genet.* **2019**, *10*, 533. [\[CrossRef\]](#)
41. Wang, S.; Meyer, E.; McKay, J.K.; Matz, M.V. 2b-RAD: A Simple and Flexible Method for Genome-Wide Genotyping. *Nat. Methods* **2012**, *9*, 808–810. [\[CrossRef\]](#)
42. Guo, Y.; Yuan, H.; Fang, D.; Song, L.; Liu, Y.; Liu, Y.; Wu, L.; Yu, J.; Li, Z.; Xu, X.; et al. An Improved 2b-RAD Approach (I2b-RAD) Offering Genotyping Tested by a Rice (*Oryza sativa* L.) F2 Population. *BMC Genom.* **2014**, *15*, 956. [\[CrossRef\]](#)
43. Toonen, R.J.; Puritz, J.B.; Forsman, Z.H.; Whitney, J.L.; Fernandez-Silva, I.; Andrews, K.R.; Bird, C.E. ezRAD: A Simplified Method for Genomic Genotyping in Non-Model Organisms. *PeerJ* **2013**, *1*, e203. [\[CrossRef\]](#)
44. Peterson, B.K.; Weber, J.N.; Kay, E.H.; Fisher, H.S.; Hoekstra, H.E. Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. *PLoS ONE* **2012**, *7*, e37135. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Cumer, T.; Pouchon, C.; Boyer, F.; Yannic, G.; Rioux, D.; Bonin, A.; Capblancq, T. Double-Digest RAD-Sequencing: Do Pre- and Post-Sequencing Protocol Parameters Impact Biological Results? *Mol. Genet. Genom.* **2021**, *296*, 457–471. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Jordon-Thaden, I.E.; Beck, J.B.; Rushworth, C.A.; Windham, M.D.; Diaz, N.; Cantley, J.T.; Martine, C.T.; Rothfels, C.J. A Basic ddRADseq Two-Enzyme Protocol Performs Well with Herbarium and Silica-Dried Tissues across Four Genera. *Appl. Plant Sci.* **2020**, *8*, e11344. [\[CrossRef\]](#)
47. He, J.; Zhao, X.; Laroche, A.; Lu, Z.-X.; Liu, H.; Li, Z. Genotyping-by-Sequencing (GBS), an Ultimate Marker-Assisted Selection (MAS) Tool to Accelerate Plant Breeding. *Front. Plant Sci.* **2014**, *5*, 484. [\[CrossRef\]](#)
48. Qin, B.; Hu, Y.; Huang, Y.; Liang, Y.; Guo, X.; Huang, B. Genotyping-by-Sequencing of *Illicium difengpi* Highlights Its Potential Genetic Diversity and Conservation Status. *Sci. Rep.* **2025**, *15*, 21015. [\[CrossRef\]](#)
49. Elshire, R.J.; Glaubitz, J.C.; Sun, Q.; Poland, J.A.; Kawamoto, K.; Buckler, E.S.; Mitchell, S.E. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLoS ONE* **2011**, *6*, e19379. [\[CrossRef\]](#) [\[PubMed\]](#)
50. Heo, M.-S.; Han, K.; Kwon, J.-K.; Kang, B.-C. Development of SNP Markers Using Genotyping-by-Sequencing for Cultivar Identification in Rose (*Rosa hybrida*). *Hortic. Environ. Biotechnol.* **2017**, *58*, 292–302. [\[CrossRef\]](#)
51. Roday, M.C.; Millard, M.J.; Glaubitz, J.C.; Peiffer, J.A.; Swarts, K.L.; Casstevens, T.M.; Elshire, R.J.; Acharya, C.B.; Mitchell, S.E.; Flint-Garcia, S.A.; et al. Comprehensive Genotyping of the USA National Maize Inbred Seed Bank. *Genome Biol.* **2013**, *14*, R55. [\[CrossRef\]](#) [\[PubMed\]](#)
52. Zila, C.T.; Ogut, F.; Roday, M.C.; Gardner, C.A.; Buckler, E.S.; Holland, J.B. Genome-Wide Association Study of Fusarium Ear Rot Disease in the U.S.A. Maize Inbred Line Collection. *BMC Plant Biol.* **2014**, *14*, 372. [\[CrossRef\]](#) [\[PubMed\]](#)

53. Peiffer, J.A.; Romay, M.C.; Gore, M.A.; Flint-Garcia, S.A.; Zhang, Z.; Millard, M.J.; Gardner, C.A.C.; McMullen, M.D.; Holland, J.B.; Bradbury, P.J.; et al. The Genetic Architecture of Maize Height. *Genetics* **2014**, *196*, 1337–1356. [[CrossRef](#)]
54. Shrestha, S.; Adhikari, L.; Crain, J.; Dreisigacker, S.; Wu, S.; Singh, R.P.; Mondal, S.; Juliana, P.; Crossa, J.; Lucas, M.; et al. Genotyping Analysis of over 130,000 CIMMYT Bread Wheat Breeding Lines: A Decade-Long Effort in Optimizing Wheat Genotyping. *Plant Genome* **2025**, *18*, e70148. [[CrossRef](#)]
55. Poland, J.A.; Brown, P.J.; Sorrells, M.E.; Jannink, J.-L. Development of High-Density Genetic Maps for Barley and Wheat Using a Novel Two-Enzyme Genotyping-by-Sequencing Approach. *PLoS ONE* **2012**, *7*, e32253. [[CrossRef](#)]
56. Morris, G.P.; Ramu, P.; Deshpande, S.P.; Hash, C.T.; Shah, T.; Upadhyaya, H.D.; Riera-Lizarazu, O.; Brown, P.J.; Acharya, C.B.; Mitchell, S.E.; et al. Population Genomic and Genome-Wide Association Studies of Agroclimatic Traits in Sorghum. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 453–458. [[CrossRef](#)]
57. Lu, F.; Romay, M.C.; Glaubitz, J.C.; Bradbury, P.J.; Elshire, R.J.; Wang, T.; Li, Y.; Li, Y.; Semagn, K.; Zhang, X.; et al. High-Resolution Genetic Mapping of Maize Pan-Genome Sequence Anchors. *Nat. Commun.* **2015**, *6*, 6914. [[CrossRef](#)]
58. Sonah, H.; Bastien, M.; Iqura, E.; Tardivel, A.; Légaré, G.; Boyle, B.; Normandeau, É.; Laroche, J.; Larose, S.; Jean, M.; et al. An Improved Genotyping by Sequencing (GBS) Approach Offering Increased Versatility and Efficiency of SNP Discovery and Genotyping. *PLoS ONE* **2013**, *8*, e54603. [[CrossRef](#)]
59. Jaganathan, D.; Thudi, M.; Kale, S.; Azam, S.; Roorkiwal, M.; Gaur, P.M.; Kishor, P.B.K.; Nguyen, H.; Sutton, T.; Varshney, R.K. Genotyping-by-Sequencing Based Intra-Specific Genetic Map Refines a “QTL-Hotspot” Region for Drought Tolerance in Chickpea. *Mol. Genet. Genom.* **2015**, *290*, 559–571. [[CrossRef](#)]
