Efficient Breeding of Crop Plants

14

Pawan L. Kulwal, Reyazul Rouf Mir, and Rajeev K. Varshney

Abstract

In crop breeding programs, the rate of genetic gain which is achieved using the traditional breeding is insufficient to meet the increased demand of food for the rapidly expanding global population. The main constraint with the conventional breeding is the time which is required in developing crosses, followed by selection and testing of the experimental cultivars. Although, using this technique, lot of progress has been made in increasing the productivity, the time has come to think beyond this and integrate the recent advances in the area of genomics, phenomics and computational biology into the conventional breeding program for increasing its efficiency. While doing this emphasis on proper characterization and use of plant genetic resources, defining the breeding objectives and use of recent advances in holistic way are also essential. Therefore, in this chapter, we first highlight the importance of plant breeding followed by significance of the plant genetic resources in the breeding program, need of ideotype breeding and the breeding objectives for important traits including resistance against various biotic and abiotic stresses. We then discuss the

---

P. L. Kulwal (✉)
Biotechnology Centre, Mahatma Phule Krishi Vidyapeeth (Agricultural University), Rahuri, Maharashtra, India

R. R. Mir
Division of Genetics and Plant Breeding, Faculty of Agriculture, SKUAST-Kashmir, Kashmir, Jammu and Kashmir, India

R. K. Varshney
Centre of Excellence in Genomics and Systems Biology, International Crops Research Institute for the Semi-Arid Tropics, Patancheru, Hyderabad, Telangana, India

State Agricultural Biotechnology Centre, Crop Research Innovation Centre, Food Futures Institute, Murdoch University, Murdoch, WA, Australia

D. K. Yadava et al. (eds.), *Fundamentals of Field Crop Breeding*,
https://doi.org/10.1007/978-981-16-9257-4_14

limitations of conventional breeding and advantages of genomics-assisted breeding. While doing this, we also discuss various molecular breeding tools and genomic resources as well as different approaches for efficient breeding including marker-assisted selection, marker-assisted recurrent selection and genomic selection. This is followed by importance of other non-conventional approaches including the recent one on gene editing, speed breeding and role of integrated data management and bioinformatics in the breeding programs. We also discuss the significance of phenomics and phenotyping platforms in the crop breeding as well as role of computational techniques like artificial intelligence and machine learning in analysing the huge data which is being generated in the breeding programs. Finally, we conclude with a brief note on the emerging challenges in breeding which need to be addressed and the thrust areas of research for the future.

Keywords

Plant breeding · Efficient breeding · Ideotype breeding · Genomics · Plant genetic resources · Phenomics · Gene editing · Artificial intelligence

14.1 Importance of Crop Breeding

The history of plant breeding has been very promising, and plant breeding has contributed immensely in consistently increasing the productivity of the crop plants during the last several centuries. This has helped in addressing the issue of hunger and malnutrition to larger extent. The achievements of plant breeding have been manifold and have resulted in tremendous increase in yield either by directly improving the yield component traits or through resistance breeding. The Green Revolution is the best example of this, through which the productivity of cereals has increased remarkably by introducing dwarf genes into wheat varieties responsive to fertilizers (Tester and Langridge 2010). In several countries including India, self-sufficiency has been achieved in almost all the major crops in the last few decades using the practices of conventional breeding coupled with better agronomic practices. However, these methods are not sufficient to produce enough food for the growing population in the coming years, particularly with the growing challenges of environmental changes (Godfray et al. 2010; Tester and Langridge 2010).

In order to achieve the breeding objectives, plant breeders often use various methods and techniques. Although the application of techniques in conventional breeding is often straight forward, they need to be refined from time to time according to the existing needs. It is under such circumstances that modern and efficient breeding approaches involving use of genomics tools should become part of the breeding programs (Kulwal et al. 2012). However, before implementing any such technique, it is essential that the breeding objectives are clearly defined.

The science of plant breeding has evolved over the years. Recently, the timeline of the plant breeding has been partitioned into four different phases (Ramstein et al. 2019). Breeding 1.0 focused on selection with unknown loci during the first 10,000 years before 1900 (early domestication phase), Breeding 2.0 relating to the Mendelian genetics and selection by controlled crosses, Breeding 3.0 involving use of genomics techniques like marker-assisted breeding and Breeding 4.0 focusing on ideotype-based selection and use of techniques like genetic transformation and gene editing (Wallace et al. 2018; Ramstein et al. 2019). Broadly speaking, we can categorize Breeding 1 and 2 as conventional approaches while Breeding 3 and 4 as efficient approaches. While we are already in the phase 4.0, the real potential of genomics tools (phase 3.0) has not yet been fully exploited in all the crops. Nevertheless, the success these efficient breeding techniques have shown (see later) has provided new dimensions to the crop improvement programs.

14.2 Importance of Plant Genetic Resources in Crop Breeding

Efficiency of any breeding program is directly dependent on the availability of the plant genetic resources (PGRs) for that crop. These PGRs include cultivated and wild species and valuable germplasm including landraces (Bošković et al. 2012). In order to deal with the complex traits and to respond to the challenges of climate change, particularly biotic and abiotic stresses, breeders need to exploit as much genetic diversity as they can (Galluzzi et al. 2020). This diversity can be exploited through use of the PGRs (landraces, crop wild relatives and other available germplasm) which are generally conserved in the gene banks. Often, the breeders use superior performing lines, albeit with limited diversity in their breeding programs. Although substantial progress has been made for yield improvement with this approach, the time has come when breeders need to exploit the true potential of the PGRs to address the challenges of the future. There are different ways through which PGRs are used in crop breeding programs. However, this depends on the breeding objective, the trait being studied and the genetic resources available (Bidinger 1992).

The three most common ways through which PGRs are being used in the conventional breeding are the (1) introgression (transfer of one or few genes from PGRs), (2) incorporation (widening of the genetic base) and (3) pre-breeding (use of exotic materials or wild relatives) (Haussmann et al. 2004). Besides this, PGRs can also be effectively used (1) as a parental genotype in developing a mapping population to be used in QTL mapping and (2) simultaneous identification and transfer of desirable QTLs from wild or unadapted germplasm into the cultivated one through advanced backcross QTL mapping and (3) for identification of genes/QTLs through the approach of association mapping/genome-wide association studies (GWAS) and allele mining (Kulwal 2016; Kulwal and Singh 2021). Large numbers of studies have been carried out in different crops, and hundreds of marker-trait associations (MTAs) have been identified for a variety of traits through GWAS (Gupta et al. 2019). In situations when all the PGRs for a crop cannot be utilized for GWAS, then

Table 14.1 Plant genetic resources available in different gene banks of CGIAR institutes

S. no.	Crop	Institute	Number of accessions
1	Wheat	CIMMYT and ICARDA	199,248
2	Rice	IRRI and Africa Rice	151,765
3	Barley	ICARDA	32,790
4	Maize	CIMMYT and IITA	30,055
5	Sorghum	ICRISAT	41,582
6	Bajra	ICRISAT	23,841
7	Small millets	ICRISAT	11,797
8	Chickpea	ICRISAT and ICARDA	36,513
9	Pigeon pea	ICRISAT	13,783
10	Soybean	IITA	4841
11	Groundnut	ICRISAT	15,622
12	Forages	CIAT, ICARDA and ILRI	70,514

core and mini-core collections are developed by identifying most diverse but representative set of germplasm from the total collection for further analysis.

There are quite a few examples where PGRs from the gene banks have been used to identify and introgress important genes/QTLs for abiotic stresses in different crops. For instance, in rice, a landrace FR13A was identified for its resilience to complete submergence. The submergence QTL (*Sub1*) from its derivatives has successfully been introduced into several lines leading to the development of many rice varieties with improved tolerance to submergence (Bailey-Serres et al. 2010). Similarly, in chickpea genotype ICC 4958 available at the ICRISAT gene bank was identified as having a profuse root system and was used in many QTL mapping studies which had led to the identification of *QTL-hotspot* for drought tolerance-related traits (Varshney et al. 2014). This QTL region was subsequently transferred in the genetic background of chickpea variety Pusa-372 leading to the development and release of drought-tolerant variety Pusa Chickpea-10,216 following the approach of marker-assisted selection (MAS). In addition, the *QTL-hotspot* has been introgressed in several genetic background, and many introgression lines have been developed (Roorkiwal et al. 2020; Bharadwaj et al. 2021). These examples clearly demonstrate the vital role PGRs play in the plant breeding programs. There are many more such examples in different crops where PGRs have been used for the improvement of the varieties.

The gene banks are invaluable treasures which contain large number of PGRs and provide an opportunity to utilize these resources to breed for the uncertainties posed by climate change (Bohra et al. 2020; Khan et al. 2020). Large numbers of accessions have been conserved at different gene banks across the world. These include major gene banks of CGIAR (https://www.genebanks.org) which house thousands of accessions collected from world over for the important crops (Table 14.1). The number of accessions per crop which are conserved in different gene banks in the CGIAR institutes is also available at https://www.genebanks.org/resources/crops/ (verified March 04, 2021). In addition, there are several national gene banks in different countries which also conserve different accessions. In India, National Bureau of Plant Genetic Resources (NBPGR) at New Delhi is the main gene bank entrusted with this responsibility. In addition, there are several crop-specific gene banks at different institutes of ICAR. While conserving the germplasm is an important step, it is equally important to characterize and utilize these PGRs effectively for use in the crop improvement programs for addressing the issue of the world's future food requirements (McCouch et al. 2020). In order to utilize this treasure of PGRs available with the national gene bank in India, recently the Department of Biotechnology, Government of India, has funded several projects for the utilization of these germplasm for their characterization and use in the breeding programs. There is no doubt that utilization of these PGRs in the breeding programs can help in achieving the Sustainable Development Goals as outlined by the United Nations (Halewood et al. 2018).

