

Chapter 28

Sequence-Based Marker Assisted Selection in Wheat



Marco Maccaferri, Martina Bruschi, and Roberto Tuberosa

Abstract Wheat improvement has traditionally been conducted by relying on artificial crossing of suitable parental lines followed by selection of the best genetic combinations. At the same time wheat genetic resources have been characterized and exploited with the aim of continuously improving target traits. Over this solid framework, innovations from emerging research disciplines have been progressively added over time: cytogenetics, quantitative genetics, chromosome engineering, mutagenesis, molecular biology and, most recently, comparative, structural, and functional genomics with all the related -omics platforms. Nowadays, the integration of these disciplines coupled with their spectacular technical advances made possible by the sequencing of the entire wheat genome, has ushered us in a new breeding paradigm on how to best leverage the functional variability of genetic stocks and germplasm collections. Molecular techniques first impacted wheat genetics and breeding in the 1980s with the development of restriction fragment length polymorphism (RFLP)-based approaches. Since then, steady progress in sequence-based, marker-assisted selection now allows for an unprecedentedly accurate ‘breeding by design’ of wheat, progressing further up to the pangenome-based level. This chapter provides an overview of the technologies of the ‘circular genomics era’ which allow breeders to better characterize and more effectively leverage the huge and largely untapped natural variability present in the Triticeae gene pool, particularly at the tetraploid level, and its closest diploid and polyploid ancestors and relatives.

Keywords Genetic diversity · Molecular marker · Mapping · QTLome · Cloning · Homoeologous loci

M. Maccaferri (✉) · M. Bruschi · R. Tuberosa
Alma Mater Studiorum–Università di Bologna, University of Bologna, Bologna, Italy
e-mail: marco.maccaferri@unibo.it; martina.bruschi3@unibo.it; roberto.tuberosa@unibo.it

28.1 Learning Objectives

- Assessing the feasibility, benefits, and shortcomings of molecular techniques and especially marker-assisted selection in wheat breeding.

28.2 Introduction

Meeting the food demand of a population of 10 billion by 2050 will require a substantial increase in genetic gain presently achieved mostly by conventional breeding approaches (see Chap. 27). In wheat and other crops, gains from selection are tapering off, also in part due to climate change effects, and will not meet the estimated 70% increase in crop productivity required by 2050 to feed mankind (see Chap. 21). This worrisome trend can be mitigated through genomics-assisted breeding, particularly through marker-assisted selection (MAS) and genomic selection, two procedures increasingly adopted to accelerate gain from selection in breeding programs worldwide [1, 2].

The success of the Green Revolution that fueled the high selection gains in the 1960s–1980s was mainly due to the previous identification, followed by deployment, of the semi-dwarf *Rht* alleles in combination with photoperiod-insensitive *Ppd* alleles which allowed for the selection of short and early flowering cultivars able to escape heat and drought and take advantage of higher nitrogen fertilization regimes (see Chap. 10). These remarkable results highlight the key role played by the genotype x environment x management (GxExM) interaction.

Additional traits played an important role towards the release of novel cultivars provided with alleles able to mitigate the negative effects of biotic (e.g., rusts, fusarium head blight, root rot, septoria tritici blotch, etc.), and abiotic (e.g., drought, heat, nutrient deficiency and toxicity, etc.) stress on yield and its stability. In both cases, the identification of beneficial alleles at the loci (genes and mostly quantitative trait loci: QTLs) governing the resistance/tolerance to such factors and their selection through MAS are being increasingly adopted to accelerate the gain from selection (see Chaps. 5 and 6). The identification of QTLs with a major effect on the target traits has been more frequently reported for biotic stress (see Chap. 19; see also Fig. 28.5), though some notable examples have been reported for abiotic stress [3, 4], particularly when targeting morpho-physiological traits (e.g., early vigor, root system architecture, staygreen, isotope discrimination, etc.) with predictive value as proxies for biomass production, water-use efficiency, yield components, yield and its stability [5].

28.3 Genetic Resources, Mapping Approaches and Database

The key role played by germplasm collections for both gene discovery and pre-breeding purposes has been highlighted in hexaploid wheat with the Watkins collection [6] and in tetraploid wheat with the Global Durum wheat Panel (GDP; [7]) and the Tetraploid wheat Global Collection (TGC; [8]) (Fig. 28.1).

Linkage mapping and association mapping also known as genome wide association study (GWAS) in wheat have been conducted using various molecular marker sets and platforms [9]. Therefore, cross-referencing loci and QTL mapping results across experiments and genetic materials is cumbersome but otherwise essential for increasing the accuracy of mapping, as well as for mapping the allele/haplotype distribution in germplasm collections and breeding pools across the QTLOme [10]. A valuable approach to prioritize the QTLs to focus on with MAS and eventually attempting their cloning is provided by meta-analyses compiling and comparing the results of multiple QTL studies, hence providing a more accurate mapping of QTLs and their overall value across environments [11, 12].

The wheat community shares the knowledge related to the various molecular marker sets used during the past 40 years, mainly through dedicated publications and the GrainGenes database (<https://wheat.pw.usda.gov/GG3/>). Widely used, common and high-quality molecular marker sets were first adopted for RFLPs and then for SSR markers. The genome density of SSR markers allowed for cross-referencing

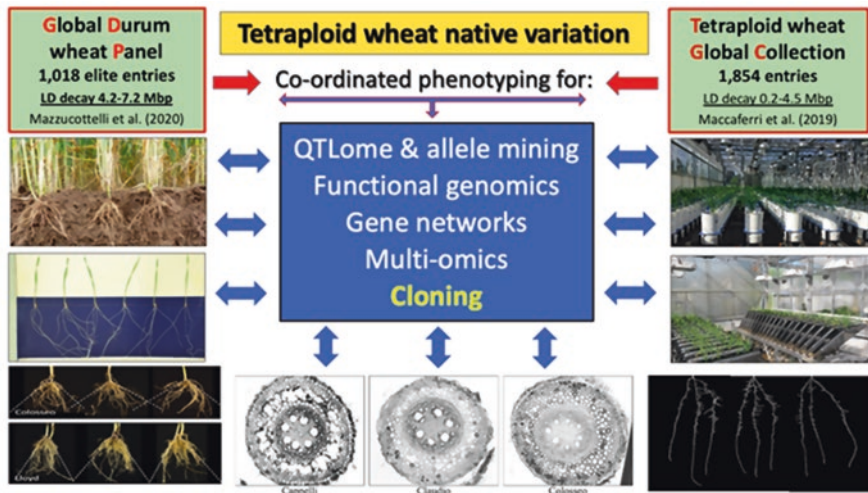


Fig. 28.1 The Global Durum wheat Panel (GDP; [7]) and the Tetraploid wheat Global Collection (TGC; [8]) are instrumental to mine the vast biodiversity present in the A and B tetraploid wheat genomes. The higher genetic variability coupled with lower linkage disequilibrium (LD) decay of the TGC indicates its suitability for QTL discovery and cloning while the GDP is more suitable for breeding purposes

across diverse linkage maps and highly polymorphic reference maps. Most important were the ITMI mapping population, a highly polymorphic map obtained from the cross of the bread wheat cv. Opata with a highly diverse wheat Synthetic line obtained from a cross between durum wheat and *Aegilops tauschii* and the Courtot x Chinese Spring intervarietal molecular marker linkage map. Subsequently, thanks to dedicated software, consensus maps providing higher genetic resolution and denser markers were assembled in both durum wheat [13] and bread wheat [14].

28.4 Dissecting the Wheat QTLome

The prevailing assumption has been that the variation in quantitative traits observed among wheat accessions is caused by the effects of multiple QTLs – mostly due to natural dominant mutations like insertion or deletion of bases (INDELS) in the regulatory gene regions – and the environment that inevitably limits our capacity for identifying QTLs, particularly under conditions of low heritability frequently present under abiotic stress ([5]; Chap. 13). Additionally, the wheat genome is huge and highly repetitive [8, 15], thus posing further difficulties in managing map-based cloning procedures that are implemented for the most interesting QTLs, clearly a very limited number (Fig. 28.2).

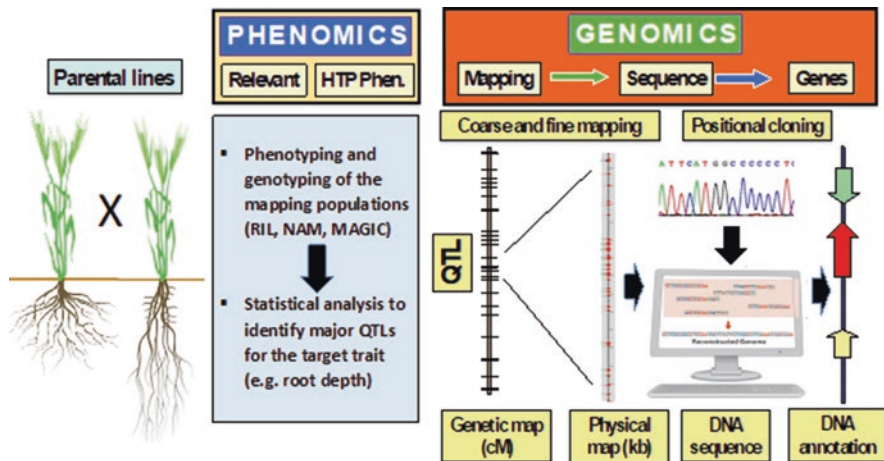


Fig. 28.2 The positional (map-based) cloning of a major QTL for a target trait (e.g., root depth) requires (1) the phenotyping and genotyping of an adequately large mapping population segregating for the trait, (2) the statistical analysis to map the QTLs and estimate their additive effect, (3) the fine mapping at high genetic resolution (possibly <0.1 cM) usually achieved with the phenotyping of a very large (from 1000 to 5000 F₂ plants depending on the heritability of the trait) population usually assembled from the cross of two near-isogenic lines contrasted for the QTL alleles. (Modified with permission from [16])

Enhancing genetic gain in wheat and other crops relies on the identification and, ideally, cloning of the loci governing the variability of the target traits followed by their selection via MAS and/or other genomic tools [2, 10].