60. Cericola, F.; Lenk, I.; Fè, D.; Byrne, S.; Jensen, C.S.; Pedersen, M.G.; Asp, T.; Jensen, J.; Janss, L. Optimized Use of Low-Depth Genotyping-by-Sequencing for Genomic Prediction Among Multi-Parental Family Pools and Single Plants in Perennial Ryegrass (*Lolium perenne* L.). *Front. Plant Sci.* **2018**, *9*, 369. [[CrossRef](#)] [[PubMed](#)]
61. Torkamaneh, D.; Laroche, J.; Belzile, F. Genome-Wide SNP Calling from Genotyping by Sequencing (GBS) Data: A Comparison of Seven Pipelines and Two Sequencing Technologies. *PLoS ONE* **2016**, *11*, e0161333. [[CrossRef](#)]
62. Brechley, R.; Spannagl, M.; Pfeifer, M.; Barker, G.L.A.; D’Amore, R.; Allen, A.M.; McKenzie, N.; Kramer, M.; Kerhornou, A.; Bolser, D.; et al. Analysis of the Bread Wheat Genome Using Whole-Genome Shotgun Sequencing. *Nature* **2012**, *491*, 705–710. [[CrossRef](#)]
63. Rimbart, H.; Darrier, B.; Navarro, J.; Kitt, J.; Choulet, F.; Leveugle, M.; Duarte, J.; Rivière, N.; Eversole, K.; on behalf of The International Wheat Genome Sequencing Consortium; et al. High Throughput SNP Discovery and Genotyping in Hexaploid Wheat. *PLoS ONE* **2018**, *13*, e0186329. [[CrossRef](#)]
64. Yang, S.; Fresnedo-Ramírez, J.; Wang, M.; Cote, L.; Schweitzer, P.; Barba, P.; Takacs, E.M.; Clark, M.; Luby, J.; Manns, D.C.; et al. A Next-Generation Marker Genotyping Platform (AmpSeq) in Heterozygous Crops: A Case Study for Marker-Assisted Selection in Grapevine. *Hortic. Res.* **2016**, *3*, 16002. [[CrossRef](#)]
65. Zou, C.; Karn, A.; Reisch, B.; Nguyen, A.; Sun, Y.; Bao, Y.; Campbell, M.S.; Church, D.; Williams, S.; Xu, X.; et al. Haplotyping the Vitis Collinear Core Genome with rhAmpSeq Improves Marker Transferability in a Diverse Genus. *Nat. Commun.* **2020**, *11*, 413. [[CrossRef](#)] [[PubMed](#)]
66. Weitemier, K.; Straub, S.C.K.; Cronn, R.C.; Fishbein, M.; Schmickl, R.; McDonnell, A.; Liston, A. Hyb-Seq: Combining Target Enrichment and Genome Skimming for Plant Phylogenomics. *Appl. Plant Sci.* **2014**, *2*, 1400042. [[CrossRef](#)] [[PubMed](#)]
67. Johnson, M.G.; Gardner, E.M.; Liu, Y.; Medina, R.; Goffinet, B.; Shaw, A.J.; Zerega, N.J.C.; Wickett, N.J. HybPiper: Extracting Coding Sequence and Introns for Phylogenetics from High-Throughput Sequencing Reads Using Target Enrichment. *Appl. Plant Sci.* **2016**, *4*, 1600016. [[CrossRef](#)] [[PubMed](#)]
68. Villaverde, T.; Pokorný, L.; Olsson, S.; Rincón-Barrado, M.; Johnson, M.G.; Gardner, E.M.; Wickett, N.J.; Molero, J.; Riina, R.; Sanmartín, I. Bridging the Micro- and Macroevolutionary Levels in Phylogenomics: Hyb-Seq Solves Relationships from Populations to Species and Above. *New Phytol.* **2018**, *220*, 636–650. [[CrossRef](#)]
69. Loiseau, O.; Olivares, I.; Paris, M.; de La Harpe, M.; Weigand, A.; Koubínová, D.; Rolland, J.; Bacon, C.D.; Balslev, H.; Borchsenius, F.; et al. Targeted Capture of Hundreds of Nuclear Genes Unravels Phylogenetic Relationships of the Diverse Neotropical Palm Tribe Geonomateae. *Front. Plant Sci.* **2019**, *10*, 864. [[CrossRef](#)]
70. Ogutcen, E.; Christe, C.; Nishii, K.; Salamin, N.; Möller, M.; Perret, M. Phylogenomics of Gesneriaceae Using Targeted Capture of Nuclear Genes. *Mol. Phylogenetics Evol.* **2021**, *157*, 107068. [[CrossRef](#)]
71. Johnson, M.G.; Pokorný, L.; Dodsworth, S.; Botigué, L.R.; Cowan, R.S.; Devault, A.; Eiserhardt, W.L.; Epitawalage, N.; Forest, F.; Kim, J.T.; et al. A Universal Probe Set for Targeted Sequencing of 353 Nuclear Genes from Any Flowering Plant Designed Using K-Medoids Clustering. *Syst. Biol.* **2019**, *68*, 594–606. [[CrossRef](#)]
72. Guenay-Greunke, Y.; Bohan, D.A.; Traugott, M.; Wallinger, C. Handling of Targeted Amplicon Sequencing Data Focusing on Index Hopping and Demultiplexing Using a Nested Metabarcoding Approach in Ecology. *Sci. Rep.* **2021**, *11*, 19510. [[CrossRef](#)] [[PubMed](#)]

73. Onda, Y.; Takahagi, K.; Shimizu, M.; Inoue, K.; Mochida, K. Multiplex PCR Targeted Amplicon Sequencing (MTA-Seq): Simple, Flexible, and Versatile SNP Genotyping by Highly Multiplexed PCR Amplicon Sequencing. *Front. Plant Sci.* **2018**, *9*, 201. [\[CrossRef\]](#)
74. Nagano, S.; Hirao, T.; Takashima, Y.; Matsushita, M.; Mishima, K.; Takahashi, M.; Iki, T.; Ishiguri, F.; Hiraoka, Y. SNP Genotyping with Target Amplicon Sequencing Using a Multiplexed Primer Panel and Its Application to Genomic Prediction in Japanese Cedar, *Cryptomeria japonica* (L.f.) D.Don. *Forests* **2020**, *11*, 898. [\[CrossRef\]](#)
75. Hawliczek, A.; Bolibok, L.; Tofil, K.; Borzęcka, E.; Jankowicz-Cieślak, J.; Gawroński, P.; Kral, A.; Till, B.J.; Bolibok-Bragoszewska, H. Deep Sampling and Pooled Amplicon Sequencing Reveals Hidden Genic Variation in Heterogeneous Rye Accessions. *BMC Genom.* **2020**, *21*, 845. [\[CrossRef\]](#) [\[PubMed\]](#)
76. Takeshima, R.; Ogiso-Tanaka, E.; Yasui, Y.; Matsui, K. Targeted Amplicon Sequencing + Next-Generation Sequencing–Based Bulk Segregant Analysis Identified Genetic Loci Associated with Preharvest Sprouting Tolerance in Common Buckwheat (*Fagopyrum esculentum*). *BMC Plant Biol.* **2021**, *21*, 18. [\[CrossRef\]](#)