14.3 Ideotype Breeding

The concept of ideotype breeding was proposed for the first time by Donald (1968) in wheat. The concept given by Donald presumed that most of the plant breeding programs are based on two main objectives including (1) defect elimination and (2) selection for yield. The term defect elimination is appropriately used when the disease resistance trait is introduced into a susceptible variety or when early maturity trait is introduced into a variety prone to water stress in late season (Donald 1968). The defect elimination involves correction of physical imperfections like fragile skin in tomatoes, weak malting performance in barley, poor flavour in potatoes and weak straw in cereals. The elimination of these defects ultimately led to the increase in yield and quality. On the other hand, selection and improving yield is considered ultimate objective of any breeding program in the world. The improvement of yield can be done through hybridization that involves crossing of superior genotypes having broad genetic base. Therefore, in summary, ideotype-based approach in plant breeding improves the efficiency of the selection (Gauffreteau 2018). The process of ideotype breeding involves (1) defining the varietal specifications; (2) designing and building an ideotype; and (3) selecting varieties according to the ideotype and assessing their ability to meet the specifications.

In ideotype breeding, the theoretical models are being developed based on our understanding, knowledge, experience and imagination. The models thus developed are known as "the breeding of model plants or ideotypes". Through this type of breeding, it is possible to design a plant that is capable of greater production in a target environment than the genotype it is to replace. For efficient ideotype breeding, we need to choose model characters. These model characters are presumably very important. For instance, in cereals stout stem is a "model character" and provides lodging resistance. Similarly, presence of awns is another model character and has a role in photosynthesis and yield. Another model character is erect foliage (upright leaves) which have shown advantage in photosynthesis. The concept of crop ideotype initially formulated for wheat crop by Donald has been used for several other crop species including barley (Rasmusson 1987), chickpea (Siddique and Sedgley 1987), forest trees (Dickmann 1985) such as spruce and pine (Kärki and Tigerstedt 1985) and fruit trees like mango and apple (Dickmann et al. 1994). The attributes of ideotypes are morphological characters based on physiological consideration. The concept of ideotype first relied primarily on the morphological traits. However, later the ideotype definition was extended, and traits like physiological, biochemical, anatomical and phenological traits have been included in defining a particular crop ideotype. However, before including any new parameter or a trait for a plant type, it is important to check its contribution towards crop yield. Thus, study and understanding the concept of plant ideotype for a given crop can prove to be an efficient in the breeding of that crop.

14.4 Breeding Objectives for Important Traits

14.4.1 Breeding Objectives for Yield

Grain yield is one of the most complex quantitative traits known in crop plants, and this trait is significantly influenced by the environment. The genetics of grain yield over the years suggested that grain yield is controlled by large number of genes/ QTLs with minor effect. The gene controlling grain yield is often influenced by the environment. Therefore, it is important to study genotype × environment interaction while studying genetics of this complex quantitative trait. In addition, extensive studies of grain yield have uncovered that genes responsible for grain yield are often interacting with other genes. This phenomenon is popularly described as “epistasis”. Different kinds of gene × gene interactions and gene × gene × environment interactions have also been discovered for grain yield. The complexity of grain yield as a trait can be realized by the fact that grain yield can be dissected into several component traits with higher heritability. In addition, it has been reported in earlier studies that individual traits showing correlation with grain yield are most often controlled by the same set of QTLs/genes.

The most important component traits of grain yield included number of spikelet’s/spike, number of grains per spike and 1000 grain/kernel weight (Mir et al. 2021b). For example, in bread wheat, the important reproductive organ harbouring grains is the spike, and therefore traits related to wheat spike are considered very important for manipulating grain yield. The published literature showed strong and positive correlation between different spike traits like spike length with grain yield and yield-related traits including 1000 grain/kernel weight (Mir et al. 2012a). Therefore, from a breeding point of view, genes/QTLs already identified or to be identified in the future for traits related to wheat spike are of importance for wheat molecular breeding programs aimed at enhancing grain yield. The domestication genes/loci like *Q*, *C* and *S* are also considered important and influence several yield component traits including rachis fragility, spike length, spike compactness, grain morphology traits, grain number, seed shape and glumes of a wheat spike.

The genetic dissection of grain yield has been attempted using a variety of approaches including QTL interval mapping and association mapping and have led to the discovery of genes/QTLs for grain yield component traits in several crops. The QTL/genes that have been identified over years need to be deployed into breeding programs through modern and efficient breeding methods including MAS (see later).

14.4.2 Grain Quality Characters

Grain quality traits are very important in any crop because the market price of the end produce is directly related to the quality parameters. A lot of emphasis is thus being given on the improvement of the quality parameters. While there are some quality parameters which are common to majority of the grain crops (e.g. grain size, shape, protein content, etc.), some are unique (e.g. pre-harvest sprouting tolerance in wheat, sorghum, mungbean; milling quality in rice, etc.). List of some important quality traits in different crops is given in Table 14.2. While a lot of breeding efforts have been made for the improvement yield of crop plant, it is now necessary that focus should be shifted towards quality parameters. Breeding for quality parameters is difficult because their nature of inheritance is complex and they are controlled by polygenes. Moreover, measuring the trait precisely is most of the time technically demanding and expensive. However, with the development of novel, rapid and cost-effective screening techniques, breeders have taken renewed interest in the breeding for quality parameters (Munck 2009).

Any breeding program aimed at improving the quality traits relies most importantly on the availability of the known sources containing genes for these traits (Varshney et al. 2021). The PGRs (germplasm, landraces, wild relatives) can prove very useful in this regard. Genes/QTLs for these traits can be identified in a mapping population developed for this purpose by crossing two genotypes differing for the trait or using germplasm collection and natural population. Several studies have been carried out in different crops where QTLs have been identified for important quality

Table 14.2 Key important quality traits in different crops

S. no.	Crop	Quality traits
1	Wheat	Grain protein content; gluten strength; pre-harvest sprouting tolerance; grain size; grain weight; kernel colour; dough making properties
2	Rice	Grain size; grain shape; aroma; whiteness; long and thin uncooked grains; amylose content; milling, physical appearance, cooking, sensory, palatability, and nutritional value
3	Maize	Protein content; amino acid content; fatty acid content; starch quality
4	Sorghum	Amylose content; starch content; crude protein; gross energy; tannin content; polyphenol content
5	Chickpea	Protein content; grain size; fibre content; carotenoid composition
6	Soybean	Seed weight; seed protein; sucrose concentration; oil content; oil composition

traits (Kulwal et al. 2005, 2010). Some of these QTLs have also been transferred into the desirable genotypes using efficient breeding programs like MAS (Varshney et al. 2021). However, success of any such program also depends on the variation explained by the identified QTL. One of the concerns in breeding for quality traits is that it is often considered that improvement in some trait (e.g. protein content) comes with yield penalty, although not always (Kulwal and Mhase 2017). For this purpose, understanding the nature of gene action and genotype × environment interaction is also essential. In recent years, the techniques like genetic engineering and gene editing offer tremendous scope for improving the quality.

14.4.3 Resistance to Biotic Stresses

Biotic stresses (pests and diseases) are known to cause significant losses to the crop plants resulting in reduction in the yield as well as quality of the produce. Depending on the existing circumstances, the degree of losses caused can vary from minor to severe. It has been reported that almost half of the total yield losses in the world are due to biotic stresses (Balconi et al. 2012). Therefore, breeding for biotic stresses is the major objective in any crop improvement program. While use of chemical pesticides is effective in controlling these stresses, they can add to the cost of production and cause significant damage not only to human and animal being but to the environment as well (Bainsla and Meena 2016). With the issues concerning climate change, the threat of new pests and diseases has also increased in recent years. In order to address this issue in a sustainable way, use of resistance sources in the breeding program coupled with deployment of efficient and precise techniques like MAS, marker-assisted gene pyramiding, genetic engineering and gene editing are very promising. MAS not only has simplified the breeding programs but has also accelerated the process of gene transfer in the desired genetic background (see later). The sources of resistance can be identified by screening large number of germplasm accessions against particular disease/pest. Exploitation of PGR can prove useful under such circumstances.

Although the conventional breeding relying on the principle of back crossing using resistant sources has proved successful in incorporating resistance, it is time-consuming and not very efficient technique. Other limitation is that it is effective only when the *R* genes are used one at a time, thus limiting gene pyramiding. There are many successful examples where the technique of MAS has successfully been used to transfer resistance genes/QTLs against various pests and diseases in the crop plants as well as in gene pyramiding (Dormatey et al. 2020). It has also been reported that the resistance achieved through pyramiding of the resistant QTL is at par or better than that conferred by the *R* genes (Richardson et al. 2006). However, MAS is useful only when the resistance source which is being used in the crossing program is compatible with the recipient genotype. In cases when the compatible resistant sources are not available, the technique like genetic engineering and gene editing (see later) can prove very effective (Dong and Ronald 2019). Several successful examples are available where these techniques have been used for incorporating resistance against particular pest (Varshney et al. 2021). One thing which is very important in resistance breeding is that with the constant pressure of emerging pests and diseases, the breeders need to keep themselves well equipped with the recent tools and one step ahead of the pest and disease in question so that new varieties with durable resistance can be developed.