More than three decades of dedicated experiments indicate that most QTL effects are small, as predicted by the so-called ‘infinitesimal’ model [17]. However, major QTLs (i.e., those accounting for >10% of the measured phenotypic variability) have also been reported and positionally cloned in wheat [18–20] which allows for designing the so-called ‘perfect marker’ for MAS (no recombination between the marker and the target locus) while advancing our understanding of the functional basis of variability of the target traits.

Once a QTL has been cloned via forward-genetics approaches, other reverse-genomics approaches (e.g., Targeting Induced Local Lesions in Genomes: TILLING, genetic engineering and gene editing, see Chap. 29) offer unprecedented opportunities to exploit native and/or artificially induced novel alleles. Considering the importance of quantitative traits for sustaining wheat performance under adverse conditions, increasing attention is being devoted to the mapping and cloning of major QTLs – hereafter defined ‘QTLome’ as a whole – which accounts for a sizeable portion of the variability targeted by breeders ([10]; Fig. 28.3).

Genomics-assisted wheat improvement is implemented in two complementary ways: (i) by targeting a limited number of well-characterized major QTLs via MAS (the tip of the iceberg in Fig. 28.3) and (ii) by leveraging the plethora of unknown

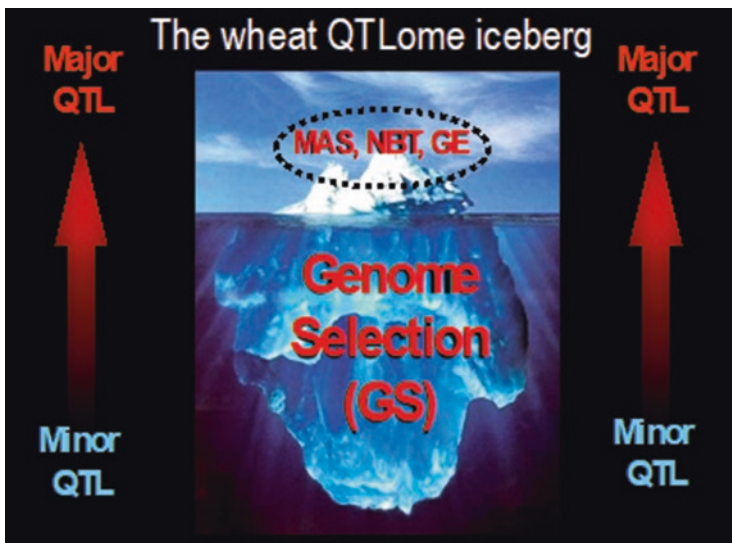


Fig. 28.3 The wheat QTLome represents the portion of QTLs with a sufficiently strong additive effect that makes their mapping and selection of the beneficial alleles via MAS possible. Only a minute fraction of these major QTLs can be cloned, hence allowing for the application of new breeding technologies (NBT; e.g., gene editing) and/or genetic engineering (GE). The vast majority of QTLs have additive effects too small to allow for their mapping. Their selection is possible through genome selection (GS)

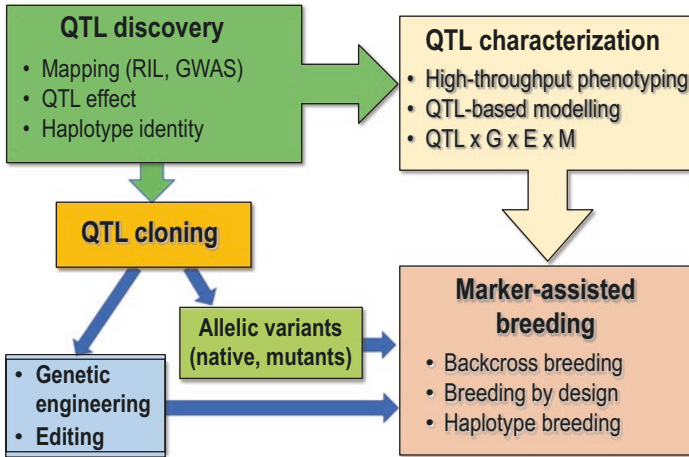


Fig. 28.4 How genomics-assisted breeding allows us to identify beneficial QTL alleles and deploys marker-assisted selection (MAS), genome editing, and/or genetic engineering (GE) to enhance the frequency of beneficial allelic variants in breeders' pools

QTLs with additive effects too small to be mapped (the submerged portion of the iceberg in Fig. 28.3) but otherwise indirectly selectable through genomic selection (GS; Chap. 30). The two approaches are complementary and their adoption – either as single approaches or in combination- should be based on a case-by-case evaluation, depending on the selection objectives, available genetic materials, and information on the genetic make-up and heritability of the target trait. The sequential or integrated adoption of both approaches (i.e., QTL-MAS followed by the application of GS models accounting for known genes/alleles as fixed effects) has been proven far more effective than GS alone [21].

Additionally, Fig. 28.4 indicates how, once a QTL has been cloned, the sequence information of the causative sequence (coding or non-coding) allows for the design of ‘perfect’ markers and the identification of rare native haplotypes present in the collection. Alternatively, the sequence of the QTL can be used to create novel alleles through gene editing (Chap. 29) and/or through genetic engineering, thereby enriching the MAS pipeline with novel alleles.

28.5 Selecting Traits and Loci for the MAS Pipeline

Choosing the traits suitable for the MAS pipeline requires a clear understating of the priorities and limiting factors of the breeding project based on (i) the prevailing environmental and phytosanitary conditions in the target environment and (ii) the concurrent effects on other traits (e.g., quality) of the targeted alleles/haplotypes *per se* caused by metabolic pleiotropy and/or caused by loci tightly linked to the allele/s targeted by MAS, the so-called ‘linkage drag’.

An aspect of paramount importance for loci/QTL/allele proper exploitation in applied breeding programs is the thorough evaluation of the QTLxGxE×M interaction that underpins the QTL effects [3]. This issue is often inadequately addressed, because an appropriate experimental design of field trials to achieve such a goal can be too expensive. Equally important when evaluating QTL effects is the concept of ‘envirotyping’ as a third ‘typing’ technology, complementing genotyping and phenotyping (Chap. 3). Envirotyping is a fundamental prerequisite to crop modeling and phenotype prediction through its functional components [22]. In this respect, modeling yield in wheat is particularly challenging due to its broad distribution across the globe and the contrasting environmental conditions under which wheat is grown (Chap. 31).

28.5.1 Loci for Phenology

The *Rht* and *Ppd* loci that fueled the Green Revolution are obvious “low-hanging fruit” for the application of MAS since heading date and height are primary determinants optimizing yield while ensuring its stability across environments. Data on the haplotype profiles at the *Rht* and *Ppd* loci are increasingly available for the founders and other modern genotypes that most frequently are used as parents to create novel segregating populations. Among the 46 currently known *Rht* genes and alleles (<https://shigen.nig.ac.jp/wheat/komugi/>), *RhtB1b* and *Rht-D1b* confer insensitivity to GA₃ and are the first two loci identified and used in the Green Revolution. However, taller and faster-growing wheat cultivars can be higher yielding than dwarf or semi-dwarf wheat genotypes under early and severe drought conditions, a finding likely related to effects of *Rht* alleles on coleoptile length and seedling/tiller vigor at early growth stages but also on root traits (e.g., root mass and depth) as shown by Beyer et al. [23]. The shorter wheat varieties are considered as better adapted to well-watered and nutrient-rich conditions rather than conditions of low soil moisture, a notable example of GxExM interaction, indicating how breeders can leverage MAS for *Rht* alleles to optimize yield and yield stability based upon the environmental conditions. As an example of wide differentiation of allelic distribution driven by adaptation and yield potential, we can consider the case of worldwide *RhtB1* allelic distribution in durum wheat. Most of the modern, highly productive durum varieties grown in the fertile and temperate areas under fall sowing and overwinter tillering are homozygous for the semi-dwarf *RhtB1b* allele while, on the contrary, this allele is rarely found in modern varieties bred for the Northern American prairies including North Dakota, Montana and Canada where extensive agriculture and short growing cycle are dominating.

Based on the environmental conditions (e.g., photoperiod, precipitation, temperature, etc.) of the target environment, breeding programs have been optimized for the alleles present at these loci in the parental lines and pre-breeding materials (Chaps. 3 and 25). Developmental regulatory networks include response to vernalization (*VRN loci*; Chap. 3) and response to day-length conditions, including *PHOTOPERIOD1*, *PHYB* or *PHYC*, *CO1*, and *CO2* as well as response to

vernalization and freezing tolerance, including *CBF* and *COLD REGULATED (COR)* genes.

A key player for the fine tuning of flowering time in both durum and bread wheat is the *Vrn-1* locus that regulates the switch from vegetative to the reproductive mode based upon the duration of the exposure to a critical threshold of number of days with temperatures between -2 and 15 °C. The regulation of flowering time in response to environmental temperature and day-length conditions is further fine-tuned by partially redundant networks, including a vernalization responsive network with four *VRN* loci: *VRN1*, *VRN2*, *VRN3=FLOWERING LOCUS T (FT)*, and *VRN4*, all amenable to MAS.