77. Kutasy, B.; Farkas, Z.; Kolics, B.; Decsi, K.; Hegedűs, G.; Kovács, J.; Taller, J.; Tóth, Z.; Kálmán, N.; Kazinczi, G.; et al. Detection of Target-Site Herbicide Resistance in the Common Ragweed: Nucleotide Polymorphism Genotyping by Targeted Amplicon Sequencing. *Diversity* **2021**, *13*, 118. [\[CrossRef\]](#)
78. Nishio, S.; Moriya, S.; Kunihisa, M.; Takeuchi, Y.; Imai, A.; Takada, N. Rapid and Easy Construction of a Simplified Amplicon Sequencing (Simplified AmpSeq) Library for Marker-Assisted Selection. *Sci. Rep.* **2023**, *13*, 10575. [\[CrossRef\]](#)
79. Loera-Sánchez, M.; Studer, B.; Kölliker, R. A Multispecies Amplicon Sequencing Approach for Genetic Diversity Assessments in Grassland Plant Species. *Mol. Ecol. Resour.* **2022**, *22*, 1725–1745. [\[CrossRef\]](#)
80. Ogiso-Tanaka, E.; Shimizu, T.; Hajika, M.; Kaga, A.; Ishimoto, M. Highly Multiplexed AmpliSeq Technology Identifies Novel Variation of Flowering Time-Related Genes in Soybean (*Glycine max*). *DNA Res.* **2019**, *26*, 243–260. [\[CrossRef\]](#) [\[PubMed\]](#)
81. Bernardo, A.; Wang, S.; Amand, P.S.; Bai, G. Using Next Generation Sequencing for Multiplexed Trait-Linked Markers in Wheat. *PLoS ONE* **2015**, *10*, e0143890. [\[CrossRef\]](#)
82. Galindo-González, L.; Pinzón-Latorre, D.; Bergen, E.A.; Jensen, D.C.; Deyholos, M.K. Ion Torrent Sequencing as a Tool for Mutation Discovery in the Flax (*Linum usitatissimum* L.) Genome. *Plant Methods* **2015**, *11*, 19. [\[CrossRef\]](#) [\[PubMed\]](#)
83. Fonseca, L.H.M.; Carlsen, M.M.; Fine, P.V.A.; Lohmann, L.G. A Nuclear Target Sequence Capture Probe Set for Phylogeny Reconstruction of the Charismatic Plant Family Bignoniaceae. *Front. Genet.* **2023**, *13*, 1085692. [\[CrossRef\]](#)
84. Baker, W.J.; Bailey, P.; Barber, V.; Barker, A.; Bellot, S.; Bishop, D.; Botigué, L.R.; Brewer, G.; Carruthers, T.; Clarkson, J.J.; et al. A Comprehensive Phylogenomic Platform for Exploring the Angiosperm Tree of Life. *Syst. Biol.* **2022**, *71*, 301–319. [\[CrossRef\]](#)
85. Yoo, M.-J.; Lee, B.-Y.; Kim, S.; Lim, C.E. Phylogenomics With Hyb-Seq Unravels Korean Hosta Evolution. *Front. Plant Sci.* **2021**, *12*, 645735. [\[CrossRef\]](#)
86. Nevill, P.G.; Zhong, X.; Tonti-Filippini, J.; Byrne, M.; Hislop, M.; Thiele, K.; van Leeuwen, S.; Boykin, L.M.; Small, I. Large Scale Genome Skimming from Herbarium Material for Accurate Plant Identification and Phylogenomics. *Plant Methods* **2020**, *16*, 1. [\[CrossRef\]](#)
87. Bohmann, K.; Mirarab, S.; Bafna, V.; Gilbert, M.T.P. Beyond DNA Barcoding: The Unrealized Potential of Genome Skim Data in Sample Identification. *Mol. Ecol.* **2020**, *29*, 2521. [\[CrossRef\]](#) [\[PubMed\]](#)
88. Cai, L.; Zhang, H.; Davis, C.C. PhyloHerb: A High-throughput Phylogenomic Pipeline for Processing Genome Skimming Data. *Appl. Plant Sci.* **2022**, *10*, e11475. [\[CrossRef\]](#)
89. Pezzini, F.F.; Ferrari, G.; Forrest, L.L.; Hart, M.L.; Nishii, K.; Kidner, C.A. Target Capture and Genome Skimming for Plant Diversity Studies. *Appl. Plant Sci.* **2023**, *11*, e11537. [\[CrossRef\]](#) [\[PubMed\]](#)
90. Campbell, E.O.; Brunet, B.M.T.; Dupuis, J.R.; Sperling, F.A.H. Would an RRS by Any Other Name Sound as RAD? *Methods Ecol. Evol.* **2018**, *9*, 1920–1927. [\[CrossRef\]](#)
91. Zhou, Z.; Jiang, Y.; Wang, Z.; Gou, Z.; Lyu, J.; Li, W.; Yu, Y.; Shu, L.; Zhao, Y.; Ma, Y.; et al. Resequencing 302 Wild and Cultivated Accessions Identifies Genes Related to Domestication and Improvement in Soybean. *Nat. Biotechnol.* **2015**, *33*, 408–414. [\[CrossRef\]](#) [\[PubMed\]](#)
92. Chia, J.-M.; Song, C.; Bradbury, P.J.; Costich, D.; de Leon, N.; Doebley, J.; Elshire, R.J.; Gaut, B.; Geller, L.; Glaubitz, J.C.; et al. Maize HapMap2 Identifies Extant Variation from a Genome in Flux. *Nat. Genet.* **2012**, *44*, 803–807. [\[CrossRef\]](#)
93. Markkandan, K.; Yoo, S.; Cho, Y.-C.; Lee, D.W. Genome-Wide Identification of Insertion and Deletion Markers in Chinese Commercial Rice Cultivars, Based on Next-Generation Sequencing Data. *Agronomy* **2018**, *8*, 36. [\[CrossRef\]](#)
94. Zhou, Y.; Chebotarov, D.; Kudrna, D.; Llaca, V.; Lee, S.; Rajasekar, S.; Mohammed, N.; Al-Bader, N.; Sobel-Sorenson, C.; Parakkal, P.; et al. A Platinum Standard Pan-Genome Resource That Represents the Population Structure of Asian Rice. *Sci. Data* **2020**, *7*, 113. [\[CrossRef\]](#)
95. Muñoz-Amatriaín, M.; Eichten, S.R.; Wicker, T.; Richmond, T.A.; Mascher, M.; Steuernagel, B.; Scholz, U.; Ariyadasa, R.; Spannagl, M.; Nussbaumer, T.; et al. Distribution, Functional Impact, and Origin Mechanisms of Copy Number Variation in the Barley Genome. *Genome Biol.* **2013**, *14*, R58. [\[CrossRef\]](#)

96. Swanson-Wagner, R.A.; Eichten, S.R.; Kumari, S.; Tiffin, P.; Stein, J.C.; Ware, D.; Springer, N.M. Pervasive Gene Content Variation and Copy Number Variation in Maize and Its Undomesticated Progenitor. *Genome Res.* **2010**, *20*, 1689–1699. [[CrossRef](#)]
97. Takagi, H.; Abe, A.; Yoshida, K.; Kosugi, S.; Natsume, S.; Mitsuoka, C.; Uemura, A.; Utsushi, H.; Tamiru, M.; Takuno, S.; et al. QTL-Seq: Rapid Mapping of Quantitative Trait Loci in Rice by Whole Genome Resequencing of DNA from Two Bulk Populations. *Plant J.* **2013**, *74*, 174–183. [[CrossRef](#)] [[PubMed](#)]