14.4.4 Resistance to Abiotic Stresses

The yield of crop plants is mostly influenced by a variety of biotic and abiotic stresses, climatic and agronomic factors. The major abiotic stresses limiting crop yields include drought, heat, cold, freezing, salinity and metal stresses. All these abiotic stresses create adverse effects on morphology, physiology and biochemistry which ultimately lead to the adverse effects on growth and yield of plants. Both vegetative and reproductive phases of plant growth are influenced by the abiotic stresses like heat, drought, cold and freezing. It is estimated that an average of 50% yield losses in agricultural crops is caused by abiotic stresses, and among these abiotic stresses high temperature stress causes 40% loss followed by salinity (20%), drought (17%) and cold/freezing (15%) (Meena et al. 2017). For example, in chickpea, one of the most important grain legume crops in the world the significant economic losses due to drought/heat (1.3 billion US dollars), cold (186 million US dollars) and salinity (354 million US dollars) have raised major concerns among the chickpea-growing countries (Jha et al. 2014; Mir et al. 2021a). This situation is further exacerbated by climate change which may cause higher intensities and frequencies of abiotic stresses, thereby necessitating the identification and development of climate-resilient cultivars having region-specific traits, which can perform well under stress.

Several physiological changes are induced in plants in response to different abiotic stresses. For instance, the physiological changes induced in plants in response to abiotic stress included wilting and abscission of the leaf, transpiration, etc. The most important abiotic stress “drought” results in decrease in turgor pressure and thus affects cell growth. It also creates disruption to water flow from xylem to neighbouring elongating cells and thus stops cell elongation. The visible changes include reduction in leaf area as well as plant height. The different mechanisms adapted by plants during abiotic stress include escape, avoidance and tolerance mechanisms. Like drought and heat stress, cold and freezing stress also adversely affects plant growth and development, membrane structure and photosynthetic activity (Mir et al. 2021a). Low temperature stress is an important issue for winter-sown crops like chickpea, lentil, pea, wheat, barley, etc. in the countries surrounding the Mediterranean Sea, the tropical highlands and temperate growing areas. The most affected regions are northern South Asia and parts of the Australia, where crop faces low-temperature stress (<15 °C) which limits growth and vigour at all phenological stages but particularly during vegetative and reproductive stages leading to chlorosis and necrosis of leaf tips, substantial loss of flowers and pod abortion, reduced pollen viability and pollen tube growth and, thus, reduced seed quality and yield potential by 30–40% (Rani et al. 2020; Mir et al. 2021a).

Keeping in view the losses due to abiotic stresses, the breeding for climatic resilient varieties having tolerance against abiotic stresses has emerged as one of the important subject areas of crop research and a major goal in plant breeding programs worldwide (Meena et al. 2017). Different breeding methods have been used to breed for different abiotic stresses in different crop plants. The methods of breeding adopted in different crop plants depend on their mode of reproduction (self-pollination, cross-pollination or asexual) and the genetic control of trait of interest. The breeding methods for drought are usually the same as that of yield. In general, the conventional breeding methods like pedigree and bulk method could be used for self-pollinated crop species, and recurrent selection could be used for cross-pollinated crop species.

In some cases if one wishes to transfer more than one trait related to drought into a high yielding variety, then back cross breeding method is considered the most appropriate. The varieties for salinity tolerance were developed through pedigree, modified bulk pedigree and anther culture approach. The different genetic resources including landraces, wild relatives, released varieties, pre-breeding lines, advanced breeding lines and mutants also serve as important sources of resistance against abiotic stresses. Wild species of crop plants serve as an important source and reservoir of genes for abiotic stress tolerance. For instance, in wheat crop, *Aegilops kotschyi*, *Ae. Squarrosa* and *Triticum urartu* serve as an excellent source of drought tolerance and *Aegilops tauschii* as source of salinity tolerance. In chickpea, *Cicer microphyllum*, *Cicer reticulatum* and *Cicer echinospermum* serve as excellent sources of cold tolerance. The genetics of abiotic stresses is considered very complex and controlled by large number of small effect genes/QTLs.

Although conventional breeding methods have been used to address the issues of abiotic stresses, to breed more efficiently, scientists all over the world are now adopting integrated genomics-based tools in their breeding pipelines. A variety of genomics, physiology and breeding approaches have been extensively deployed for the genetic dissection of abiotic stress adaptation (Mir et al. 2012b). Several genomics approaches including QTL interval mapping, GWAS, transcriptomics, etc. have been deployed for the discovery of genes/QTLs/transcripts/markers for their use in genomics-assisted breeding for development of abiotic stress-tolerant crop varieties.

14.5 Crop Breeding in the Era of Genomics

14.5.1 Limitations of Conventional Breeding

While the conventional breeding is being practiced since centuries and has immensely helped in development of large number of varieties in different crops, it has its own limitations. Since the conventional breeding programs are dependent on phenotypic selection, which basically is more of an art than science, the effectiveness of such program is less. The major limitations which hamper the success of conventional breeding are (a) the time required for improvement of a trait is very long; (b) the accuracy with which the desirable genotypes could be selected is very less (this happens mainly due to linkage drag when many traits are transferred along with the trait(s) of interest—including those traits that have undesirable effects); and (c) the efficiency with which the selection for the trait can be done is also very less. These three limitations not only increase the time and cost required for development of a variety, but many times, the new genotype becomes susceptible to new pest or disease by the time it is ready for release as a new variety.

14.5.2 Advantages of Molecular Breeding/Genomics-Assisted Breeding

The three key limitations associated with the conventional breeding discussed above can effectively be overcome by the use of molecular breeding/genomics-assisted breeding (GAB). This can also be called as precision breeding (Kulwal et al. 2012). The three key advantages of molecular breeding over the conventional breeding are the following.

14.5.2.1 Time Saving

Using molecular breeding, the time and labour savings may arise from the substitution of difficult or time-consuming field trials (that need to be conducted at particular times of year or at specific locations or are technically complicated) with DNA marker tests. This is an important advantage of the molecular breeding, since lesser time will also curtail the cost of the breeding program and selections can be done any time during the year.

14.5.2.2 Increased Efficiency

Molecular breeding can greatly increase the efficiency and effectiveness of breeding compared to conventional breeding because it is simpler compared to phenotypic screening, selection may be carried out at seedling stage and single plants may be selected with high reliability. These key advantages may translate into greater efficiency or accelerated line development in the plant breeding programs.

14.5.2.3 Increased Accuracy

It is well-known that effect/influence of environmental factors on field trials is more, thereby limiting the accuracy of selection. However, selection based on DNA markers is more reliable as effect of environment is not there. Moreover, the total number of lines that need to be tested may be reduced, and specific genotypes can easily be identified and selected.

With the advances in the area of genomics coupled with the availability of large number of genomic resources in different crops (see next section), molecular breeding or GAB has now become an integral part of the crop improvement program (Varshney et al. 2021).

14.5.3 Molecular Breeding Tools for Efficient Breeding

As already discussed in the above section that there are several advantages of molecular breeding over conventional plant breeding, the traditional/conventional crop improvement techniques are now being replaced by GAB approaches. Recent advances in genomics tools and techniques have facilitated the development of large number of different types of molecular markers, genetic and physical maps, high-throughput and precise genotyping platforms and rapid discovery of genes using different approaches in almost all important crop plants (Bohra et al. 2020). Therefore, efficiency and precision of crop improvement could increase using these genomic tools and techniques. This has become primarily possible due to the advances in the next-generation DNA sequencing technologies and development of large numbers of molecular markers. The efficient breeding program today thus relies on several tools which are described in the following sections.

14.5.3.1 Molecular Markers

The recent years have witnessed the development of large-scale genomics and genetic resources including variety of molecular markers, expressed sequence tags (ESTs) or transcript reads, bacterial artificial chromosome (BAC) libraries, genetic and physical maps and genetic stocks with rich genetic diversity, such as core reference sets and introgression lines in majority of the crop plants (Varshney et al. 2010). The DNA-based markers also known as molecular markers are being used for the detection of genome sequence level differences between two or more than two individuals. The discovery of molecular markers has revolutionized crop improvement programs by providing quick and sophisticated/reliable crop improvement tools and techniques. The marker technology although discovered in the 1980s with the discovery of RFLPs has witnessed continuous evolution from hybrization-based markers to GBS-based markers (Mir et al. 2013; Mir and Varshney 2013; Gupta et al. 2013a, b; Kumar et al. 2021; Tyagi et al. 2019, 2021). A number of classifications have been proposed to classify molecular markers. Some of the classifications of markers include hybridization-based vs. non-hybridization-based, first-generation vs. second-generation vs. third-generation markers, past vs. present vs. future, low-throughput vs. medium-throughput vs. high-throughput markers, sequence based vs. non-sequence based and array-based vs. non-array-based markers (Gupta et al. 2008; Mir et al. 2013; Mir and Varshney 2013).

The advances in genomics has not only resulted in development of large numbers of markers in important crops but also in in the once considered orphan and resource poor crops like chickpea, pigeon pea and groundnut. Now in these crops, thousands of all important types of molecular markers including SSR, diversity arrays technology (DArT), single nucleotide polymorphism (SNP), different SNP platforms, micro-array-based markers, GBS, InDel markers, etc. are available. For instance, over the years, >2000 SSRs in chickpea, >3000 in pigeon pea and >2500 in groundnut have become available using different approaches of marker development. Similarly, thousands of DArT and SNP markers are available in these crops.