A similar situation has been reported for the *Ppd1* locus, also present with three homeologs (*Ppd-A1*, *Ppd-B1*, and *Ppd-D1*) of different strength and with different alleles, including various *Ppd*-insensitive dominant mutations in the gene promoter regions at all the three homeologs that were rapidly selected by breeders due to their positive effects in temperate environments. Additionally, copy number variation is another major cause of natural allelic variation in *VRN* and *PPD* genes. The *VRN* and *PPD* allelic combinations consciously or unconsciously selected by breeders at the three *VRN* genes, *Ppd1*, and at their homeologs, respectively, have been surveyed in both tetraploid and hexaploid wheat [24].

28.5.2 *Loci for the Root System Architecture (RSA)*

Notwithstanding the well-demonstrated importance of the *Rht* loci, increasing attention is being devoted to the loci that control RSA, particularly root mass and root depth, both of which have been shown to play a pivotal role in capturing soil moisture and nutrients [25]. Selection and breeding for RSA traits have been documented to be effective under conditions where plants complete their cycle based on stored soil water, a condition where deeper roots allow the plants to access deeper soil layers where more residual moisture is available as compared to upper soil layers. A marker-assisted approach targeting plants enriched in alleles conferring deeper roots would expedite the release of drought-tolerant cultivars under such conditions when residual moisture is more likely available at depth around anthesis and grain-filling when surface layers become dry [26].

28.5.3 *Loci for Disease Resistance*

Nowadays, MAS for resistance to fungal diseases, mainly rusts, fusarium head blight and root rot, septoria tritici blotch, and powdery mildew accounts for the vast majority of the MAS activities, particularly marker-assisted backcrossing, routinely carried out in wheat breeding programs worldwide.

The release of new cultivars during the Green Revolution largely relied on three-way (top crosses) and less from simple crosses. At CIMMYT, the three-way crosses (top crosses) approach was mostly effective in introducing and immediately recombine new innovative and beneficial alleles at multiple loci for plant architecture (*Rht*), phenology (*Ppd*), and rust (*Lr*, *Yr*, and *Sr*) resistance. Importantly, this approach resembled the three-way cross already adopted by the early Italian wheat geneticist and breeder Nazareno Strampelli to develop a first series of innovative wheat varieties in the 1920s that in Italy supported the ‘Battle for Grain’ launched in 1925 and eventually allowed the country to become self-sufficient in wheat production. The many Strampelli’s innovative varieties selected from the cross ‘Rieti/Wilhelmina//Akakomugi’ carried out in 1913, later spread worldwide, particularly in South America and China [27].

Increasing attention and effort are devoted to the identification of markers associated to loci for resistance to viruses (e.g., *SBCMV*) and/or insects (e.g., Hessian fly) whose diffusion and damaging effects are being increased by global warming. An example is provided by the search of markers linked to the loci for resistance to soil-borne cereal mosaic virus (*SBCMV*) which has been shown to reduce yield by 40–50% in susceptible commercial winter wheat cultivars in UK up to 70% in durum wheat in Italy [28].

28.6 Molecular Marker Technologies for MAS

A summary of marker technologies and their pros and cons is reported in Table 28.1. The ‘first generation markers’ developed at the onset of MAS in the late 1980s was based on RFLP, a very expensive and time-consuming technology. The advent of the PCR technique ushered in a number of much cheaper and faster ‘second generation markers’ such as random amplified polymorphic DNA (RAPD), and derived markers such as sequence characterized amplified regions (SCAR). Previous studies conducted to dissect the QTLome of soil-borne cereal mosaic virus (*SBCMV*) resistance in durum wheat were based on SSR and Diversity Arrays Technology (DArT) markers [28]. However, these marker classes present a series of constraints: low throughput (SSR markers), genome density insufficient for fine mapping (SSR and DArT markers [29]) and limited informativeness (DArT markers in their original version). In the past decade, efficient use of SNPs has become possible thanks to the development of arrays like the Illumina 90K [30]. Based on the wide use of the Illumina 90K wheat array worldwide, Maccaferri et al. [31] developed a consensus map for tetraploid wheat harboring 30,144 markers in which the high density of gene-derived SNPs provides useful anchor points for positional cloning. The abundance of SNPs in the wheat genome, together with the possibility of coupling them with high-throughput genotyping technologies, like KASP (Kompetitive Allele Specific PCR; Chap. 18) makes them suitable for fine mapping which requires the sampling of thousands of plants.

Table 28.1 Time-course progress in molecular marker technologies with their pros and cons

Marker type	Platform/Technology	Time interval	Time/assay	Marker numerosity	Pros and cons/Multiplexability
<i>Hybridization-based markers</i>					
Restriction fragment length polymorphism (RFLP)	Enzymatic restriction + Probe hybridization + Electrophoresis + Southern Blot + radioactive/chemiluminescent detection	'70-'90	Several days	Hundreds	Very laborious and expensive. Mainly codominant. Multiplex (yes, 2–3 assay)
<i>Polymerase chain reaction (PCR)-based markers</i>					
Simple PCR from expressed sequence tags (EST)	PCR + horizontal electrophoresis	'80-now	One day	Hundreds, or custom design	One day. Allows to detect insertion/deletions (INDEL). Multiplex (no). Codominant
Cleaved amplified polymorphisms (CAPS)	PCR + Enzymatic restriction + electrophoresis	'80-now	One day	Hundreds, or custom design	One day. Allows to detect single nucleotide substitutions. Multiplex (no) Codominant
Allele specific oligonucleotide PCR (ASO)	PCR + electrophoresis	'80-now	One day	Hundreds, or custom design	One day. Allows to detect substitutions or INDEL. Multiplex (no). Codominant
Random Amplified Polymorphic DNA (RAPD)	PCR + electrophoresis	'80-'90	One day	Hundreds	One day. Multiplex (yes). Robustness drawback. Dominant
Amplified Fragment Polymorphisms (AFLP)	PCR + vertical electrophoresis (polyacrylamide gel fragment analysis)	'80-'90	One day	Hundreds	One day. Multiplex (yes). Robust. Dominant
Simple Sequence Repeat (SSR)	PCR + vertical electrophoresis	'80-now	One day	Thousands	One day. Multiplex (yes, 3–6 assay). Robust. Codominant
Real-Time hydrolysis probes (TaqMan®)	SNP Fluorescent detection	'90-now	One day	Thousands. Custom design	One day. Multiplex (no). Robust. Codominant. Custom design

Marker type	Platform/Technology	Time interval	Time/assay	Marker numerosity	Pros and cons/Multiplexability
Kompetitive Allele Specific PCR (KASP®)	SNP Fluorescent detection	2000-now	One day	Thousands. Custom design	One day. Multiplex (no). Robust. Codominant. Custom design
Single Nucleotide Polymorphism arrays (SNP arrays)	DNA microarrays. Illumina and Affimetrix technology	2000-now	Short turn around time	Highly flexible based on discovery panels	One week. High Multiplex of 10K-50K-90K-280K-660K-800K. Robust. Codominant. Suitable for whole genome analysis (GWAS, GS)
Targeted resequencing	Illumina Next Generation massive sequencing	2010	Short turn around time	Highly flexible based on discovery panels	One week. Multiplex of up to 1-10K Robust. Codominant. Suitable for whole genome analysis

The adoption of the SNP array technology and genotyping-by-sequencing (GBS) allowed for an unprecedented level of marker density and mapping quality [32]. A few arrays were quickly adopted by the wheat community, like the Illumina iSelect wheat 9K and 90K arrays [30] and the Affymetrix 35K array [33]. This allowed for the accumulation of mapping data sufficient to generate a newer, highly dense generation of reference and consensus maps. These maps reached a density of 1–10 marker/cM across the entire genome. Among those maps, the SNP-based durum consensus map assembled by Maccaferri et al. [31], joined all previous SSR- and SNP-based mapping information from tetraploid wheat. Reference consensus maps were quickly and widely adopted for (i) projecting QTL mapping results and QTL confidence intervals from multiple experiments into reference consensus maps/assembled genomes and (ii) providing a framework for assisting the wheat genome sequence assembly procedures/pipelines.

28.7 Reference Genome Assembly

Gold-standard wheat genome assemblies have been obtained for the hexaploid wheat Chinese Spring [15] and the tetraploid wheat Svevo [8]. Second-generation, highly accurate, platinum-standard genome assemblies are being developed based on the integration of Optical Mapping (Bionano) and third generation long-read sequencing technology (PacBio), as recently shown with the release of the hexaploid wheat pangenome based on 10 high-quality genome assemblies from highly diverse and widely used cultivars worldwide (<http://www.10wheatgenomes.com>). The release of these highly contiguous wheat genomes allows to accurately project most of the molecular marker sets irrespectively of the marker technology adopted (DArT and SSR markers, SNP array, GBS, etc.) and represent the best reference for investigating the wheat QTLome [8].