98. Huang, X.; Kurata, N.; Wei, X.; Wang, Z.-X.; Wang, A.; Zhao, Q.; Zhao, Y.; Liu, K.; Lu, H.; Li, W.; et al. A Map of Rice Genome Variation Reveals the Origin of Cultivated Rice. *Nature* **2012**, *490*, 497–501. [[CrossRef](#)]
99. Makarevitch, I.; Waters, A.J.; West, P.T.; Stitzer, M.; Hirsch, C.N.; Ross-Ibarra, J.; Springer, N.M. Transposable Elements Contribute to Activation of Maize Genes in Response to Abiotic Stress. *PLoS Genet.* **2015**, *11*, e1004915. [[CrossRef](#)]
100. Zhao, Q.; Feng, Q.; Lu, H.; Li, Y.; Wang, A.; Tian, Q.; Zhan, Q.; Lu, Y.; Zhang, L.; Huang, T.; et al. Pan-Genome Analysis Highlights the Extent of Genomic Variation in Cultivated and Wild Rice. *Nat. Genet.* **2018**, *50*, 278–284. [[CrossRef](#)]
101. Geibel, J.; Reimer, C.; Pook, T.; Weigend, S.; Weigend, A.; Simianer, H. How Imputation Can Mitigate SNP Ascertainment Bias. *BMC Genom.* **2021**, *22*, 340. [[CrossRef](#)]
102. Simpson, J.T.; Wong, K.; Jackman, S.D.; Schein, J.E.; Jones, S.J.M.; Birol, I. ABySS: A Parallel Assembler for Short Read Sequence Data. *Genome Res.* **2009**, *19*, 1117–1123. [[CrossRef](#)] [[PubMed](#)]
103. Ross, M.G.; Russ, C.; Costello, M.; Hollinger, A.; Lennon, N.J.; Hegarty, R.; Nusbaum, C.; Jaffe, D.B. Characterizing and Measuring Bias in Sequence Data. *Genome Biol.* **2013**, *14*, R51. [[CrossRef](#)] [[PubMed](#)]
104. Huang, X.; Wei, X.; Sang, T.; Zhao, Q.; Feng, Q.; Zhao, Y.; Li, C.; Zhu, C.; Lu, T.; Zhang, Z.; et al. Genome-Wide Association Studies of 14 Agronomic Traits in Rice Landraces. *Nat. Genet.* **2010**, *42*, 961–967. [[CrossRef](#)] [[PubMed](#)]
105. Watowich, M.M.; Chiou, K.L.; Graves, B.; Montague, M.J.; Brent, L.J.N.; Higham, J.P.; Horvath, J.E.; Lu, A.; Martinez, M.I.; Platt, M.L.; et al. Best Practices for Genotype Imputation from Low-Coverage Sequencing Data in Natural Populations. *Mol. Ecol. Resour.* **2025**, *25*, e13854. [[CrossRef](#)]
106. Baraja-Fonseca, V.; Arrones, A.; Vilanova, S.; Plazas, M.; Prohens, J.; Bombarely, A.; Gramazio, P. Benchmarking of Low Coverage Sequencing Workflows for Precision Genotyping in Eggplant. *BMC Plant Biol.* **2025**, *25*, 1125. [[CrossRef](#)]
107. Difabachew, Y.F.; Frisch, M.; Langstroff, A.L.; Stahl, A.; Wittkop, B.; Snowdon, R.J.; Koch, M.; Kirchhoff, M.; Cselényi, L.; Wolf, M.; et al. Genomic Prediction with Haplotype Blocks in Wheat. *Front. Plant Sci.* **2023**, *14*, 1168547. [[CrossRef](#)]
108. Biagini, S.A.; Becelaere, S.; Aerden, M.; Jatsenko, T.; Hannes, L.; Damme, P.V.; Breckpot, J.; Devriendt, K.; Thienpont, B.; Vermeesch, J.R.; et al. Genotype Imputation from Low-Coverage Data for Medical and Population Genetic Analyses. *Genome Res.* **2025**, *35*, 1929–1941. [[CrossRef](#)]
109. Kostyukova, V.; Pozharskiy, A.; Dulat, B.; Gritsenko, D. Isolation and Molecular Identification of *Monilinia fructigena* in Almaty Region of Kazakhstan. *Horticulturae* **2025**, *11*, 1029. [[CrossRef](#)]
110. Song, J.-M.; Guan, Z.; Hu, J.; Guo, C.; Yang, Z.; Wang, S.; Liu, D.; Wang, B.; Lu, S.; Zhou, R.; et al. Eight High-Quality Genomes Reveal Pan-Genome Architecture and Ecotype Differentiation of *Brassica napus*. *Nat. Plants* **2020**, *6*, 34–45. [[CrossRef](#)]
111. Gao, L.; Gonda, I.; Sun, H.; Ma, Q.; Bao, K.; Tieman, D.M.; Burzynski-Chang, E.A.; Fish, T.L.; Stromberg, K.A.; Sacks, G.L.; et al. The Tomato Pan-Genome Uncovers New Genes and a Rare Allele Regulating Fruit Flavor. *Nat. Genet.* **2019**, *51*, 1044–1051. [[CrossRef](#)]
112. Belser, C.; Istace, B.; Denis, E.; Dubarry, M.; Baurens, F.-C.; Falentin, C.; Genete, M.; Berrabah, W.; Chèvre, A.-M.; Delourme, R.; et al. Chromosome-Scale Assemblies of Plant Genomes Using Nanopore Long Reads and Optical Maps. *Nat. Plants* **2018**, *4*, 879–887. [[CrossRef](#)] [[PubMed](#)]
113. Daccord, N.; Celton, J.-M.; Linsmith, G.; Becker, C.; Choisne, N.; Schijlen, E.; van de Geest, H.; Bianco, L.; Micheletti, D.; Velasco, R.; et al. High-Quality de Novo Assembly of the Apple Genome and Methylome Dynamics of Early Fruit Development. *Nat. Genet.* **2017**, *49*, 1099–1106. [[CrossRef](#)] [[PubMed](#)]
114. Wang, L.; Wang, Z.-W.; Ma, H.-Y.; Sun, Z.-M.; Zhang, R.-G.; Qin, X.; Jia, K.-H.; Liu, D. A Near-Complete Reassembled Haplotype-Resolved Reference Genome of *Acer truncatum*. *Sci. Data* **2025**, *12*, 1990. [[CrossRef](#)] [[PubMed](#)]
115. Simpson, J.T.; Workman, R.E.; Zuzarte, P.C.; David, M.; Dursi, L.J.; Timp, W. Detecting DNA Cytosine Methylation Using Nanopore Sequencing. *Nat. Methods* **2017**, *14*, 407–410. [[CrossRef](#)]
116. Schreiber, J.; Wescoe, Z.L.; Abu-Shumays, R.; Vivian, J.T.; Baatar, B.; Karplus, K.; Akeson, M. Error Rates for Nanopore Discrimination among Cytosine, Methylcytosine, and Hydroxymethylcytosine along Individual DNA Strands. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 18910–18915. [[CrossRef](#)]
117. Wenger, A.M.; Peluso, P.; Rowell, W.J.; Chang, P.-C.; Hall, R.J.; Concepcion, G.T.; Ebler, J.; Fungtammasan, A.; Kolesnikov, A.; Olson, N.D.; et al. Accurate Circular Consensus Long-Read Sequencing Improves Variant Detection and Assembly of a Human Genome. *Nat. Biotechnol.* **2019**, *37*, 1155–1162. [[CrossRef](#)]