Several genotypic platforms including Kompetitive Allele Specific PCR (KASP) assays, GoldenGate assays, Vera Code assays and 60 K SNP chips using Affymetrix SNP platform and Axiom SNP array with thousands of SNPs uniformly distributed across the genome are available now (Thudi et al. 2021). These marker resources have been used in the study of genetic diversity, population structure, development of genetic maps and QTL mapping/GWAS for key traits in all major food crops. The genes/QTLs once identified are being deployed into molecular breeding programs aimed at enhancing targeted traits in different crop plants through MAS, marker-assisted recurrent selection (MARS) and genomic selection (GS). It is expected that the improved versions of next-generation crop varieties could be developed with enhanced quality traits, better yield and disease resistance.

14.5.3.2 Genetic/Linkage Maps

The molecular genetic map or linkage map refers to linear arrangement of genetic markers (loci) on the genome obtained on the basis of estimates of recombination fractions among the markers. The genetic map may be thought of as a “road map” of the chromosomes developed for a mapping population derived from two different parents. It indicates the position and relative genetic distances between markers along chromosomes, which is analogous to signs or landmarks along a highway. The concept of genetic linkage is known since the studies of Morgan 1911 and Sturtevant published genetic map of chromosome X of Drosophila in 1913. The first partial genetic map of maize was published by Emerson and colleagues in 1935. These linkage maps were prepared by analysing segregating populations derived from crosses of genetically diverse parents and estimating the recombination frequency (RF) among genetic loci. The distance between the markers on a genetic map is related to the RF between the markers. Lower the frequency of recombination between two markers, the closer they are situated on a chromosome, and hence greater the frequency of recombination, more is the genetic distance. Markers that have a recombination frequency of 50% are described as “unlinked” and assumed to be located far apart on the same chromosome or on different chromosomes. Different chromosomal regions vary in their recombination frequency. Because of this, genetic maps cannot be used to measure physical distance between markers on the genome and only provide an approximation of physical distance, as well as a representation of marker order along the chromosome. Genetic maps have in general several functions including (1) providing an insight into genome organization, (2) the evolution of species, (3) synteny between related species, (4) rearrangement across taxa and more importantly (5) identification of genes/QTL for a trait of interest.

The molecular genetic maps based on DNA markers are now available in almost all plants of significant academic and economic interest, and the list of plants is growing regularly. Different types of molecular markers have been used for development of genetic maps, and sometimes genetic maps of only individual chromosomes have been constructed on the basis of needs of researchers. However, development of whole genome maps covering all the chromosomes in genome is always desirable. Initially RFLP-based genetic maps were developed, but with the discovery of markers of choice like SSRs and SNPs, now almost all maps are based on these markers (Thudi et al. 2021). More recently, several high-throughput genotyping platforms have been developed based on SNP markers, and the use of these genotyping platforms has facilitated the development of high-resolution and high-density genetic maps. These high-density genetic maps prove useful in fine mapping of genes. The different maps developed over the years have been used for mapping, tagging, cloning and characterization of large number of genes/QTLs in all important crop plants, and the information so generated can facilitate efficient breeding.

14.5.3.3 Use of Genome Sequence Information for Crop Breeding

The science of plant genomics research had its beginning with the publication of genome sequence of *Arabidopsis thaliana* in the year 2000. Nowadays, the genome sequence of almost all the major crop plants including cereals and legumes have become available (Thudi et al. 2021). After sequencing of Arabidopsis genome, rice genome was sequenced in year 2005. After sequencing of rice genomes, draft genomes of several cultivars among ssp. *japonica* and ssp. *indica* have also become available. Draft genomes of Australian wild A genome taxa including *O. rufipogon* and *O. meridionalis*, other wild species, core collections and mini-core core collections have also become available in recent years (for a review, see Thudi et al. 2021).

The other most important cereal crop "maize" genome was sequenced using a most widely used female parent "B73" for developing maize hybrids and study of maize genetics (Schnable et al. 2009). Followed by the sequencing of B73 genome, draft genome sequence of other maize inbred including Mo17, W22, HZS, SK and K0326Y were also generated. Bread wheat, one of the most important cereal crops with complex and huge genome size, has also been sequenced by the International Wheat Genome Sequencing Consortium (IWGSC) more than a decade after the initial drafts of the rice genome (International Wheat Genome Sequencing Consortium, IWGSC, http://www.wheatgenome.org/). Similarly, barley genome was also sequenced using six-row malting cultivar Morex (Mascher et al. 2017) followed by sequencing of Tibetan hulless barley (Zeng et al. 2015, 2018) and wild barley species AWCS276.

Like cereals, major legume genomes have also been sequenced. For instance, draft genome of cultivated soybean Williams 82 (Schmutz et al. 2010) and undomesticated ancestor of *G. max*, the *G. soja* (Kim et al. 2010), have become available. The genome sequence of two diploid progenitor species of groundnut (*A. duranensis* V14167 and *A. ipaensis* K30076) was reported by the International Peanut Genome Initiative (IPGI) through the Peanut Genome Consortium (PGC) (Bertioli et al. 2016). This was followed by sequencing of several other groundnut genomes (Chen et al. 2016; Lu et al. 2018; Yin et al. 2018; Bertioli et al. 2019; Zhuang et al. 2019). The common bean genome sequence has also become available for Andean inbred landrace "G19833" (Schmutz et al. 2014). For grain legume crop chickpea, a draft genome sequence for Kabuli genotype, CDC Frontier was generated by Varshney et al. (2013b) and for desi chickpea genotype ICC 4958 by Jain et al. (2013). For pigeon pea, draft genome assembly was developed for variety Asha (ICPL 87119) (Singh et al. 2012; Varshney et al. 2012).

The sequencing of these plant genomes has played very crucial role in discovering important genes and understanding their biological functions. Due to advances in genome sequencing technologies, the speed of sequencing has increased, and cost of sequencing has drastically decreased. This has resulted in the sequencing of draft genome of >800 plant species, and the number is continuously increasing (Thudi et al. 2021). The sequencing of crop genomes and the information derived have been utilized in both basic and applied research. For instance, this information has been utilized in working out evolutionary relationships, developing better phylogenetic classification and discovery of genes, alleles, markers, etc. The sequencing of large number of genomes has also resulted in the introduction of the concept of plant pan genome, each composed of "core genome" and "dispensable genome". The recent advances in genomics tools and techniques have helped in the development of genomics resources in all major crop plants in the world. Several databases like Legume Information System (LIS https://legumeinfo.org/; LegumeIP, https://plan.tgrn.noble.org/ LegumeIP; and Know Pulse, https://knowpulse.usask.ca) have been developed for providing genomic information.

14.6 Approaches for Efficient Breeding

14.6.1 Genomics-Assisted Breeding

With the advances in the area of genomics in last few years coupled with the availability of large numbers of genomic resources in terms of mapping populations, different types of molecular markers, genome sequence, linkage maps and identified QTLs/MTAs have changed the way plant breeding is being practiced. The important advantage the science of genomics has brought is the increased efficiency of the breeding programs leading to the proper understanding of the genetic architecture of the traits and incorporating this information in the varietal development programs. There are different ways through which this can be accomplished and are discussed below.

14.6.1.1 Marker-Assisted Selection (MAS)

The simple meaning of MAS is the use of molecular marker linked with the QTL/gene of interest as the substitute for making selection for a desirable genotype under laboratory conditions rather than cumbersome field-based phenotypic screening. However, before using the markers in the breeding program, there are few key things which need to be satisfied. These are (1) identification of the marker (s) associated with the trait of interest, (2) validation or testing suitability of these markers in the desirable genetic background and (3) use of these markers in the breeding program though marker-assisted backcrossing (MABC) for transfer of the required QTL/gene in the desirable genetic background (Kulwal et al. 2012). The general pre-requites for undertaking MAS have also been discussed elsewhere (Jiang 2015). The success of MAS entirely depends on the accuracy with which MTAs are identified. This can be accomplished by the development of linkage maps (preferably high-density maps) using the genotypic data of large numbers of markers generated for a biparental mapping population developed by crossing two genotypes differing for the trait of interest. The biparental mapping populations which can be used for this purpose are F2, doubled haploids (DH) and the recombinant inbred lines (RILs). The advantage of having linkage maps is that, one can place the markers on different chromosomes in linear order and assign distances between these markers. When development of biparental mapping population is not possible/ feasible (particularly in case of tree species), one can use the germplasm or natural population and genotype them using molecular markers for identification of the QTLs/MTAs following the approach of association mapping (AM) or GWAS.

It is equally important that the population for which genotypic data has been generated is phenotyped precisely for the trait of interest so that the data can be used in conjunction with the genotypic data for the identification of QTLs/MTAs. Ideally, phenotypic data recorded over seasons and locations is desirable so that QTL × environment interactions can also be worked out. Several approaches of QTL analysis have been proposed for the analysis of the data (reviewed by Kulwal 2018). Large numbers of studies have been carried out following the approach of biparental QTL mapping and GWAS in different crops for variety of traits (Gupta et al. 2011, 2014, 2019). However, not all the markers linked with these identified QTLs can be used in the MAS program. The underlying criterion for this therefore is that the QTL which is to be transferred through MAS must be a major effect QTL. However, in recent years with the availability of wealth of data generated through large numbers of QTL and GWA studies as well as meta-analysis, it has become easy to identify major QTL and choose the type of marker for the MAS program. The advantage which MAS offers is that it can effectively be utilized for traits having low heritability. Large numbers of studies are now available where MAS has been utilized for the transfer of useful QTL in the desirable genetic backgrounds in different crops leading to the development of superior breeding lines and varieties (Varshney et al. 2013a). Given the fact that there has been tremendous reduction in the costs of marker genotyping and the advantages which MAS offers, it is anticipated that MAS will be used on large scale by the breeders for crop improvement in the future.