28.8 Handling Sequence Data for Developing KASP Markers

For over a decade, SSR markers have provided a highly accurate and sufficiently dense marker framework that allowed for the development of many MAS protocols [34]. The drawback of SSR genotyping is that it required high-resolution polyacrylamide gel electrophoresis, no longer required with SNP technology where alternative alleles are discriminated by fluorescence. However, the entire molecular marker detection technology had to be revisited to adapt to the requirement of the SNP substitution detection, that does not involve differences in molecular weight between the alternative alleles. Discriminating the alternative SNP alleles requires ‘allele-specific’ recognition assays, with discrimination based on in-plate direct fluorescence reading, usually detected on real-time PCR (also known as quantitative PCR, qPCR) machines or plate fluorescence readers, which bypasses electrophoresis and

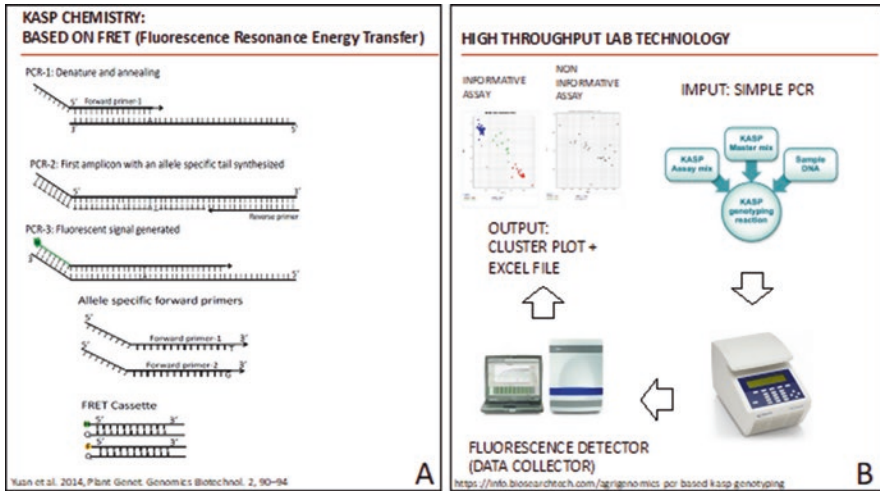


Fig. 28.5 (a) Schematic of KASP PCR (reprinted with permission from [36]). In evidence, the two allele-specific primer and the FRET cassette containing HEX and FAM fluorochromes. 1. The allele-specific primer anneals to the complementary sample DNA. 2. The first amplicon with allele-specific tail is synthesized. 3. The subsequent PCR cycles synthesize complements of the allele-specific tail sequence enabling the FRET cassette to bind the DNA and to emit allele-specific fluorescence based on the sample genotype formula. (b) Workflow of KASP genotyping technique. 1. Reagents required for KASP PCR. 2. Thermal cycler used to perform the reaction. 3. Detection of fluorescence during multiple amplifications performed in a Real-Time PCR instrument 4. Software output. See also <https://info.biosearchtech.com/agri-genomics-pcr-based-kasp-genotyping>

allows for the automation and high-throughput robotization of the assays. Together with the already well-established TaqMan technology, the KASP assay technology was progressively adopted due to the optimal combination of accuracy, easy implementation, and cost-effectiveness. Both technologies can accurately genotype SNPs based on either allele-specific probes (Taqman) or primers (KASP) (reviewed in [35]). Figures 28.5, 28.6, and 28.7 explain the main technical steps to design suitable KASP primers and implement the assays.

While in diploids the development of KASP assay is straightforward, this task poses several problems in tetra- and hexaploid wheat due to the high rate of similarity among the homeolog genome sequences adjacent to the varietal SNP. This entails a ‘dilution effect’ of the fluorescent signals that makes it progressively more difficult to accurately discriminate the target allelic variants that are genome-specific. Therefore, for allopolyploids, KASP primer design requires due attention to Mendelize the assay, i.e., making the assay as much genome-specific as possible, with primers being both allele- and genome-specific.

This requires multiple alignments of the two or three reference genomes in the SNP region in order to identify the position and sequence of both the varietal SNP (target- and genome-specific) and of the neighbor homeolog SNPs/INDELs that locally differentiate the genomes. Subsequently, an accurate design of the

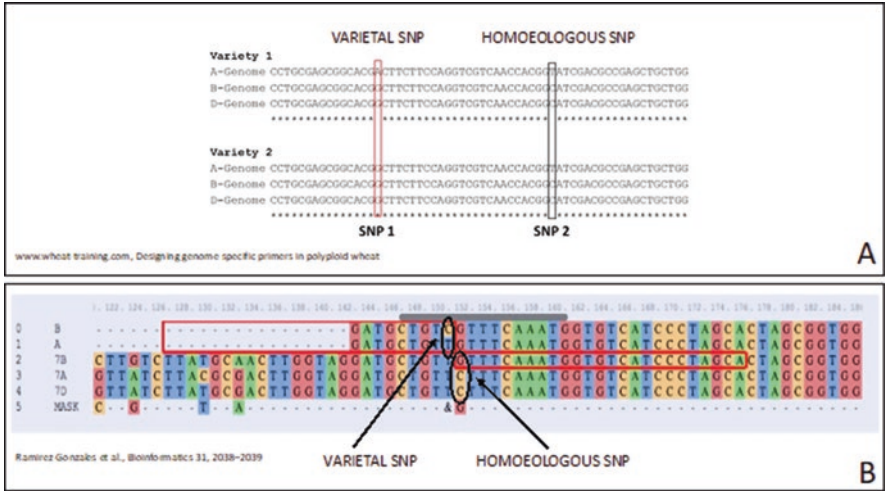


Fig. 28.6 (a) Example of hexaploid wheat sequence containing varietal (SNP 1) and homoelogenous (SNP 2) SNPs from www.wheat-training.com. Varietal SNPs are polymorphisms between varieties while homoelogenous SNPs are polymorphisms between genomes of a polyploid individual and typically non-polymorphic, though heterozygous, among varieties. A reliable genotype call can be obtained only by ensuring a sufficient NGS Illumina read depth on the polymorphic region (e.g., >8 reads). (b) Example of alignment performed by PolyMarker, a primer design pipeline for polyploids. KASP allele-specific primers are designed based on the varietal SNP, while the common primer is based on the homoelogenous SNP and gives genome specificity to the KASP assay. (Modified with permission from [37])

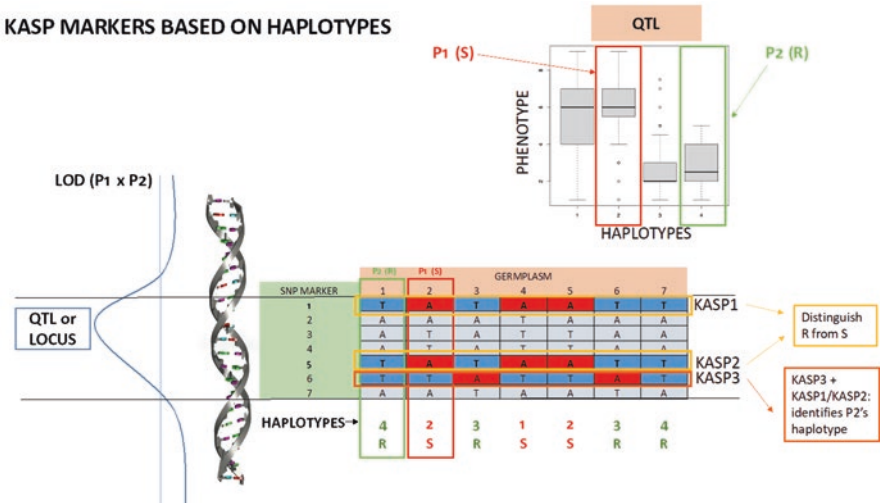


Fig. 28.7 Haplotype-based development of KASP assays for a disease resistance QTL. Haplotypes of resistant/susceptible parental lines can be used to develop diagnostic KASP assays that are predictive of multiallelic haplotypes (four haplotypes are represented). *P1* parental line 1, *P2* parental line 2, *R* resistant, *S* susceptible

allele-specific primers and of the genome-specific common primers is implemented with the support of dedicated primer-design software [38].

Additionally, the use of the reference genomes is relevant also to check genome-wide for off-target priming sites to prevent designing potentially non-specific primers on SNPs at loci other than those being targeted.

28.9 Examples of MAS

In wheat, protocols for tagging beneficial alleles suitable to MAS have been published in dedicated journals and made available to public and private research institutions and breeders worldwide since the late 1980s. Apart from specific literature searches using scientific publication browsers, effort has been made to provide access to this vast albeit fragmented knowledge. In particular, websites and databases specifically cataloguing MAS results are available at Graingenes (<https://wheat.pw.usda.gov/GG3/>), Komugi (<https://shigen.nig.ac.jp/wheat/komugi/>) as well as the Catalogue of Gene Symbols for Wheat. Even more focused websites are MAS-WHEAT (<https://maswheat.ucdavis.edu/>), the T3 Triticeae Toolbox website (<https://wheat.triticeatoolbox.org/>), and CIMMYT publications (Laboratory Protocols, CIMMYT Applied Molecular Genetics Laboratory). In particular, MAS-Wheat provides a concise and informative report for each locus of breeding interest, assembled by the original study authors and including a short locus and allele description, molecular marker protocols, primer sequences, and expected amplification/hybridization allelic results. To date, 65 protocols are stored in MAS-wheat and more are expected. Additionally, databases are being developed to store, classify and manage the QTL results that are continuously published, either in the form of meta-QTL studies for several traits or more recently, as QTL databases: see T3 and WheatQTLdb [24].

Once the target locus/QTL has been identified, either through linkage or association mapping, geneticists and breeders develop one and preferably multiple user-friendly molecular marker assays useful for tracing the beneficial alleles through MAS. Due to the inherent difficulty in understanding the nature/localization of the causative gene and causative polymorphism (i.e., quantitative trait nucleotide, QTN), most molecular assays for MAS have been developed from the same original markers (SNP and/or INDEL) used in the mapping study, provided they are linked (<5 cM) or preferably tightly linked (<1 cM) to the locus/QTL peak.