118. Travers, K.J.; Chin, C.-S.; Rank, D.R.; Eid, J.S.; Turner, S.W. A Flexible and Efficient Template Format for Circular Consensus Sequencing and SNP Detection. *Nucleic Acids Res.* **2010**, *38*, e159. [[CrossRef](#)]

119. Jain, M.; Koren, S.; Miga, K.H.; Quick, J.; Rand, A.C.; Sasani, T.A.; Tyson, J.R.; Beggs, A.D.; Dillthey, A.T.; Fiddes, I.T.; et al. Nanopore Sequencing and Assembly of a Human Genome with Ultra-Long Reads. *Nat. Biotechnol.* **2018**, *36*, 338–345. [\[CrossRef\]](#)
120. Espinosa, E.; Bautista, R.; Larrosa, R.; Plata, O. Advancements in Long-Read Genome Sequencing Technologies and Algorithms. *Genomics* **2024**, *116*, 110842. [\[CrossRef\]](#)
121. Matthews, C.A.; Watson-Haigh, N.S.; Burton, R.A.; Sheppard, A.E. A Gentle Introduction to Pangenomics. *Brief. Bioinform.* **2024**, *25*, bbae588. [\[CrossRef\]](#)
122. Shi, J.; Tian, Z.; Lai, J.; Huang, X. Plant Pan-Genomics and Its Applications. *Mol. Plant* **2023**, *16*, 168–186. [\[CrossRef\]](#)
123. Wang, W.; Mauleon, R.; Hu, Z.; Chebotarov, D.; Tai, S.; Wu, Z.; Li, M.; Zheng, T.; Fuentes, R.R.; Zhang, F.; et al. Genomic Variation in 3010 Diverse Accessions of Asian Cultivated Rice. *Nature* **2018**, *557*, 43–49. [\[CrossRef\]](#)
124. Lin, M.-J.; Iyer, S.; Chen, N.-C.; Langmead, B. Measuring, Visualizing, and Diagnosing Reference Bias with Biastools. *Genome Biol.* **2024**, *25*, 101. [\[CrossRef\]](#)
125. Sirén, J.; Monlong, J.; Chang, X.; Novak, A.M.; Eizenga, J.M.; Markello, C.; Sibbesen, J.A.; Hickey, G.; Chang, P.-C.; Carroll, A.; et al. Pangenomics Enables Genotyping of Known Structural Variants in 5202 Diverse Genomes. *Science* **2021**, *374*, abg8871. [\[CrossRef\]](#)
126. Garrison, E.; Sirén, J.; Novak, A.M.; Hickey, G.; Eizenga, J.M.; Dawson, E.T.; Jones, W.; Garg, S.; Markello, C.; Lin, M.F.; et al. Variation Graph Toolkit Improves Read Mapping by Representing Genetic Variation in the Reference. *Nat. Biotechnol.* **2018**, *36*, 875–879. [\[CrossRef\]](#)
127. Ahmad, B.; Su, Y.; Hao, Y.; Razzaq, T.; Arshad, R.; Zhang, Y.; Zhang, Y.; Wang, X.; Huang, G.; Su, X.; et al. Mango Pangenome Reveals Dramatic Impacts of Reference Bias on Population Genomic Analyses. *Hortic. Res.* **2025**, *12*, uhaf166. [\[CrossRef\]](#) [\[PubMed\]](#)
128. Rakocevic, G.; Semenyuk, V.; Lee, W.-P.; Spencer, J.; Browning, J.; Johnson, I.J.; Arsenijevic, V.; Nadj, J.; Ghose, K.; Suci, M.C.; et al. Fast and Accurate Genomic Analyses Using Genome Graphs. *Nat. Genet.* **2019**, *51*, 354–362. [\[CrossRef\]](#) [\[PubMed\]](#)
129. Zhu, X.; Yang, R.; Liang, Q.; Yu, Y.; Wang, T.; Meng, L.; Wang, P.; Wang, S.; Li, X.; Yang, Q.; et al. Graph-Based Pangenome Provides Insights into Structural Variations and Genetic Basis of Metabolic Traits in Potato. *Mol. Plant* **2025**, *18*, 590–602. [\[CrossRef\]](#)
130. Qin, P.; Lu, H.; Du, H.; Wang, H.; Chen, W.; Chen, Z.; He, Q.; Ou, S.; Zhang, H.; Li, X.; et al. Pan-Genome Analysis of 33 Genetically Diverse Rice Accessions Reveals Hidden Genomic Variations. *Cell* **2021**, *184*, 3542–3558.e16. [\[CrossRef\]](#) [\[PubMed\]](#)
131. Walkowiak, S.; Gao, L.; Monat, C.; Haberer, G.; Kassa, M.T.; Brinton, J.; Ramirez-Gonzalez, R.H.; Kolodziej, M.C.; Delorean, E.; Thambugala, D.; et al. Multiple Wheat Genomes Reveal Global Variation in Modern Breeding. *Nature* **2020**, *588*, 277–283. [\[CrossRef\]](#)
132. Hirsch, C.N.; Foerster, J.M.; Johnson, J.M.; Sekhon, R.S.; Muttoni, G.; Vaillancourt, B.; Peñagaricano, F.; Lindquist, E.; Pedraza, M.A.; Barry, K.; et al. Insights into the Maize Pan-Genome and Pan-Transcriptome. *Plant Cell* **2014**, *26*, 121–135. [\[CrossRef\]](#)
133. Bayer, P.E.; Petereit, J.; Durant, É.; Monat, C.; Rouard, M.; Hu, H.; Chapman, B.; Li, C.; Cheng, S.; Batley, J.; et al. Wheat Panache: A Pangenome Graph Database Representing Presence–Absence Variation across Sixteen Bread Wheat Genomes. *Plant Genome* **2022**, *15*, e20221. [\[CrossRef\]](#)
134. Liang, Y.-Y.; Liu, H.; Lin, Q.-Q.; Shi, Y.; Zhou, B.-F.; Wang, J.-S.; Chen, X.-Y.; Shen, Z.; Qiao, L.-J.; Niu, J.-W.; et al. Pan-Genome Analysis Reveals Local Adaptation to Climate Driven by Introgression in Oak Species. *Mol. Biol. Evol.* **2025**, *42*, msaf088. [\[CrossRef\]](#)
135. Ubbens, J.; Parkin, I.; Eynck, C.; Stavness, I.; Sharpe, A.G. Deep Neural Networks for Genomic Prediction Do Not Estimate Marker Effects. *Plant Genome* **2021**, *14*, e20147. [\[CrossRef\]](#)
136. He, F.; Xu, M.; Long, R.; Zhu, K.; Du, M.; Ma, W.; Xue, H.; Peng, Y.; Chen, L.; Kang, J.; et al. Integrative Multi-Omics and Genomic Prediction Reveal Genetic Basis of Salt Tolerance in Alfalfa. *J. Genet. Genom.* **2025**, *53*, 447–457. [\[CrossRef\]](#) [\[PubMed\]](#)
137. Li, J.; He, Z.; Zhou, G.; Yan, S.; Zhang, J. DeepAT: A Deep Learning Wheat Phenotype Prediction Model Based on Genotype Data. *Agronomy* **2024**, *14*, 2756. [\[CrossRef\]](#)