14.6.2 Marker-Assisted Recurrent Selection (MARS)

It is well established that majority of the traits of interest are polygenic in nature and are controlled by many genes/QTLs, each having minor effect. The problem with deployment of these minor QTLs in the breeding program through MABC is that they are not expressed consistently over different seasons. It therefore become difficult to introgress multiple QTLs in a common background. Although in conventional breeding recurrent selection has been suggested as an effective strategy for improving the polygenic traits by accumulating the favourable alleles in the population, the strategy is not very effective due to effect of environment on the phenotype and the long time required for genotypic selection (typically 2–3 crop seasons per cycle) (Godiki et al. 2016). In order to address this issue and to utilize these minor effect QTLs in the breeding program, MARS has been proposed and allows genotypic selection and intercrossing in the same crop season (Bernardo and Charcosset 2006).

MARS utilizes markers initially for the identification and then selection of several genomic regions associated with the complex trait(s). This then can be used to assemble the best-performing genotype within a single or across related populations (Ribaut et al. 2010; Jiang 2015). The advantage of MARS over MABC is that the genetic gain achieved through the former method are more as compared to the later since MARS deals with transfer of several QTLs as against that of only selected QTL in the MABC. There are several successful examples in crop plants where MARS has shown increased efficiency of selection in the breeding programs. Some of these examples include in maize for yield improvement (Johnson 2004), for improving grain yield under drought stress (Beyene et al. 2016), in wheat for bread making-related traits (Charmet et al. 2001) and crown rot resistance (Rahman et al. 2020). There are many other examples where MARS has been successfully utilized in different crops.

14.6.3 Genomic Selection (GS)

Although conventional breeding coupled with modern breeding techniques have helped in increasing the genetic gain to a considerable extent, it is necessary that the rate of gain should be increased further to address the challenges of food security (Xu et al. 2020). Genomic selection (GS) also referred to as genome-wide selection (GWS) or genomic prediction is one of the forms of MAS, in which large numbers of markers covering the entire genome are used to predict the genetic value of a trait or individual. Selection for the desirable individuals is based on computing the genomic estimated breeding values (GEBVs) using the markers across the genome (Meuwissen et al. 2001; Crossa et al. 2017). This in contrast to MAS in which only the markers tightly linked with the trait are used in the breeding program

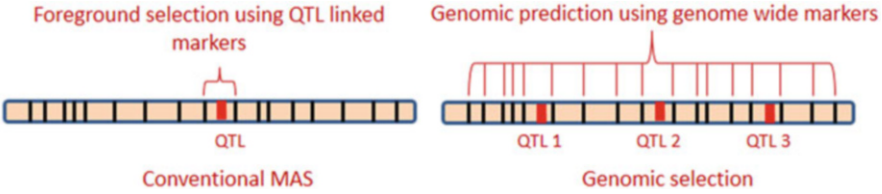

Fig. 14.1 Figure depicting the difference between marker-assisted selection and genomic selection; the horizontal bar represent the chromosome on which mapped markers are shown in black vertical lines, while position of QTL is shown with red rectangle

(Fig. 14.1). The underlying criteria in GS is that rather than focusing on only important or major QTLs in the breeding programs, it utilizes all the QTLs (minor and major) by using the genome-wide markers while making the prediction. In order to increase the accuracy of GEBV and GS, large number of markers across the genome is thus essential so that all the QTLs are in LD with at least one marker (Meuwissen 2007; Jiang 2015).

For undertaking the GS, a training population (TP) for which genotypic and phenotypic data is generated is essential so that a prediction model for understanding the relationship between the two can be developed. The genotypic data of the breeding population is then fed into this model to calculate GEBVs for these lines (Heffner et al. 2010). There are different ways through which these GEBVs can be calculated. The GEBVs calculated represents the sum of the effects of all QTLs across the genome. GS thus outperforms MAS in terms of its effectiveness (Kulwal et al. 2012). Since there are different statistical models to estimate the breeding values, each having its own superiority and limitations under the given scenario, one model cannot fit all the situations (Rahim et al. 2020). It is therefore difficult to suggest which model will work under the given scenario.

Although the technique of GS was initially proposed for use in animal breeding (Meuwissen et al. 2001), in recent years, emphasis is given on the use of GS in the breeding of many crops. It has also been reported that the cost per unit gain was lower up to 55% using GS than the phenotypic selection (PS) in case of oil palm and that GS was superior to MARS and PS (Wong and Bernardo 2008). Similarly, superiority of GS over PS has also been reported in many crops (reviewed by Jiang 2015; Pandey et al. 2020; Rahim et al. 2020). Although the technique of GS appears to be promising, its success primarily depends on the size of the TP which is used for identifying the associations. This TP should be updated frequently by incorporation of new genotypes in the analysis so as to maintain the prediction accuracy (Rahim et al. 2020). However, given the fact that not many breeders are trained in the use of molecular markers on large scale and computer programs dealing with estimation of breeding values, it is necessary that breeder friendly software packages should be developed for such purpose. In addition, developing the efficient models which can involve genotype × environment interaction and achieve greater prediction accuracy are therefore required (Xu et al. 2020). It is expected that with the reduction in the cost of marker genotyping due to advances in the next-generation sequencing techniques, accompanied with advances in the computational analysis, GS will become an integral part of the crop breeding programs (Desta and Ortiz 2014).

14.6.4 Gene Editing in Plant Breeding

Crop improvement is a continuous process for ensuring food and nutritional security for burgeoning world population. Great success has been achieved through conventional plant breeding and through use of transgenics. The advantages, success stories and limitations of transgenic technology have been reviewed in detail in several earlier publications (see Datta et al. 2004; Husaini et al. 2011; Ahmar et al. 2020). Due to several concerns associated with transgenic technology, new plant breeding techniques including RNA interference (RNAi), gene silencing and gene editing are gaining worldwide attention in different crop improvement programs. Among the different new breeding technologies, gene/genome editing (GE) making use of site-directed nucleases (SDNs) is considered the most important and one of the promising technologies that can overcome the inherent limitations associated with classical/conventional plant breeding and transgenic technology. The GE tools and techniques are considered effective for modifying the target genome and creating desired and novel new traits/phenotypes in crop plants. The breakthrough technology started with sequence-specific nucleases including zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and now the most important and emerging clustered regularly interspaced short palindromic repeats, CRISPR/Cas technology (Lloyd et al. 2005; Cermak et al. 2011; Mahfouz et al. 2011; Li et al. 2013; Nekrasov et al. 2013; Shan et al. 2013; Tan et al. 2020).

Discovered by Emmanuelle Charpentier and Jennifer A. Doudna (Nobel Laureates in chemistry for the year 2020), CRISPR/Cas9 genetic scissor is considered one of gene technology's sharpest tools that can help to change the DNA of crop plants with extremely high precision. The noble prize winning technology "CRISPR/Cas9" can help to change the code of life of crop plants over the course of a few weeks. CRISPR technology has been used extensively in plant genome editing over the past few years and has a great potential for precision breeding (Zhang et al. 2020). However, this system is also subject to some technical limitations and some other main obstacles including consumer preference of gene-edited food product that may hinder its application in food crops (Ahmar et al. 2020). The progress and perspectives of the use of genomic editing tools and technologies for their use in plant breeding have been extensively reviewed (see Varshney et al. 2019; Zimny et al. 2019; Zhang et al. 2019, 2020).

The technology has been mostly used for improvement of crop yields, quality and stress resistance by simply knocking out one or more than one genes that control a particular trait (Zhang et al. 2018). For instance, knocking out of genes (*Gn1a*, *DEP1* and *GS3)* in cereal crop rice led to the increase in grain number, dense erect panicles and increased grain size (Zhang et al. 2020). Similarly, in maize disruption of waxy gene "*Wx1*" resulted in increase in concentration of amylopectin with enhanced digestibility of grain. In wheat and tomato, knocking out of gene/allele "*MLO*" resulted in the development of powdery mildew-resistant wheat and tomato plants (for a review, see Zhang et al. 2020). The technology has recently revolutionized agriculture by helping in fixing heterosis in rice hybrids (Khanday et al. 2019; Wang et al. 2019), which is otherwise lost in subsequent generations due to segregation of alleles in F_1 hybrids.

In these studies involving fixing heterosis in rice, a genotype known as *MiMe* (Mitosis instead of Meiosis) was produced by targeting important genes responsible for meiosis, and thereby haploid plants were developed using CRISPR technology. When *MiMe* genotype was combined with haploidy in hybrid rice, the clonal progeny maintained genome-wide parental heterozygosity that demonstrates the possibility of asexual reproduction through seed propagation in crops (Khanday et al. 2019; Wang et al. 2019). The technology of fixing heterosis in rice hybrids will allow the maintenance of rice hybrids during propagation to subsequent generations. This revolutionary process can reduce the cost of hybrid seed production as well will allow farmers to produce their own hybrid seeds, reducing dependence of subsistence farmers on commercial seed producers. In a study in wheat, *A3A-PBE*-mediated cytidine base editing was used for editing of all the six *TaALS* alleles that resulted in the development of nicosulfuron-resistant wheat lines (Zong et al. 2018). In an another study considered one of the best examples of the use of CRISPR/Cas technology for crop improvement is CRISPR-mediated gene regulation in tomato, where CRISPR technology was used to mutate the promoters of genes responsible for controlling the most important quantitative traits including fruit size, inflorescence branching and plant architecture (Rodriguez-Leal et al. 2017). In summary, the technique has already been used in ~20 crop species improvement programs for the development of crops that possess high yield and withstand biotic and abiotic stresses. The list will keep on increasing in more crops involving improvement of important traits in the near future.