These newly or re-designed single marker assays are immediately available for the MAS of plants with the desired allele/s. However, there are cases where these single assays are not acceptable for their weak diagnostic power and excess of false positives. This discrepancy is proportional to their distance from the target locus, since assayed markers still recombine with the causative gene, and to their capacity to discriminate the functional haplotypes at the causal loci. To limit the impact of recombination, it is always advisable to rely on at least a couple of markers flanking the target locus/QTL peak. In this case, the frequency of false positives can be

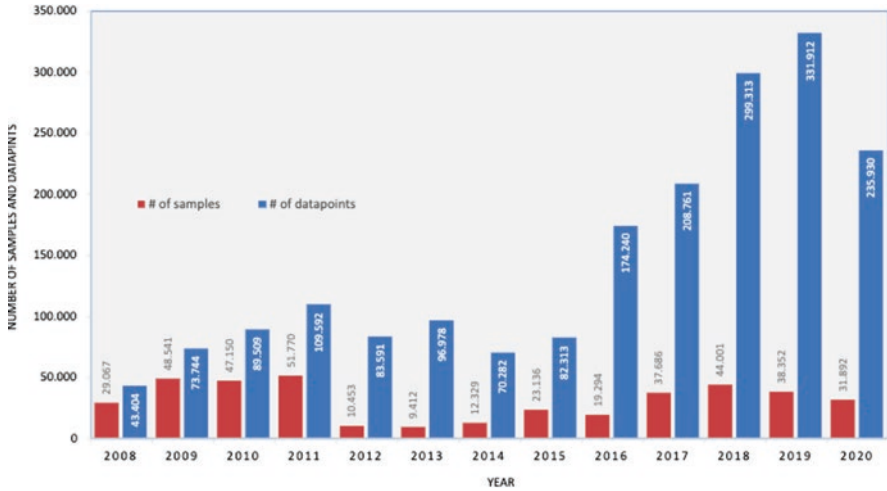


Fig. 28.8 Number of DNA samples and molecular marker assays used for MAS by CIMMYT's Global Wheat Program from 2008 to 2020. (Courtesy of the CIMMYT's Global Wheat Program)

predicted based upon the product of the distances of the two markers flanking the locus being selected. As an example, the rate of false positives of a MAS relying on two markers flanking the target locus at 2.2 and 1.6 cM will be slightly lower (due to negative crossover interference) than $2.2\% \times 1.6\% = 0.35\%$.

A more subtle albeit critical aspect is that the natural variation found at causal genes occurs at multi-allelic haplotypes in a gene. Multiple mutations with diverse phenotypic effects occur at different times in the promoter, exons, and introns of causal genes and are typically organized in haplotypes and haplogroups. Notably, single bi-allelic markers do not have enough discriminant power to trace the haplotypes of interest and a single SNP usually pre-dates or post-dates the haplotypes of interest. Therefore, precise MAS applications require the use of haplotypes comprising multiple SNPs in the target regions rather than single SNPs. In durum wheat, haplotype discrimination was adopted for the MAS of *Lr14a* based on SSR markers [29] and it is now increasingly adopted thanks to the rapid accumulation of genotypic data.

At CIMMYT, MAS was introduced around 2006 to select parental lines with the beneficial alleles at key loci for phenology, mainly *Vrn*, *Rht*, *Ppd* for phenology, resistance to rust (*Lr34*) and fusarium head blight (*Fhb*), and quality (*Glu-1*). The program quickly scaled up to segregating materials once more user-friendly markers became available and were adopted for multiplexing multiple traits. Figure 28.8 clearly indicates this trend after 2012 when each DNA sample was probed, on average, for up to 7 loci.

Table 28.2 presents a synopsis of the main loci targeted to develop MAS protocols which are being implemented in pre-breeding and/or breeding programs in wheat. The details of the references reporting the loci targeted by MAS are reported in Gupta et al. [1], Kumar et al. [39], and King et al. (Chap. 18).

Table 28.2 Targets of marker-assisted selection in wheat for loci/QTL present in the primary gene pool, native alleles, and number of loci with markers and protocols available in the public domain. Loci reported in **bold** have been directly identified through positional cloning

Traits and loci amenable to marker-assisted selection	Bread wheat	Durum wheat	Loci
<i>Plant architecture</i>			
Reduced height (<i>Rht</i>)	***	**	<i>RhtB1, Rht12, Rht18</i>
<i>Phenology, days to flowering</i>			
Vernalization (<i>VRN</i>)	***	***	<i>VRN1, VRN2, VRN3, VRN4</i>
Sensitivity to photoperiod (<i>Ppd</i>)	***	***	<i>Ppd1</i>
Earliness <i>per se</i> (<i>Eps</i>)	***		<i>ELF3</i>
<i>Resistance to foliar pathogens</i>			
<i>Biotrophs</i>			
Yellow rust – <i>Yr</i>	***	***	80 > loci named; 17 reported on MAS-WHEAT: <i>Yr5a and b; Yr7; Yr10; Yr18; YR36 = WHEAT KINASE START1 (WKS1); Yr46; Wtk1(Yr15, YrG303, and YrH52); YrAS2388</i>
Black or Stem rust – <i>Sr</i>	***	***	60 > loci named; 28 reported on MAS-WHEAT: <i>Sr2; Sr13; Sr21; Sr22; Sr26; Sr33; Sr35; Sr45; Sr46; Sr50; Sr60/WTK2; Sr61; SrTA1662</i>
Brown rust – <i>Lr</i>	***	***	80 > loci named; 16 reported on MAS-WHEAT: <i>Lr1; Lr10; Lr13; Lr14; Lr21; Lr22a; Lr34; Lr46; Lr67</i>
Powdery mildew – <i>Pm</i>	***	***	100 > loci named; KASP available for several loci in literature. <i>Pm1a; Pm2; Pm3; Pm4; Pm5e</i>
<i>Emi-biotrophs</i>			
<i>Septoria tritici</i> blotch (<i>Stb</i>)	***	***	20 > loci named; <i>Stb6-Stb16</i>
<i>Parastagonospora nodorum</i> (<i>Snn</i> and <i>Tsn-Tox</i>)	***	**	Nine host gene–necrotrophic effector interaction identified
<i>P. tritici-repentis</i> and <i>Bipolaris sorokiniana</i>	***	*	<i>Tsn1</i> to <i>Tsn5; ToxA</i> to <i>ToxD</i> <i>Tsn1/SnToxA; Snn1/SnTox1;</i> <i>Snn3-D1/SnTox3; Snn5/SnTox5</i>
Fusarium head blight (FHB)	***	***	<i>Fhb1</i> to <i>Fhb7. Fhb1, Fhb7. TaUGT6.</i> Several QTLs reported including <i>Qfhs.ifa-5A</i>
Fusarium root rot (<i>FRR</i>)	**	**	<i>FRR</i> QTLs
<i>Viruses</i>			
Soil Borne Viruses	**	***	<i>Sbm1, Sbm2.</i> QTLs for wheat yellow mosaic
Aphids-transmitted	***	**	QTLs for barley yellow dwarf virus
<i>Pests</i>			
Hessian Fly (<i>HF</i>)	***	***	35 HF resistance genes identified. <i>H</i> genes
Orange Wheat Blossom Midge	***	*	<i>Sm1</i> QTLs
Wheat Stem Sawfly	***	***	<i>SSt1</i> (<i>Solid stem locus</i>)

(continued)

Table 28.2 (continued)

Traits and loci amenable to marker-assisted selection	Bread wheat	Durum wheat	Loci
<i>Grain yield components</i>			
Spike development/shape	*	*	<i>AP2L2</i> and <i>Q. microRNA172</i> , <i>FRIZZY PANICLE (WFZP)</i>
Spike fertility (spikelet number per spike, fertile florets per spikelet)	**	*	Several QTLs. <i>WAP0-A1. GNII, GNI2</i>
Grain size and weight	***	**	Grain Weight (<i>GW1</i> to <i>GW8</i>), Grain width, Grain Length, and Grain Size genes. <i>TaGW2-A1. TaGS3</i> . Several additional QTLs identified
Carbon metabolism	***	**	Trehalose 6-phosphate synthase and sucrose synthase. ADP glucose pyrophosphorylase (AGPase)
<i>Abiotic stress tolerance</i>			
Nitrogen use (NU) and nitrogen-use efficiency (NUE)	***	**	<i>GS. Fd-GOGAT</i> . QTLs
Resilience to low temperatures/frost	*	*	<i>DREB/CBF</i> factors. <i>COR. VRN</i> QTLs.
Resilience to water and heat stress	**	*	<i>DREB/CBF</i> factors. QTLs. Heat shock proteins
Root system traits	**	**	<i>TaVSR-B1. QTLs</i> Root Growth Angle (<i>RG</i> A), root biomass and root elongation, root/shoot ratio. QTLs
Resistance to herbicides (metribuzin and imazamox)	***	*	<i>AhasL-D1</i> and <i>AhasL-B1</i> (resistance to imidazolinones). QTLs
Resistance to boron	***	*	<i>Bo1</i>
Aluminium tolerance	***	*	<i>TaALMT1</i>
Salinity stress tolerance	***	**	<i>Kna1 = TaHKT1;5-D, Nax2 = TmHKT1;5-A</i>
<i>Technological and nutritional quality</i>			
Grain Storage proteins. Quality and quantity. Gluten strength.	***	***	<i>Glu-1, Bx7, Gpc-B1 (=TaNAC)</i> . QTLs for GlutoPeak parameters and for gluten strength
Starch quality and quantity, amylose (<i>resistant starch</i>) content	**	***	<i>SBEII, GBSS (Wx1)</i>
Carotenoids and luteins synthesis and degradation (<i>grain yellow pigment content</i>)	*	***	<i>Psy1, Zds1, CYP, LCYE, LCYB, Lpx1-3, Ppo1-2</i> . QTLs and causative genes in the terpenoid pathway
Grain texture (<i>Hardness</i>)	***	*	<i>PINA, PINB</i>
Pre-Harvest Sprouting (<i>PHS</i>)	***	*	Several <i>PHS</i> loci and QTLs tagged by <i>KASP TaPHS1, TaMFT, Myb10-D; TaABI4</i>

(continued)

Table 28.2 (continued)

Traits and loci amenable to marker-assisted selection	Bread wheat	Durum wheat	Loci
Cadmium accumulation (reduced cadmium concentration)	*	***	<i>Cdu1</i>
Antinutritional factors	**	**	α -Amylase/Trypsin inhibitors

***wide interest and high relevance worldwide

**high interest at the local level, possible increase of relevance in the future

*low interest

The more tightly associated the markers are to the causative gene and the more markers that are being developed at the target locus, the more the combined SNP assays (haplotypes or ‘haplo-markers’) are diagnostic of the functional alleles at the causative gene and can be considered as ‘diagnostic’ or ‘predictive’ of the favorable alleles and phenotypes in various genetic backgrounds/crosses. These markers can be considered highly reliable and as such, widely recommended and used. Haplotype-based breeding is thus one of the most advanced areas for MAS [40].