138. Zhang, A.; Pérez-Rodríguez, P.; San Vicente, F.; Palacios-Rojas, N.; Dhliwayo, T.; Liu, Y.; Cui, Z.; Guan, Y.; Wang, H.; Zheng, H.; et al. Genomic Prediction of the Performance of Hybrids and the Combining Abilities for Line by Tester Trials in Maize. *Crop J.* **2022**, *10*, 109–116. [\[CrossRef\]](#)
139. Khanna, A.; Anumalla, M.; Catolos, M.; Bhosale, S.; Jarquin, D.; Hussain, W. Optimizing Predictions in IRRI’s Rice Drought Breeding Program by Leveraging 17 Years of Historical Data and Pedigree Information. *Front. Plant Sci.* **2022**, *13*, 983818. [\[CrossRef\]](#)
140. Lozada, D.N.; Mason, R.E.; Sarinelli, J.M.; Brown-Guedira, G. Accuracy of Genomic Selection for Grain Yield and Agronomic Traits in Soft Red Winter Wheat. *BMC Genet.* **2019**, *20*, 82. [\[CrossRef\]](#)
141. Novielli, P.; Romano, D.; Pavan, S.; Losciale, P.; Stellacci, A.M.; Diacono, D.; Bellotti, R.; Tangaro, S. Explainable Artificial Intelligence for Genotype-to-Phenotype Prediction in Plant Breeding: A Case Study with a Dataset from an Almond Germplasm Collection. *Front. Plant Sci.* **2024**, *15*, 1434229. [\[CrossRef\]](#)
142. Lourenço, V.M.; Ogutu, J.O.; Rodrigues, R.A.P.; Posekany, A.; Piepho, H.-P. Genomic Prediction Using Machine Learning: A Comparison of the Performance of Regularized Regression, Ensemble, Instance-Based and Deep Learning Methods on Synthetic and Empirical Data. *BMC Genom.* **2024**, *25*, 152. [\[CrossRef\]](#)

143. MacNish, T.R.; Danilevich, M.F.; Bayer, P.E.; Bestry, M.S.; Edwards, D. Application of Machine Learning and Genomics for Orphan Crop Improvement. *Nat. Commun.* **2025**, *16*, 982. [\[CrossRef\]](#)
144. Meuwissen, T.H.E.; Hayes, B.J.; Goddard, M.E. Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics* **2001**, *157*, 1819–1829. [\[CrossRef\]](#)
145. Pérez-Rodríguez, P.; Crossa, J.; Rutkoski, J.; Poland, J.; Singh, R.; Legarra, A.; Autrique, E.; de los Campos, G.; Burgueño, J.; Dreisigacker, S. Single-Step Genomic and Pedigree Genotype $\times$ Environment Interaction Models for Predicting Wheat Lines in International Environments. *Plant Genome* **2017**, *10*, plantgenome2016.09.0089. [\[CrossRef\]](#)
146. Montesinos-López, O.A.; Montesinos-López, A.; Pérez-Rodríguez, P.; Barrón-López, J.A.; Martini, J.W.R.; Fajardo-Flores, S.B.; Gaytan-Lugo, L.S.; Santana-Mancilla, P.C.; Crossa, J. A Review of Deep Learning Applications for Genomic Selection. *BMC Genom.* **2021**, *22*, 19. [\[CrossRef\]](#) [\[PubMed\]](#)
147. Habier, D.; Fernando, R.L.; Kizilkaya, K.; Garrick, D.J. Extension of the Bayesian Alphabet for Genomic Selection. *BMC Bioinform.* **2011**, *12*, 186. [\[CrossRef\]](#)
148. Azodi, C.B.; Bolger, E.; McCarren, A.; Roantree, M.; de los Campos, G.; Shiu, S.-H. Benchmarking Parametric and Machine Learning Models for Genomic Prediction of Complex Traits. *G3 Genes | Genomes | Genet.* **2019**, *9*, 3691–3702. [\[CrossRef\]](#)
149. Lundberg, S.M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J.M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.-I. From Local Explanations to Global Understanding with Explainable AI for Trees. *Nat. Mach. Intell.* **2020**, *2*, 56–67. [\[CrossRef\]](#) [\[PubMed\]](#)
150. Katiyar, S.K.; Das, R.R.; Pazhamala, L.T.; Bartholomé, J.; Chandel, G.; Bilaro, A.; Asante, M.D.; Iftekharuddaula, K.M.; Islam, M.M.; Yadaw, R.B.; et al. Accelerated Breeding Modernization: A Global Blueprint for Driving Genetic Gains, Climate Resilience, and Food Security in Rice. *Theor. Appl. Genet.* **2025**, *138*, 293. [\[CrossRef\]](#) [\[PubMed\]](#)
151. Baertschi, C.; Cao, T.-V.; Bartholomé, J.; Ospina, Y.; Quintero, C.; Frouin, J.; Bouvet, J.-M.; Grenier, C. Impact of Early Genomic Prediction for Recurrent Selection in an Upland Rice Synthetic Population. *G3 Genes | Genomes | Genet.* **2021**, *11*, jkab320. [\[CrossRef\]](#)
152. Neyhart, J.L.; Tiede, T.; Lorenz, A.J.; Smith, K.P. Evaluating Methods of Updating Training Data in Long-Term Genomewide Selection. *G3 Genes | Genomes | Genet.* **2017**, *7*, 1499–1510. [\[CrossRef\]](#) [\[PubMed\]](#)
153. Juliana, P.; Montesinos-López, O.A.; Crossa, J.; Mondal, S.; González Pérez, L.; Poland, J.; Huerta-Espino, J.; Crespo-Herrera, L.; Govindan, V.; Dreisigacker, S.; et al. Integrating Genomic-Enabled Prediction and High-Throughput Phenotyping in Breeding for Climate-Resilient Bread Wheat. *Theor. Appl. Genet.* **2019**, *132*, 177–194. [\[CrossRef\]](#) [\[PubMed\]](#)
154. Li, S.; Yan, B.; Li, T.K.T.; Lu, J.; Gu, Y.; Tan, Y.; Gong, F.; Lam, T.-W.; Xie, P.; Wang, Y.; et al. Ultra-Low-Coverage Genome-Wide Association Study—Insights into Gestational Age Using 17,844 Embryo Samples with Preimplantation Genetic Testing. *Genome Med.* **2023**, *15*, 10. [\[CrossRef\]](#)
155. Sthapit, S.R.; Crain, J.; Larson, S.; Anderson, J.A.; Bajgain, P.; DeHaan, L.R.; Poland, J. A Low-coverage Skim-sequencing and Imputation Pipeline for Genomic Selection. *Plant Genome* **2025**, *18*, e70139. [\[CrossRef\]](#)
156. Jensen, S.E.; Charles, J.R.; Muleta, K.; Bradbury, P.J.; Casstevens, T.; Deshpande, S.P.; Gore, M.A.; Gupta, R.; Ilut, D.C.; Johnson, L.; et al. A Sorghum Practical Haplotype Graph Facilitates Genome-Wide Imputation and Cost-Effective Genomic Prediction. *Plant Genome* **2020**, *13*, e20009. [\[CrossRef\]](#)