14.6.5 Role of Bioinformatics in Breeding

Bioinformatics is multi-disciplinary science, integrating biology, statistics, mathematics, computer science, etc. that helps in solving the biological problems. The need for the science of bioinformatics was strongly felt with the evolution of next-generation sequencing technologies that helped in sequencing of genomes of almost all important crop species. The sequencing of genomes through genome sequencing projects led to the explosion of sequencing data produced and hence demands conception/creation of novel discipline called “bioinformatics”. The aim of bioinformatics is primarily storage, acquisition, analysis, distribution and modelling of various types of sequence data (Aslam et al. 2004). Therefore, computational biology and bioinformatics have their roots in life science and helps to find out the function of macromolecules, biological sequence data and genome content/genes.

In the last two to three decades, the science of bioinformatics has emerged as a significant tool for the use of large volumes of data that have been generated using different omic-technologies. The analysis of data through different software programs provided vital role in extracting useful information, interpretation of data and future decision-making process (Batley and Edwards 2008, 2009; Aslam et al. 2004). The analysis of genome sequencing data and other genotyping data through bioinformatics tools have helped in discovery of thousands of genes in important crop species. The identified genes will prove useful in plant breeding programs aimed at enhancing trait performance of varieties leading to the development of next-generation crop varieties.

The huge volume of genome sequencing data generated required development of databases for its storage, and therefore in 1986 largest sequence databases were developed in association of GenBank with European molecular biology laboratory

(EMBL). A number of bioinformatics online databases including crop-specific databases are now available including BGI Rice Information System, Gateway of Brassica Genome, ChloroplastDB, EMBL, GRAINGENES, GRAMENE, GRIN, NCBI, LIS, KOME database, KEGG PLANT, OryGenesDb, TAIR, TREP, etc. The complete list has been tabulated elsewhere (Aslam et al. 2004). The details of classification of databases, databases for transcription factor in plants, small RNA databases, genomic databases, crop-specific databases and list of different bioinformatics tools for analysis of NGS data, dbEST available in different crops, etc. are also available in different review articles (see Vassilev et al. 2005; Agarwal and Narayan 2015; Aslam et al. 2004; Kushwaha et al. 2017). In general, there are three primary sequence databases including GenBank (NCBI), the Nucleotide Sequence Database (EMBL) and the DNA Databank of Japan (DDBJ). These databases are actually repositories used for storing of raw sequence data. However, each data entry is also extensively annotated, and important properties and features of each sequence are also highlighted. These three important databases exchange data routinely. Similarly, databases are also available for storing protein sequence like SWISS-PROT and TrEMBL. The constant serge in the omics data and the emergence of molecular breeding technologies coupled with advances in genomics and computational biology provide ample opportunities for bioinformatics to develop efficient approaches for plant breeding.

14.6.6 Integrated System of Data Management and Delivery

With the challenges posed by the varying environments, the activities of plant breeding have expanded in the last few decades. This has resulted in multi-disciplinary and multi-institutional collaborations and generation of large-scale phenotypic data, typically collected over different environments having increased dimensionality due to use of sensors and techniques like phenomics. This is in addition to the large-scale genomics information which is being generated routinely due to the availability of high-throughput techniques resulting in big data. This big wealth of data on one hand has increased the capabilities of the breeders in achieving their goals and on the other hand has made it necessary to think about proper management of the information. While the scientific community spends most part of their time on generation of valuable data, only limited time is spent on proper documentation, analysis and interpretation process.

Although there is no doubt that the key to success for any breeding program is careful collection of the data, it is equally important that there should be proper integration of the other parameters in the analysis process leading to the meaningful interpretation of the results. Any wrong decision or improper interpretation of the data can cause huge loss to any breeding program. It is therefore very important that there should be an integrated system of data management and delivery in a plant breeding program so that one can access, analyse and recombine the vast wealth of data (Kuriakose et al. 2020). The advantage with this system is that, the information generated in the experiment is not only stored carefully, but can be retrieved at any given time as per the need. This generally is not possible in the traditional way of collecting the data. Therefore the success of any future breeding program will depend on how strong the data management and delivery system is in place. An excellent overview of this aspect is discussed by Kuriakose et al. (2020).

This type of setup is generally seen in many private seed companies, but lacked in the public sector breeding programs in developing countries. It is therefore envisioned that the primary challenge for the plant breeders in coming years will be to design the efficient system to handle and analyse the massive amounts of multifarious data that is generated in the breeding programs rather than access to the modern technology (Cobb et al. 2019).

14.6.7 Speed Breeding

Ideally, any conventional breeding program involve crossing and/or selection for successive generations (typically 4–6) followed by yield evaluation trial before a variety is released for cultivation. Majority of the varieties so far in the world have been released using the same approach during the last several decades. Generally 1–2 generations/cycles are possible per year in majority of the crop plants. With this slow speed, the time required in a breeding process is too long and also slows down the process of variety development. Although the alternate techniques like shuttle breeding and doubled haploid can be used to shorten the time required in a breeding program, they have their own limitations. Therefore, in order to accelerate the process of breeding and generation advancement, a technique called speed breeding was proposed recently (Hickey et al. 2019). As the name suggests, speed breeding relies on use of environment-controlled growth chambers equipped with artificial lights which can accelerate the plant growth and development so that multiple generations of crop plants can be advanced per year (Ahmar et al. 2020). This is very much required in today’s context because in order to produce more to feed the growing world population, there is an urgent need to accelerate the rate of genetic gain. While the molecular breeding techniques are efficient in introgression of the desired gene/QTL, their utility will be enhanced only if more number of generations are advanced per year. Speed breeding enhances the growth of the plant by regulating light and temperature, thereby promoting early flowering and rapid generation advancement (Bhatta et al. 2019).

Since each crop plant has differential requirement of photoperiod for normal growth and development, the protocol established for one crop plant may not be suitable for the other. The proper understanding of the physiology of the plant is thus necessary. Ideally, in speed breeding, vegetative growth is enhanced by increasing the temperature in the chamber, while it is lowered during the reproductive growth (Hickey et al. 2019). However any such improvement comes with a cost. For instance, flowering can be hastened in the environment-controlled chambers; however, the total biomass and yield will be impacted due to this (Bhatta et al. 2019). However, the success achieved in achieving four to six generations per year in crops like wheat, barley and canola (Hickey et al. 2017; Ghosh et al. 2018; Watson et al.

2018) using speed breeding shows the great promise this technique offers in the crop improvement programs and in accelerating the speed of variety development. However, one need to weigh the cost involved in developing the facility and the associated gains through it before investment is made on this technique.

14.7 Advances in High-Throughput Phenotyping

One of the major factors limiting progress in GAB is lack of precise phenotypic data. Therefore, plant phenomics is considered one of the most important factors for translating the progress made in the area of plant genomics. The area of plant phenotyping has made huge progress in the last decade by replacing invasive or destructive methods of phenotyping by the high-throughput precise non-destructive methods of phenotyping (Mir et al. 2019). The advancement made has revolutionized crop phenomics and allowed screening of large germplasm (mapping populations, core collections and breeding material) with high precision/accuracy with less efforts, time and labour. These advances not only have generated huge amount of information but have also necessitated use of novel techniques for the analysis of the big data. These issues have been discussed in the following sections.

14.7.1 Plant Phenotyping Platforms

The non-destructive high-throughput phenotyping (HTP) platforms developed and used routinely include infrared cameras, fluorescent microscopy/spectroscopy, three-dimensional camera, lidars (light detection and ranging), magnetic resonance imaging (MRI) and positron emission tomography (PET), canopy spectral reflectance (SR) and infrared thermography (IRT), nuclear magnetic resonance (NMR) and digital imaging (see Mir et al. 2019 for review). The use of these HTP platforms helps in recording trait data on thousands of plants in a single day similar to next-generation sequencing technology in the field of plant genomics (Finkel 2009). A number of state-of-the-art international phenomics centres/facilities have been developed for precisely recording high-throughput phenotyping data in cost-effective manner. Some of the important phenomics facilities include the Plant Accelerator in Adelaide, Australia (http://www.plantaccelerator.org.au/); High Resolution Plant Phenomics Centre (http://www.plantphenomics.org/HRPPC) in South Australia; the Jülich Plant Phenotyping Centre (http://www.fz-juelich.de/ibg/ibg-2/EN/methods_jppc/methods_node.html) in Jülich, Germany; Leibniz Institute of Plant Genetics and Crop Plant Research in Gatersleben, Germany; and the National Plant Phenomics Centre (http://www.phenomics.org.uk/temp-site/about.html) in the UK to name a few (Gupta et al. 2012; Mir et al. 2015, 2019). The relevant information about plant phenotyping is being provided by the world's major plant phenotyping centres "International Plant Phenotyping Network (IPPN)" (https://www.plant-phenotyping.org/). In addition, different private companies like LemnaTec, Phenokey, PhenoSpex, Photon System Instruments, Wiwam and We

Provide Solutions are offering large-scale, custom high-throughput phenotyping platforms for the field and laboratory (Mir et al. 2019).