Once the causative gene is cloned either by positional cloning or by functional analysis of candidate genes in the locus region and the causative polymorphisms are identified, it is possible to design the so-called ‘perfect markers’, i.e., one or more molecular assays diagnostic for the causative alleles and phenotypes, and coinciding with the functional haplotypes at the causative gene and as such not subjected to recombination.

Identifying the causal gene underlying a locus/QTL is a long-time, resource-demanding procedure, albeit highly rewarding in the case of loci of paramount breeding importance. In the past decade, due to the huge and complex wheat genome this goal has only been reached for few genes (*Bo1*, *GPC*, *Lr1*, *Lr10*, *Ppd*, *Q*, *Tsn1*, *VRN*, and *Yr36*). Importantly, the international efforts aiming at developing the genomic resources in wheat have shown an impressive acceleration in the last 5 years [8]. This recently led to the isolation of the causative genes for several loci in a few years, with *Fhb1* being one of the most relevant ones, followed by *Cdu1*, *Fhb7*, *MIWE18*, *Lr14*, *Pm4*, *Sst1*, and *Yr15* as well as several *Lr*, *Sr*, and *Yr* genes. First, the isolation of the causative genes at several loci of breeding interest allowed to develop so-called ‘perfect molecular markers’ designed rightly on the nucleotide polymorphisms causative of the phenotype, and therefore highly diagnostics and not subjected to recombination. Secondly isolation of causative genes allows us to better appreciate the range and the complexity of the mutations causing the functional native allelic diversity. A notable example is the *Fhb1* locus, an example of complex locus including natural variation at a causative gene for which the wheat reference genome Chinese Spring was uninformative [41]. Additionally, it has been shown that presence/absence variants (PAV) and copy number variation (CNV) as frequent causal polymorphism of native variation at *Bo1* and *Sst1*.

Importantly, the allopolyploid nature of wheat entails the presence of two and three copies of the same gene (called homeologous copies) in tetraploid and

hexaploid wheat, respectively. Gene functionality is usually retained, although differentiation is common in terms of genes silencing, sub-functionalization, and even neo-functionalization [42], hence introducing a wider variation in effects and a dosage effects not observed in diploids. QTLs can be found at the AA, BB, and DD homeologs, as in the case of *VRNI* and *Ppd1*. Another side effect of allopolyploidy is that most of the natural variation with an appreciable phenotypic effect is mostly caused by dominant mutations, particularly in the regulative gene regions (promoter region, first intron) while recessive mutation effects are frequently hidden by the presence of at least one functional gene copy.

28.10 MAS for Transferring Beneficial Haplotypes from Wheat Wild Relatives

Historically, wheat breeding has leveraged the wide diversity present in wheat wild relatives (Chap. 18). The Triticeae tribe is huge, with many diverse species well adapted to a wide range of environments, each showing specific peculiarities. Targets for chromatin transfer from wheat wild relatives are (i) resistance to several diseases, mainly rusts, powdery mildew, and fusarium head blight, (ii), grain quality, (iii) male sterility, (iv) resilience to abiotic stress, and (v) perenniality.

Both close and distant (alien) relatives have been largely used across decades (Chap. 18). Among the close relatives, *T. urartu*, *T. dicoccoides*, *T. monococcum*, *Aegilops speltoides*, and *Ae. tauschii*, and among the distant relatives, *Ae. geniculata*, *Ae. longissima*, *Ae. ventricosa*, *Haynaldia villosa*, *Secale cereale*, *Thinopyrum elongatum*, *Th. intermedium*, and *Th. ponticum*, were more frequently used.

The effective transfer and recombination of alien chromatin from distant wild relatives heavily relied on chromosome engineering techniques, most exclusively with the use of mutations at wheat *Ph* (*Pairing homoeologous*) genes, mainly *Ph1*. Chromosome engineering programs and main results are reviewed in King et al. (Chap. 18). While chromosome engineering holds great promises for transferring traits absent in cultivated wheat and potentially of major breeding impact on a major drawback is the linkage drag caused by the alien chromatin segments, often inducing negative features such as sterility, reduced seed germination, segregation distortion, anomalies of plant growth habit, etc., which often reduce grain size and other yield components.

The linkage drag effects are proportional to the segment size of the transferred chromatin. The transfer of alien chromatin in wheat through chromosome engineering generally involves first the transfer of single wide segments from the donor species, mainly through translocation. Additional local recombinations are induced to reduce the size of the alien chromatin around the target locus. This can be considered as a pre-breeding activity where a crucial role is played by the use of fluorescence *in situ* hybridization (FISH) PCR-based molecular markers functional in both

Triticum and the alien species and well distributed in the target region. The development of the high-density SNP array technology and subsequently, the KASP technology and the accumulation of massive genome sequence data allowed to specifically design probe sets for targeting and tracking introgressions from several wild relatives (Chap. 18).

28.11 Next Generation Sequencing (NGS) Technologies to Enhance MAS Effectiveness

The advancement in molecular technologies continuously affects how MAS can be implemented more cost-effectively. KASP panels covering the majority of the assays to target and trace beneficial alleles at the most relevant loci have been specifically designed for a first genome-wide characterization of the breeders' ready germplasm. At a higher throughput level, targeted Illumina-based re-sequencing of polymorphisms used for standard single-assay marker development has been proposed to streamline and increase the throughput of screening germplasm for the presence of beneficial alleles. The developed techniques are either targeted amplicon sequencing or direct multiplexed SNP interrogation, already offered by several private providers, combined with sample barcoding for efficient exploitation of the NGS sequencing capacity.

28.12 Integration Between MAS and Genomic Selection in Breeding Programs

While the concepts of MAS and Genomic Selection (GS) appear rather independent because they tap into two distinct portions of the wheat genome (see Fig. 28.3), the most successful genomic-assisted breeding programs combine both approaches in a synergic integration. Therefore, the role of MAS in pre-breeding is and will remain unique to rapidly introgress in breeding-compatible genetic stocks the new sources of variation made available through research and pre-breeding activities.

Once the novel beneficial alleles are introgressed and fixed in elite populations, this germplasm is ideal to implement genomic selection (GS) to efficiently tap into the plethora of minor QTLs. Due to its high efficiency, a well-managed GS program leads to selection and increase of the beneficial alleles more rapidly than conventional breeding programs. Hence, the importance of continuously refueling the program with novel beneficial allelic variants to be progressively cumulated into breeders' germplasm under selection (see Fig. 28.4).

28.13 Key Concepts

Future genetic gain in wheat will rely on a more effective application of marker-assisted selection and genomics approaches leveraging the fast-increasing capacity to sequence the entire wheat genome which will eventually provide a glimpse on the structural complexity of the wheat pangenome.

The effectiveness and success of genomics-assisted wheat breeding will depend on the following factors/issues:

- Availability of well-characterized germplasm collections capturing the biodiversity present worldwide in both tetraploid and hexaploid wheat and closely related species.
- Capacity to accurately phenotype, preferably in high-throughput fashion, large populations under controlled and field conditions.
- Apply (i) linkage mapping based on sufficiently large segregating populations and (ii) genome-wide association (GWA) mapping based on germplasm collections with low linkage disequilibrium (LD) decay to accurately dissect the QTLome and fine-map major QTLs of key proxy traits for yield and yield stability.
- Based on the above, implement MAS with the closest markers flanking the target locus.
- Deploy forward- and reverse-genetics approaches to clone the functional sequences governing the target trait. Cloning allows for the design of ‘perfect’ markers ideal for an error-free MAS.
- The availability of high-quality genome assemblies greatly facilitates the identification of candidate genes and the design of high-throughput and precise KASP markers diagnostic for inter-varietal and homelog-SNPs.
- The two domestication bottlenecks undergone by the A and B wheat genomes make tetraploid and durum wheat germplasm resources a particularly suitable biodiversity source to identify novel, underexploited beneficial alleles.
- Overall, the development of organized, informative, and user-friendly dedicated genomic databases is relevant for all the above-mentioned activities. The number and variety of discovered, marker-tagged, and cloned loci are already huge and the available scientific information is fragmented and not filtered by quality parameters. Databasing and database-interconnection are crucial aspects to be addressed.

28.14 Conclusions

Marker-assisted selection (MAS) started in parallel with the earliest achievements in genetic mapping and isolation of the most relevant loci for wheat biology, genetics, and improvement. Today, wheat breeding benefits from a full range of techniques and genomic resources, including the recently completed wheat pangenome,

for developing sequence-based molecular assays to enable high-throughput MAS. Importantly, the number of cloned wheat loci/QTLs, novel MAS protocols and genetic stocks developed in the last 5 years has grown steadily. As main achievements, the release of the reference gold standard wheat genome sequences paved the way to streamline genetic studies and MAS applications. Nowadays, the integrated and combined use of gene/QTL discovery, MAS in pre-breeding and breeding programs, together with genomic selection and gene editing are key for more effectively leveraging and bridging of biodiversity of the tetraploid with hexaploid A and B genomes while contributing to advance our knowledge in and understanding of wheat functional genomics.