157. Jordan, K.W.; Bradbury, P.J.; Miller, Z.R.; Nyine, M.; He, F.; Fraser, M.; Anderson, J.; Mason, E.; Katz, A.; Pearce, S.; et al. Development of the Wheat Practical Haplotype Graph Database as a Resource for Genotyping Data Storage and Genotype Imputation. *G3 Genes | Genomes | Genet.* **2022**, *12*, jkab390. [\[CrossRef\]](#)
158. Cuperus, J.T. Single-Cell Genomics in Plants: Current State, Future Directions, and Hurdles to Overcome. *Plant Physiol.* **2022**, *188*, 749–755. [\[CrossRef\]](#)
159. Moskal, K.; Puchta-Jasińska, M.; Bolc, P.; Motor, A.; Frankowski, R.; Pietrusińska-Radzio, A.; Rucińska, A.; Tomiczak, K.; Boczkowska, M. Why “Where” Matters as Much as “How Much”: Single-Cell and Spatial Transcriptomics in Plants. *Int. J. Mol. Sci.* **2025**, *26*, 11819. [\[CrossRef\]](#) [\[PubMed\]](#)
160. Marand, A.P.; Zhang, X.; Nelson, J.; dos Reis, P.A.B.; Schmitz, R.J. Profiling Single-Cell Chromatin Accessibility in Plants. *STAR Protoc.* **2021**, *2*, 100737. [\[CrossRef\]](#)
161. Liew, L.C.; You, Y.; Auroux, L.; Oliva, M.; Peirats-Llobet, M.; Ng, S.; Tamiru-Oli, M.; Berkowitz, O.; Hong, U.V.T.; Haslem, A.; et al. Establishment of Single-Cell Transcriptional States during Seed Germination. *Nat. Plants* **2024**, *10*, 1418–1434. [\[CrossRef\]](#)
162. Marand, A.P.; Jiang, L.; Gomez-Cano, F.; Minow, M.A.A.; Zhang, X.; Mendieta, J.P.; Luo, Z.; Bang, S.; Yan, H.; Meyer, C.; et al. The Genetic Architecture of Cell Type-Specific Cis Regulation in Maize. *Science* **2025**, *388*, eads6601. [\[CrossRef\]](#)
163. Engelhorn, J.; Snodgrass, S.J.; Kok, A.; Seetharam, A.S.; Schneider, M.; Kiwit, T.; Singh, A.; Banf, M.; Doan, D.T.H.; Khaipho-Burch, M.; et al. Genetic Variation at Transcription Factor Binding Sites Largely Explains Phenotypic Heritability in Maize. *Nat. Genet.* **2025**, *57*, 2313–2322. [\[CrossRef\]](#)
164. Zhu, T.; Li, T.; Lü, P.; Li, C. Single-Cell Omics in Plant Biology: Mechanistic Insights and Applications for Crop Improvement. *Adv. Biotechnol.* **2025**, *3*, 20. [\[CrossRef\]](#)

165. Li, H.; Feng, H.; Guo, C.; Yang, S.; Huang, W.; Xiong, X.; Liu, J.; Chen, G.; Liu, Q.; Xiong, L.; et al. High-Throughput Phenotyping Accelerates the Dissection of the Dynamic Genetic Architecture of Plant Growth and Yield Improvement in Rapeseed. *Plant Biotechnol. J.* **2020**, *18*, 2345–2353. [[CrossRef](#)] [[PubMed](#)]
166. Muraya, M.M.; Chu, J.; Zhao, Y.; Junker, A.; Klukas, C.; Reif, J.C.; Altmann, T. Genetic Variation of Growth Dynamics in Maize (*Zea mays* L.) Revealed through Automated Non-Invasive Phenotyping. *Plant J.* **2017**, *89*, 366–380. [[CrossRef](#)] [[PubMed](#)]
167. Campbell, M.T.; Du, Q.; Liu, K.; Brien, C.J.; Berger, B.; Zhang, C.; Walia, H. A Comprehensive Image-Based Phenomic Analysis Reveals the Complex Genetic Architecture of Shoot Growth Dynamics in Rice (*Oryza sativa*). *Plant Genome* **2017**, *10*, plantgenome2016.07.0064. [[CrossRef](#)]
168. Marand, A.P.; Chen, Z.; Gallavotti, A.; Schmitz, R.J. A Cis-Regulatory Atlas in Maize at Single-Cell Resolution. *Cell* **2021**, *184*, 3041–3055.e21. [[CrossRef](#)] [[PubMed](#)]
169. Zhou, S.; Kremling, K.A.; Bandillo, N.; Richter, A.; Zhang, Y.K.; Ahern, K.R.; Artyukhin, A.B.; Hui, J.X.; Younkin, G.C.; Schroeder, F.C.; et al. Metabolome-Scale Genome-Wide Association Studies Reveal Chemical Diversity and Genetic Control of Maize Specialized Metabolites. *Plant Cell* **2019**, *31*, 937–955. [[CrossRef](#)]
170. Cubizolles, N.; Rey, E.; Choulet, F.; Rimbart, H.; Laugier, C.; Balfourier, F.; Bordes, J.; Poncet, C.; Jack, P.; James, C.; et al. Exploiting the Repetitive Fraction of the Wheat Genome for High-Throughput Single-Nucleotide Polymorphism Discovery and Genotyping. *Plant Genome* **2016**, *9*, plantgenome2015.09.0078. [[CrossRef](#)]
171. Xiong, Z.; Pires, J.C. Karyotype and Identification of All Homoeologous Chromosomes of Allopolyploid Brassica Napus and Its Diploid Progenitors. *Genetics* **2011**, *187*, 37–49. [[CrossRef](#)]
172. Hoopes, G.; Meng, X.; Hamilton, J.P.; Achakkagari, S.R.; Guesdes, F.d.A.F.; Bolger, M.E.; Coombs, J.J.; Esselink, D.; Kaiser, N.R.; Kodde, L.; et al. Phased, Chromosome-Scale Genome Assemblies of Tetraploid Potato Reveal a Complex Genome, Transcriptome, and Predicted Proteome Landscape Underpinning Genetic Diversity. *Mol. Plant* **2022**, *15*, 520–536. [[CrossRef](#)]
173. Phillips, A.R. Variant Calling in Polyploids for Population and Quantitative Genetics. *Appl. Plant Sci.* **2024**, *12*, e11607. [[CrossRef](#)]
174. Njuguna, J.N.; Clark, L.V.; Lipka, A.E.; Anzoua, K.G.; Bagmet, L.; Chebukin, P.; Dwiyantri, M.S.; Dzyubenko, E.; Dzyubenko, N.; Ghimire, B.K.; et al. Impact of Genotype-Calling Methodologies on Genome-Wide Association and Genomic Prediction in Polyploids. *Plant Genome* **2023**, *16*, e20401. [[CrossRef](#)] [[PubMed](#)]
175. Clark, L.V.; Lipka, A.E.; Sacks, E.J. polyRAD: Genotype Calling with Uncertainty from Sequencing Data in Polyploids and Diploids. *G3 Genes | Genomes | Genet.* **2019**, *9*, 663–673. [[CrossRef](#)]