The different HTP phenotyping platforms that have been developed for recording data on variety of traits in almost all crop species include “LEAF-E” developed and used for the analyses of leaf growth parameters, “Zeppelin NT aircraft” used aerial phenotyping “Phenovator” and “GROWSCREEN FLUORO” used for phenotyping for photosynthesis and growth “TRiP (Tracking Rhythms in Plants)” used for determination of circadian period. Similarly, image-based phenotyping methods have been developed and used for measuring plant stresses including cold tolerance and spikelet anthesis. The other phenomics platforms like” PHENOPSIS” was used to dissect plant responses to soil water deficit, and “Unmanned Aerial Platforms (UAP)” was used for measuring low-nitrogen (low-N) stress tolerance. The “Hyperspectral Imaging (HIS)” was used to determine spectral changes on the leaf and cellular level in plants during resistance reactions/host-pathogen interactions. Like phenotypic platforms, a number of different software programs have been also developed for recording trait data on variety of traits. A list of software programs and phenotyping platforms for high-throughput precise phenotyping being used in several laboratories across the world are available elsewhere (see Mir et al. 2019).

The different precise high-throughput phenomics methods/platforms already developed have been used for trait phenotyping of variety of traits including growth traits, phenological traits, physiological traits, scoring disease incidence, insect damage, drought tolerance and recording data on different plant organs like roots, seeds and shoots (for review see Mir et al. 2019). For instance, in crops like rice, wheat, barley, maize, pea, Arabidopsis, potato, soybean, etc., different phenotyping platforms have been used, and data has been recorded for spikelet anthesis, circadian period, plant height, leaf growth parameters including leaf area, area phenotyping of canopies, photosynthesis, photosynthesis efficiency, chlorophyll content, leaf nitrogen content and canopy height (see Mir et al. 2019 for more details). Phenomics has also been used for the study of plant responses to various abiotic stresses including drought, heat, cold tolerance, salinity and nutrient-starving. For drought tolerance, trait phenotyping either in glasshouse or in field have been conducted and approaches like osmotic balance in hydroponics to conveyer systems in glass house to rainout shelters in the field have been used very extensively.

Several important methods and platforms that are now routinely being used for precise high-throughput phenotyping of drought tolerance have been discussed in detail elsewhere (Mir et al. 2012b). These methods are based on imaging, robotics and computers that allow recording of trait data of thousands of plants in a day in non-destructive manner. Like abiotic stresses, phenomics platforms have been also used for recording trait data on biotic stresses like insect pests. For instance, automated video tracking “a phenomics platform” has been developed and used to record the aphid feeding behaviour on leaf discs that is helpful to measure plant resistance. This platform of automated video tracking can be also used to measure data on aphids and other piercing-sucking insects in plants in high-throughput manner. Like insects, the platforms can be used for recording data on disease reactions and for characterization/selection of resistant plants against fungal pathogens. In summary, the phenomics platforms/methods/software have been used to record data in high-throughput fashion for variety of traits in almost all important crops, and the trait evaluation has also led to the genetic dissection leading to discovery of genes/QTLs for several traits including root system architecture traits, seed shape, osmotic tolerance and biomass traits in crops like rice, wheat, barley and mustard.

14.7.2 Applications of Artificial Intelligence (AI) and Machine Learning (ML) in Crop Breeding

Classical plant breeding techniques mainly focus on estimation of genetic diversity, analysis of stability for different traits over different seasons and environments, hybrid prediction using different parental combinations and related things and rely on routine statistical methods for analysis of the data (Niazian and Niedbała 2020). Besides this, in order to identify the desirable plant, breeder often needs to take repeated observations in the field and make careful selection. This not only requires lot of time, but skill and experience of a breeder. It has now been realized that in any plant breeding program, rapid and precise phenotyping for the desired trait is very essential. Since this involves recoding thousands of data points in shorter time, in recent years a shift from traditional way of phenotyping to use of sensor based phenotyping has been seen. This has been facilitated by the advances in the area of phenomics and phenotyping platforms as discussed above. The important advantage with these techniques is that it can enable collection of enormous and high-dimensional data in a very short span of time. Similarly, with the advances in the omics techniques (genomics, proteomics, metabolomics, epigenomics), the volume of data which is being generated in any such experiment is huge. In order to handle this huge amount of data efficiently in a breeding program and to make meaningful interpretations, the methods which involve minimal human efforts are required for analysing the data with increased precision (Harfouche et al. 2019).

Modern technology has been of great help to the breeders in this endeavour. For instance, digital images of standing crop in the fields are taken from the surface or through air with the help of drones or unmanned aerial vehicles (UAVs). This not only saves time in recording the data but also reduces the error associated with the manual way of recording observations. The large numbers of images or data points so captured can be analysed using computer tools for understanding the traits and interpreting the results. Artificial intelligence (AI) and machine learning (ML) and variants thereof (neural networks, deep learning, etc.) are considered as the important breakthrough in dealing with this big-data. The ML tools can collectively analyse the phenotypic, genetic and environmental data to help breeders better understand the relationships between genetics, environment and plant performance. While doing this, ML method uses approximations to find out the patterns which are embedded in the data so that it can be used to predict the future data (Murphy 2012). Thus, AI and ML can be used practically in all aspects of breeding (including prediction of phenotype, image identification, disease identification and genomics experiment including GWAS and GS studies) (Harfouche et al. 2019). This not only will accelerate the process of breeding but will allow screening of large number of accessions in a breeding experiment with increased precision in shorter span of time. This can enable breeders in identifying the desirable plant suited for a particular climate and soil type and identifying desirable cross combination of genes for increased yield (Beans 2020). In recent years, different approaches of ML have been proposed and used in plant breeding programs (Parmley et al. 2019; Kuriakose et al. 2020; Niazian and Niedbała 2020).

Although sensors and machines can increase the volume of data, they cannot replace experience of a breeder. Moreover, since breeders generally are not experienced in the algorithms which are used in ML, therefore, any such experiment requires close cooperation between statisticians, IT specialists and experienced breeders. It is expected that with the growing awareness about AI and ML and quest for better algorithms involved in analysis and interpretation of the data, they will become an integral part of breeding programs in the future.

14.8 Emerging Challenges at National and International Level

Although plant breeding has contributed immensely and resulted in achieving self-sufficiency in many crops in several countries, with the pressure of ever-growing world population and the impact of climate change resulting in uncertain environments, the task of plant breeders has become more challenging. In the future, the main task before the breeders will be to develop varieties with higher productivity having better adaptability to the changing climates. The important challenge which breeders need to address is to develop the varieties which are suited to the specific agroecological regions rather than developing mega-varieties. In addition, emphasis need to be given on varieties which offer better nutrient use efficiency and requiring less resources so that they can offer economic benefits to the farmers. Similarly, another challenge before the breeders will be to develop varieties which can offer food security for the increasing world population and sustainable agriculture.

14.9 Future Thrust Areas and Conclusions

While plant breeding is often considered as an art and science of genetically improving plants and no technique can substitute plant breeding, it needs to adapt to the advances in digital revolution so as to integrate it with molecular techniques for the benefit of humankind. For this purpose, plant breeders need to sync themselves with the advances in the area of genomics and phenomics and embrace modern techniques like AI and ML to increase its effectiveness and to address the challenges of the future. Although the impact of plant breeding in increasing the food productivity is known to everyone, it is still necessary that the subject should be considered as priority by all. Integration of promising techniques like MAS and GS

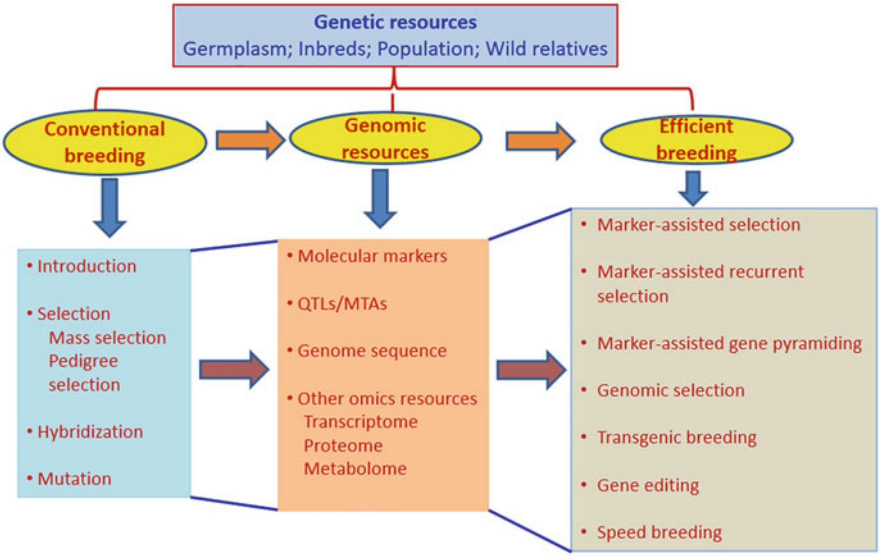

Fig. 14.2 Figure depicting different breeding schemes from conventional to efficient using genomic resources

in the breeding program can offer increased genetic gains. This probably seems to be the only way through which we can ensure/achieve food security in more sustainable way. In addition, emphasis on exploiting the potential of the PGRs is very essential. Moreover, the novel methods like speed breeding and genome editing technique like CRISPR/Cas can allow rapid generation advancements of a breeding cycle and development of genetic diversity for breeding purpose, respectively, and are very promising if used in proper manner. To conclude, rather than using technologies in isolation, integration of modern and efficient techniques in the breeding program can help in achieving food security in a sustainable manner in long run (Fig. 14.2).