References

1. Gupta PK, Balyan HS, Sharma S, Kumar R (2020) Genetics of yield, abiotic stress tolerance and biofortification in wheat (*Triticum aestivum* L.). *Theor Appl Genet* 133:1569–1602. <https://doi.org/10.1007/s00122-020-03583-3>
2. Langridge P, Reynolds M (2021) Breeding for drought and heat tolerance in wheat. *Theor Appl Genet* 134:1753–1769. <https://doi.org/10.1007/s00122-021-03795-1>
3. Collins NC, Tardieu F, Tuberosa R (2008) Quantitative trait loci and crop performance under abiotic stress: where do we stand? *Plant Physiol* 147:469–486. <https://doi.org/10.1104/pp.108.118117>
4. Gabay G, Zhang J, Burguener GF, Howell T, Wang H, Fahima T, Lukaszewski A, Moriconi JI, Santa Maria GE, Dubcovsky J (2021) Structural rearrangements in wheat (1BS)–rye (1RS) recombinant chromosomes affect gene dosage and root length. *Plant Genome* 14:e20079. <https://doi.org/10.1002/tpg2.20079>
5. Tuberosa R (2012) Phenotyping for drought tolerance of crops in the genomics era. *Front Physiol* 3:1–26. <https://doi.org/10.3389/fphys.2012.00347>
6. Winfield MO, Allen AM, BurrIDGE AJ, Barker GLA, Benbow HR, Wilkinson PA, Coghill J, Waterfall C, Davassi A, Scopes G, Pirani A, Webster T, Brew F, Bloor C, King J, West C, Griffiths S, King I, Bentley AR, Edwards KJ (2016) High-density SNP genotyping array for hexaploid wheat and its secondary and tertiary gene pool. *Plant Biotechnol J* 14:1195–1206. <https://doi.org/10.1111/pbi.12485>
7. Mazzucotelli E, Sciarra G, Mastrangelo AM, Desiderio F, Xu SS, Faris J, Hayden MJ, Tricker PJ, Ozkan H, Echenique V, Steffenson BJ, Knox R, Niane AA, Udupa SM, Longin FCH, Marone D, Petruzzino G, Corneti S, Ormanbekova D, Pozniak C, Roncallo PF, Mather D, Able JA, Amri A, Braum H, Ammar K, Baum M, Cattivelli L, Maccaferri M, Tuberosa R, Bassi FM (2020) The global durum wheat panel (GDP): an international platform to identify and exchange beneficial alleles. *Front Plant Sci* 11:2036. <https://doi.org/10.3389/fpls.2020.569905>
8. Maccaferri M, Harris NS, Twardziok SO, Pasam RK, Gundlach H, Spannagl M, Ormanbekova D, Lux T, Prade VM, Milner S, Himmelbach A, Mascher M, Bagnaresi P, Faccioli P, Cozzi P, Lauria M, Lazzari B, Stella A, Manconi A, Gnocchi M, Moscatelli M, Avni R, Deek J, Biyiklioglu S, Frascaroli E, Corneti S, Salvi S, Sonnante G, Desiderio F, Marè C, Crosatti C, Mica E, Özkan H, Kilian B, De Vita P, Marone D, Joukhadar R, Mazzucotelli E, Nigro D, Gadaleta A, Chao S, Faris JD, Melo A, Pumphrey M, Pecchioni N, Milanese L, Wiebe K, Ens J, MacLachlan RP, Clarke JM, Sharpe AG, Koh CS, Liang K, Taylor GJ, Knox R, Budak H, Mastrangelo AM, Xu SS, Stein N, Hale J, Distelfeld D, Hayden MJ, Tuberosa R, Walkowiak S, Mayer KFX, Ceriotti A, Pozniak CJ, Cattivelli L (2019) Durum wheat genome highlights past domestication signatures and future improvement targets. *Nat Genet* 51:885–895. <https://doi.org/10.1038/s41588-019-0381-3>

9. Adhikari S, Saha S, Biswas A, Rana TS, Bandyopadhyay TK, Ghosh P (2017) Application of molecular markers in plant genome analysis: a review. *Nucleus* 60:283–297. <https://doi.org/10.1007/s13237-017-0214-7>
10. Salvi S, Tuberosa R (2015) The crop QTLome comes of age. *Curr Opin Biotechnol* 32:179–185. <https://doi.org/10.1016/j.copbio.2015.01.001>
11. Venske E, dos Santos RS, Farias D d R, Rother V, da Maia LC, Pegoraro C, de Oliveira A (2019) Meta-analysis of the QTLome of Fusarium head blight resistance in bread wheat: refining the current puzzle. *Front Plant Sci* 10:727. <https://doi.org/10.3389/fpls.2019.00727>
12. Shariatipour N, Heidari B, Richards CM (2021) Meta-analysis of QTLome for grain zinc and iron contents in wheat (*Triticum aestivum* L.). *Euphytica* 217:86. <https://doi.org/10.1007/s10681-021-02818-8>
13. Maccaferri M, Canè MA, Sanguineti MC, Salvi S, Colalongo MC, Massi A, Clarke F, Knox R, Pozniak CJ, Clarke JM, Fahima T, Dubcovsky J, Xu S, Ammar K, Karsai I, Vida G, Tuberosa R (2014) A consensus framework map of durum wheat (*Triticum durum* Desf.) suitable for linkage disequilibrium analysis and genome-wide association mapping. *BMC Genomics* 15:873. <https://doi.org/10.1186/1471-2164-15-873>
14. Quraishi UM, Pont C, Ain Q, Flores R, Burlot L, Alaux M, Quesneville H, Salse J (2017) Combined genomic and genetic data integration of major agronomical traits in bread wheat (*Triticum aestivum* L.). *Front Plant Sci* 8:1843. <https://doi.org/10.3389/fpls.2017.01843>
15. The International Wheat Genome Sequencing Consortium, Appels R, Eversole K, Stein N, Feuillet C, Keller B, Rogers J, Pozniak CJ, Choulet F, Distelfeld A, Poland J, Ronen G, Sharpe AG, Barad O, Baruch K, Keeble-Gagnère G, Mascher M, Ben-Zvi G, Josselin A-A, Himmelbach A, Balfourier F, Gutierrez-Gonzalez J, Hayden M, Koh C, Muehlbauer G, Pasam RK, Paux E, Rigault P, Tibbits J, Tiwari V, Spannagl M, Lang D, Gundlach H, Haberer G, Mayer KFX, Ormanbekova D, Prade V, Šimková H, Wicker T, Swarbreck D, Rimbart H, Felder M, Guilhot N, Kaithakottil G, Keilwagen J, Leroy P, Lux T, Twardziok S, Venturini L, Juhász A, Abrouk M, Fischer I, Uauy C, Borrill P, Ramirez-Gonzalez RH, Arnaud D, Chalabi S, Chalhoub B, Cory A, Datla R, Davey MW, Jacobs J, Robinson SJ, Steuernagel B, van Ex F, Wulff BBH, Benhamed M, Bendahmane A, Concia L, Latrasse D, Bartoš J, Bellec A, Berges H, Doležel J, Frenkel Z, Gill B, Korol A, Letellier T, Olsen O-A, Singh K, Valárik M, van der Vossen E, Vautrin S, Weining S, Fahima T, Glikson V, Raats D, Čížková J, Toegelová H, Vrána J, Sourdille P, Darrier B, Barabaschi D, Cattivelli L, Hernandez P, Galvez S, Budak H, Jones JDG, Witek K, Yu G, Small I, Melonek J, Zhou R, Belova T, Kanyuka K, King R, Nilsen K, Walkowiak S, Cuthbert R, Knox R, Wiebe K, Xiang D, Rohde A, Golds T, Čížková J, Akpinar BA, Biyikliglu S, Gao L, N'Daiye A, Kubaláková M, Šafář J, Alfama F, Adam-Blondon A-F, Flores R, Guerche C, Loaec M, Quesneville H, Condie J, Ens J, Maclachlan R, Tan Y, Alberti A, Aury J-M, Barbe V, Couloux A, Cruaud C, Labadie K, Mangenot S, Wincker P, Kaur G, Luo M, Sehgal S, Chhuneja P, Gupta OP, Jindal S, Kaur P, Malik P, Sharma P, Yadav B, Singh NK, Khurana JP, Chaudhary C, Khurana P, Kumar V, Mahato A, Mathur S, Sevanthi A, Sharma N, Tomar RS, Holušová K, Plíhal O, Clark MD, Heavens D, Kettleborough G, Wright J, Balcárková B, Hu Y, Salina E, Ravin N, Skryabin K, Beletsky A, Kadnikov V, Mardanov A, Nesterov M, Rakitin A, Sergeeva E, Handa H, Kanamori H, Katagiri S, Kobayashi F, Nasuda S, Tanaka T, Wu J, Cattonaro F, Jiumeng M, Kugler K, Pfeifer M, Sandve S, Xun X, Zhan B, Batley J, Bayer PE, Edwards D, Hayashi S, Tulpová Z, Visendi P, Cui L, Du X, Feng K, Nie X, Tong W, Wang L (2018) Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361:eaar7191. <https://doi.org/10.1126/science.aar7191>
16. Varshney RK, Tuberosa R (2007) Genomics-assisted crop improvement. Springer
17. Visscher PM, Goddard ME (2019) From R.A. Fisher's 1918 paper to GWAS a century later. *Genetics* 211:1125–1130. <https://doi.org/10.1534/genetics.118.301594>
18. Uauy C, Distelfeld A, Fahima T, Blechl A, Dubcovsky J (2006) A NAC gene regulating senescence improves grain protein, zinc, and iron content in wheat. *Science* 314:1298–1301. <https://doi.org/10.1126/science.1133649>
19. Pallotta M, Schnurbusch T, Hayes J, Hay A, Baumann U, Paull J, Langridge P, Sutton T (2014) Molecular basis of adaptation to high soil boron in wheat landraces and elite cultivars. *Nature* 514:88–91. <https://doi.org/10.1038/nature13538>