176. Gerard, D.; Ferrão, L.F.V.; Garcia, A.A.F.; Stephens, M. Genotyping Polyploids from Messy Sequencing Data. *Genetics* **2018**, *210*, 789–807. [[CrossRef](#)] [[PubMed](#)]
177. Voorrips, R.E.; Gort, G.; Vosman, B. Genotype Calling in Tetraploid Species from Bi-Allelic Marker Data Using Mixture Models. *BMC Bioinform.* **2011**, *12*, 172. [[CrossRef](#)]
178. Serang, O.; Mollinari, M.; Garcia, A.A.F. Efficient Exact Maximum a Posteriori Computation for Bayesian SNP Genotyping in Polyploids. *PLoS ONE* **2012**, *7*, e30906. [[CrossRef](#)] [[PubMed](#)]
179. Mollinari, M.; Olukolu, B.A.; Pereira, G.d.S.; Khan, A.; Gemenet, D.; Yencho, G.C.; Zeng, Z.-B. Unraveling the Hexaploid Sweetpotato Inheritance Using Ultra-Dense Multilocus Mapping. *G3 Genes | Genomes | Genet.* **2020**, *10*, 281–292. [[CrossRef](#)] [[PubMed](#)]
180. de Bem Oliveira, I.; Resende, M.F.R., Jr.; Ferrão, L.F.V.; Amadeu, R.R.; Endelman, J.B.; Kirst, M.; Coelho, A.S.G.; Munoz, P.R. Genomic Prediction of Autotetraploids; Influence of Relationship Matrices, Allele Dosage, and Continuous Genotyping Calls in Phenotype Prediction. *G3 Genes | Genomes | Genet.* **2019**, *9*, 1189–1198. [[CrossRef](#)]
181. Elias, M.; Chere, D.; Lule, D.; Serba, D.; Tirfessa, A.; Gelmessa, D.; Tesso, T.; Bantte, K.; Menamo, T.M. Multi-Locus Genome-Wide Association Study Reveal Genomic Regions Underlying Root System Architecture Traits in Ethiopian Sorghum Germplasm. *Plant Genome* **2024**, *17*, e20436. [[CrossRef](#)] [[PubMed](#)]
182. Josephs, E.B.; Stinchcombe, J.R.; Wright, S.I. What Can Genome-Wide Association Studies Tell Us about the Evolutionary Forces Maintaining Genetic Variation for Quantitative Traits? *New Phytol.* **2017**, *214*, 21–33. [[CrossRef](#)] [[PubMed](#)]
183. Lv, Y.; Dong, L.; Wang, X.; Shen, L.; Lu, W.; Si, F.; Zhao, Y.; Zhu, G.; Ding, Y.; Cao, S.; et al. Single- and Multi-Locus Genome-Wide Association Study Reveals Genomic Regions of Thirteen Yield-Related Traits in Common Wheat. *BMC Plant Biol.* **2024**, *24*, 1228. [[CrossRef](#)]
184. Gahlaut, V.; Jaiswal, V.; Singh, S.; Balyan, H.S.; Gupta, P.K. Multi-Locus Genome Wide Association Mapping for Yield and Its Contributing Traits in Hexaploid Wheat under Different Water Regimes. *Sci. Rep.* **2019**, *9*, 19486. [[CrossRef](#)] [[PubMed](#)]
185. Heffner, E.L.; Sorrells, M.E.; Jannink, J.-L. Genomic Selection for Crop Improvement. *Crop Sci.* **2009**, *49*, 1–12. [[CrossRef](#)]
186. Wang, X.; Xu, Y.; Hu, Z.; Xu, C. Genomic Selection Methods for Crop Improvement: Current Status and Prospects. *Crop J.* **2018**, *6*, 330–340. [[CrossRef](#)]
187. Rutkoski, J.; Poland, J.; Mondal, S.; Autrique, E.; Pérez, L.G.; Crossa, J.; Reynolds, M.; Singh, R. Canopy Temperature and Vegetation Indices from High-Throughput Phenotyping Improve Accuracy of Pedigree and Genomic Selection for Grain Yield in Wheat. *G3 Genes | Genomes | Genet.* **2016**, *6*, 2799–2808. [[CrossRef](#)]

188. Cao, S.; Loladze, A.; Yuan, Y.; Wu, Y.; Zhang, A.; Chen, J.; Huestis, G.; Cao, J.; Chaikam, V.; Olsen, M.; et al. Genome-Wide Analysis of Tar Spot Complex Resistance in Maize Using Genotyping-by-Sequencing SNPs and Whole-Genome Prediction. *Plant Genome* **2017**, *10*, plantgenome2016.10.0099. [[CrossRef](#)]
189. Wang, P.; Lehti-Shiu, M.D.; Lotreck, S.; Segura Abá, K.; Krysan, P.J.; Shiu, S.-H. Prediction of Plant Complex Traits via Integration of Multi-Omics Data. *Nat. Commun.* **2024**, *15*, 6856. [[CrossRef](#)]
190. Westhues, M.; Schrag, T.A.; Heuer, C.; Thaller, G.; Utz, H.F.; Schipprack, W.; Thiemann, A.; Seifert, F.; Ehret, A.; Schlereth, A.; et al. Omics-Based Hybrid Prediction in Maize. *Theor. Appl. Genet.* **2017**, *130*, 1927–1939. [[CrossRef](#)]
191. Burgueño, J.; de los Campos, G.; Weigel, K.; Crossa, J. Genomic Prediction of Breeding Values When Modeling Genotype $\times$ Environment Interaction Using Pedigree and Dense Molecular Markers. *Crop Sci.* **2012**, *52*, 707–719. [[CrossRef](#)]
192. Jia, Y.; Jannink, J.-L. Multiple-Trait Genomic Selection Methods Increase Genetic Value Prediction Accuracy. *Genetics* **2012**, *192*, 1513–1522. [[CrossRef](#)]
193. Montesinos-López, O.A.; Montesinos-López, A.; Crossa, J.; Gianola, D.; Hernández-Suárez, C.M.; Martín-Vallejo, J. Multi-Trait, Multi-Environment Deep Learning Modeling for Genomic-Enabled Prediction of Plant Traits. *G3 Genes | Genomes | Genet.* **2018**, *8*, 3829–3840. [[CrossRef](#)] [[PubMed](#)]
194. Sandhu, K.S.; Lozada, D.N.; Zhang, Z.; Pumphrey, M.O.; Carter, A.H. Deep Learning for Predicting Complex Traits in Spring Wheat Breeding Program. *Front. Plant Sci.* **2021**, *11*, 613325. [[CrossRef](#)] [[PubMed](#)]
195. Zingaretti, L.M.; Gezan, S.A.; Ferrão, L.F.V.; Osorio, L.F.; Monfort, A.; Muñoz, P.R.; Whitaker, V.M.; Pérez-Enciso, M. Exploring Deep Learning for Complex Trait Genomic Prediction in Polyploid Outcrossing Species. *Front. Plant Sci.* **2020**, *11*, 25. [[CrossRef](#)] [[PubMed](#)]

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.