20. Rawat N, Pumphrey MO, Liu S, Zhang X, Tiwari VK, Ando K, Trick HN, Bockus WW, Akhunov E, Anderson JA, Gill BS (2016) Wheat Fhb1 encodes a chimeric lectin with agglutinin domains and a pore-forming toxin-like domain conferring resistance to Fusarium head blight. *Nat Genet* 48:1576–1580. <https://doi.org/10.1038/ng.3706>
21. Sallam AH, Conley E, Prakapenka D, Da Y, Anderson JA (2020) Improving prediction accuracy using multi-allelic haplotype prediction and training population optimization in wheat. *G3 Genes/Genomes/Genetics* 10:2265–2273. <https://doi.org/10.1534/g3.120.401165>
22. Rasheed A, Xia X (2019) From markers to genome-based breeding in wheat. *Theor Appl Genet* 132:767–784. <https://doi.org/10.1007/s00122-019-03286-4>
23. Beyer S, Daba S, Tyagi P, Bockelman H, Brown-Guedira G, Mohammadi M, IWGSC (2019) Loci and candidate genes controlling root traits in wheat seedlings—a wheat root GWAS. *Funct Integr Genomics* 19:91–107. <https://doi.org/10.1007/s10142-018-0630-z>
24. Singh K, Batra R, Sharma S, Saripalli G, Gautam T, Singh R, Pal S, Malik P, Kumar M, Jan I, Singh S, Kumar D, Pundir S, Chaturvedi D, Verma A, Rani A, Kumar A, Sharma H, Chaudhary J, Kumar K, Kumar S, Singh VK, Singh VP, Kumar S, Kumar R, Gaurav SS, Sharma S, Sharma PK, Balyan HS, Gupta PK (2021) WheatQTLdb: a QTL database for wheat. *Mol Gen Genomics*. <https://doi.org/10.1007/s00438-021-01796-9>
25. Tuberosa R, Frascaroli E, Maccaferri M, Salvi S (2021) Understanding and exploiting the genetics of plant root traits. In: Gregory PJ (ed) *Understanding and improving crop root function*. Burleigh Dodds Science Publishing, Cambridge
26. Ober ES, Alahmad S, Cockram J, Forestan C, Hickey LT, Kant J, Maccaferri M, Marr E, Milner M, Pinto F, Rambla C, Reynolds M, Salvi S, Sciarra G, Snowdon RJ, Thomelin P, Tuberosa R, Uauy C, Voss-Fels KP, Wallington E, Watt M (2021) Wheat root systems as a breeding target for climate resilience. *Theor Appl Genet* 134:1645–1662. <https://doi.org/10.1007/s00122-021-03819-w>
27. Hao C, Jiao C, Hou J, Li T, Liu H, Wang Y, Zheng J, Liu H, Bi Z, Xu F, Zhao J, Ma L, Wang Y, Majeed U, Liu X, Appels R, Maccaferri M, Tuberosa R, Lu H, Zhang X (2020) Resequencing of 145 landmark cultivars reveals asymmetric sub-genome selection and strong founder genotype effects on wheat breeding in China. *Mol Plant* 13:1733–1751. <https://doi.org/10.1016/j.molp.2020.09.001>
28. Maccaferri M, Ratti C, Rubies-Autonell C, Vallega V, Demontis A, Stefanelli S, Tuberosa R, Sanguineti MC (2011) Resistance to soil-borne cereal mosaic virus in durum wheat is controlled by a major QTL on chromosome arm 2BS and minor loci. *Theor Appl Genet* 123:527–544. <https://doi.org/10.1007/s00122-011-1605-9>
29. Terracciano I, Maccaferri M, Bassi F, Mantovani P, Sanguineti MC, Salvi S, Šimková H, Doležel J, Massi A, Ammar K, Kolmer J, Tuberosa R (2013) Development of COS-SNP and HRM markers for high-throughput and reliable haplotype-based detection of *Lr14a* in durum wheat (*Triticum durum* Desf.). *Theor Appl Genet* 126:1077–1101. <https://doi.org/10.1007/s00122-012-2038-9>
30. Wang S, Wong D, Forrest K, Allen A, Chao S, Huang BE, Maccaferri M, Salvi S, Milner SG, Cattivelli L, Mastrangelo AM, Whan A, Stephen S, Barker G, Wieseke R, Plieske J, Consortium IWGS, Lillemo M, Mather D, Appels R, Dolferus R, Brown-Guedira G, Korol A, Akhunova AR, Feuillet C, Salse J, Morgante M, Pozniak C, Luo M-C, Dvorak J, Morell M, Dubcovsky J, Ganal M, Tuberosa R, Lawley C, Mikoulitch I, Cavanagh C, Edwards KJ, Hayden M, Akhunov E (2014) Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnol J* 12:787–796. <https://doi.org/10.1111/pbi.12183>
31. Maccaferri M, Ricci A, Salvi S, Milner SG, Noli E, Martelli PL, Casadio R, Akhunov E, Scalabrini S, Vendramin V, Ammar K, Blanco A, Desiderio F, Distelfeld A, Dubcovsky J, Fahima T, Farris J, Korol A, Massi A, Mastrangelo AM, Morgante M, Pozniak C, N'Diaye A, Xu S, Tuberosa R (2015) A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. *Plant Biotechnol J* 13:648–663. <https://doi.org/10.1111/pbi.12288>

32. van Poecke RMP, Maccaferri M, Tang J, Truong HT, Janssen A, van Orsouw NJ, Salvi S, Sanguineti MC, Tuberosa R, van der Vossen EAG (2013) Sequence-based SNP genotyping in durum wheat. *Plant Biotechnol J* 11:809–817. <https://doi.org/10.1111/pbi.12072>
33. Sun C, Dong Z, Zhao L, Ren Y, Zhang N, Chen F (2020) The Wheat 660K SNP array demonstrates great potential for marker-assisted selection in polyploid wheat. *Plant Biotechnol J* 18:1354–1360. <https://doi.org/10.1111/pbi.13361>
34. Gupta PK, Langridge P, Mir RR (2010) Marker-assisted wheat breeding: present status and future possibilities. *Mol Breed* 26:145–161. <https://doi.org/10.1007/s11032-009-9359-7>
35. Ayalew H, Tsang PW, Chu C, Wang J, Liu S, Chen C, Ma X-F (2019) Comparison of TaqMan, KASP and rhAmp SNP genotyping platforms in hexaploid wheat. *PLoS One* 14:1–9. <https://doi.org/10.1371/journal.pone.0217222>
36. Yuan J, Wen Z, Gu C, Wang D (2014) Introduction of high throughput and cost effective SNP genotyping platforms in soybean. *Plant Genet Genomics Biotechnol* 2:90–94
37. Ramirez-Gonzalez RH, Uauy C, Caccamo M (2015) PolyMarker: a fast polyploid primer design pipeline. *Bioinformatics* 31:2038–2039. <https://doi.org/10.1093/bioinformatics/btv069>
38. Makhoul M, Rambla C, Voss-Fels KP, Hickey LT, Snowdon RJ, Obermeier C (2020) Overcoming polyploidy pitfalls: a user guide for effective SNP conversion into KASP markers in wheat. *Theor Appl Genet* 133:2413–2430. <https://doi.org/10.1007/s00122-020-03608-x>
39. Kumar A, Saripalli G, Jan I, Kumar K, Sharma PK, Balyan HS, Gupta PK (2020) Meta-QTL analysis and identification of candidate genes for drought tolerance in bread wheat (*Triticum aestivum* L.). *Physiol Mol Biol Plants* 26:1713–1725. <https://doi.org/10.1007/s12298-020-00847-6>
40. Thudi M, Palakurthi R, Schnable JC, Chitkineni A, Dreisigacker S, Mace E, Srivastava RK, Satyavathi CT, Odeny D, Tiwari VK, Lam H-M, Bin HY, Singh VK, Li G, Xu Y, Chen X, Kaila S, Nguyen H, Sivasankar S, Jackson SA, Close TJ, Shubo W, Varshney RK (2021) Genomic resources in plant breeding for sustainable agriculture. *J Plant Physiol* 257:153351. <https://doi.org/10.1016/j.jplph.2020.153351>
41. Ma Z, Xie Q, Li G, Jia H, Zhou J, Kong Z, Li N, Yuan Y (2020) Germplasms, genetics and genomics for better control of disastrous wheat Fusarium head blight. *Theor Appl Genet* 133:1541–1568. <https://doi.org/10.1007/s00122-019-03525-8>
42. Adamski NM, Borrill P, Brinton J, Harrington SA, Marchal C, Bentley AR, Bovill WD, Cattivelli L, Cockram J, Contreras-Moreira B, Ford B, Ghosh S, Harwood W, Hassani-Pak K, Hayta S, Hickey LT, Kanyuka K, King J, Maccaferri M, Naamati G, Pozniak CJ, Ramirez-Gonzalez RH, Sansaloni C, Trevasik B, Wingen LU, Wulff BBH, Uauy C (2020) A roadmap for gene functional characterisation in crops with large genomes: lessons from polyploid wheat. *elife* 9:e55646. <https://doi.org/10.7554/eLife.55646>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