Acknowledgements PLK would like to thank Department of Biotechnology, Govt. of India for the research grants during the course of writing this article. RKV thanks Bill and Melinda Gates Foundation, USA, for supporting several projects related to genomics-assisted breeding at ICRISAT and acknowledges Science and Engineering Research Board (SERB), Department of Science and Technology, Government of India, for awarding JC Bose National Fellowship.

References

Agarwal R, Narayan J (2015) Unraveling the impact of bioinformatics and omics in agriculture. Int J Plant Biol Res 3:1039

Ahmar S, Gill RA, Jung KH et al (2020) Conventional and molecular techniques from simple breeding to speed breeding in crop plants: recent advances and future outlook. Int J Mol Sci 21: 2590

Aslam Z, Khattak JZK, Ahmed M et al (2004) A role of bioinformatics in agriculture. In: Ahmed M, Stockle CO (eds) Quantification of climate variability, adaptation and mitigation for agricultural sustainability, pp 414–430

Bailey-Serres J, Fukao T, Ronald P et al (2010) Submergence tolerant rice: *SUB1*'s journey from landrace to modern cultivar. Rice 3:138–147

Bainsla NK, Meena HP (2016) Breeding for resistance to biotic stresses in plants. In: Yadav P, Kumar S, Jain V (eds) Recent advances in plant stress physiology. Daya Publishing House, New Delhi, pp 379–411

Balconi C, Stevanato P, Motto M, Biancardi E (2012) Breeding for biotic stress resistance/tolerance in plants. In: Ashraf M et al (eds) Crop production for agricultural improvement. Springer, pp 57–114

Batley J, Edwards D (2008) Bioinformatics: fundamentals and applications in plant genetics and breeding. In: Kole C, Abbott AG (eds) Principles and practices of plant molecular mapping and breeding. Science Publishers, Enfield, pp 269–302

Batley J, Edwards D (2009) Mining for single nucleotide polymorphism (SNP) and simple sequence repeat (SSR) molecular genetic markers. In: Posada D (ed) Bioinformatics for DNA sequence analysis. Humana Press, New York, pp 303–322

Beans C (2020) Inner workings: crop researchers harness artificial intelligence to breed crops for the changing climate. Proc Natl Acad Sci U S A 117:27066–27069

Bernardo R, Charcosset A (2006) Usefulness of gene information in marker-assisted recurrent selection: a simulation appraisal. Crop Sci 46:614–621

Bertioli DJ, Cannon SB, Froenicke L et al (2016) The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. Nat Genet 48:438

Bertioli DJ, Jenkins J, Clevenger J et al (2019) The genome sequence of peanut (*Arachis hypogaea*), a segmental allotetraploid. Nat Genet 51:877–884

Beyene Y, Semagn K, Crossa J et al (2016) Improving maize grain yield under drought stress and non-stress environments in sub-Saharan Africa using marker-assisted recurrent selection. Crop Sci 56:344–353

Bharadwaj C, Tripathi S, Soren KR et al (2021) Introgression of "QTL-hotspot" region enhances drought tolerance and grain yield in three elite chickpea cultivars. Plant Genome 14:e20076

Bhatta M, Sandro P, Smith MR et al (2019) Need for speed: manipulating plant growth to accelerate breeding cycles. Curr Opin Plant Biol 60:101986. https://doi.org/10.1016/j.pbi.2020.101986

Bidinger FR (1992) Utilization of plant genetic resources in crop breeding programmes. Dinteria 23:70–81

Bohra A, Chand Jha U, Godwin ID et al (2020) Genomic interventions for sustainable agriculture. Plant Biotechnol J 18:2388–2405

Bošković J, Zečević V, Coghill TG et al (2012) The importance of plant genetic resources in agroecosystem. Rev Agric Rural Dev 1:302–308

Cermak T, Doyle EL, Christian M et al (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res 39:e82

Charmet G, Robert N, Perretant MR et al (2001) Marker assisted recurrent selection for cumulating QTLs for bread-making related traits. Euphytica 119:89–93

Chen X, Li H, Pandey MK et al (2016) Draft genome of the peanut A-genome progenitor (*Arachis duranensis*) provides insights into geocarpy, oil biosynthesis, and allergens. Proc Natl Acad Sci U S A 113:6785–6790

Cobb JN, Biswas PS, Platten JD (2019) Back to the future: revisiting MAS as a tool for modern plant breeding. Theor Appl Genet 132:647–667

Crossa J, Pérez-Rodríguez P, Cuevas J et al (2017) Genomic selection in plant breeding: methods, models, and perspectives. Trends Plant Sci 22:961–975

Datta SK, Baisakh N, Ramanathan V et al (2004) Transgenics in crop improvement plant. In: Jain HK, Kharkwal MC (eds) Breeding—Mendelian to molecular approaches, pp 333–371

Desta ZA, Ortiz R (2014) Genomic selection: genome-wide prediction in plant improvement. Trends Plant Sci 19:592–601

Dickmann DI (1985) The ideotype concept applied to forest trees. In: Cannell MGR, Jackson JE (eds) Attributes of trees as crop plants. Institute of Terrestrial Ecology, Penicuik, Scotland, pp 89–101

Dickmann DI, Gold MA, Flore JA (1994) The ideotype concept and the genetic improvement of tree crops. Plant Breed Rev 12:163–193

Donald CM (1968) The breeding of crop ideotypes. Euphytica 17:385–403

Dong OX, Ronald PC (2019) Genetic engineering for disease resistance in plants: recent progress and future perspectives. Plant Physiol 180:26–38

Dormatey R, Sun C, Ali K et al (2020) Gene pyramiding for sustainable crop improvement against biotic and abiotic stresses. Agronomy 10:1255

Finkel E (2009) With ‘phenomics,’ plant scientists hope to shift breeding into overdrive. Science 325:380–381

Galluzzi G, Seyoum A, Halewood M et al (2020) The role of genetic resources in breeding for climate change: the case of public breeding programmes in eighteen developing countries. Plan Theory 9:1129. https://doi.org/10.3390/plants9091129

Gauffreteau A (2018) Using ideotypes to support selection and recommendation of varieties. OCL 25(6):D602. https://doi.org/10.1051/ocl/2018042

Ghosh S, Watson A, Gonzalez-Navarro OE et al (2018) Speed breeding in growth chambers and glasshouses for crop breeding and model plant research. Nat Protoc 13:2944–2963

Godfray HCJ, Beddington JR, Crute IR et al (2010) Food security: the challenge of feeding 9 billion people. Science 327:812–818

Godiki Y, Bhanu AN, Singh MN (2016) Marker assisted recurrent selection: an overview. Adv Life Sci 5:6493–6499

Gupta PK, Rustgi S, Mir RR (2008) Array-based high-throughput DNA markers for crop improvement. Heredity 101:5–18

Gupta PK, Balyan HS, Mir RR et al (2011) QTL analysis, association mapping and marker-assisted selection for some quality traits in bread wheat—an overview of the work done at CCS university, Meerut. J Wheat Res 3:1–11

Gupta PK, Balyan HS, Gahlaut V et al (2012) Phenotyping, genetic dissection, and breeding for drought and heat tolerance in common wheat: status and prospects. Plant Breed Rev 36:85

Gupta PK, Kulwal PL, Mir RR (2013a) QTL mapping: methodology and applications in cereal breeding. In: Gupta PK, Varshney RK (eds) Cereal genomics II. Springer, Dordrecht, pp 275–318

Gupta PK, Rustgi S, Mir RR (2013b) Array-based high-throughput DNA markers and genotyping platforms for cereal genetics and genomics. In: Gupta PK, Varshney RK (eds) Cereal genomics II. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-6401-9_2

Gupta PK, Kulwal PL, Jaiswal V (2014) Association mapping in crop plants: opportunities and challenges. Adv Genet 85:109–147

Gupta PK, Kulwal PL, Jaiswal V (2019) Association mapping in plants in the post-GWAS genomics era. Adv Genet 104:75–154. https://doi.org/10.1016/bs.adgen.2018.12.001

Halewood M, Chiurugwi T, Hamilton RS, Kurtz B, Marden E, Welch E, Michiels F, Mozafari J, Sabran M, Patron N, Kersey P, Bastow R, Dorius S, Dias S, McCouch S, Powel W (2018) Plant genetic resources for food and agriculture: opportunities and challenges emerging from the science and information technology revolution. New Phytol 217:1407–1419

Harfouche AL, Jacobson DA, Kainer D et al (2019) Accelerating climate resilient plant breeding by applying next-generation artificial intelligence. Trends Biotechnol 37(11):1217–1235

Haussmann BI, Parzies HK, Presterl T et al (2004) Plant genetic resources in crop improvement. Plant Genet Resour 2:3–21

Heffner EL, Lorenz AJ, Jannink J-L et al (2010) Plant breeding with genomic selection: gain per unit time and cost. Crop Sci 50:1681–1690

Hickey LT, Germa SE, Diaz JE et al (2017) Speed breeding for multiple disease resistance in barley. Springer, New York

Hickey LT, Hafeez AN, Robinson H et al (2019) Breeding crops to feed 10 billion. Nat Biotechnol 37:744–754

Husaini AM, Rashid Z, Mir RR et al (2011) Approaches for gene targeting and targeted gene expression in plants. GM Crops 2:150–162

Jain M, Misra G, Patel RK et al (2013) A draft genome sequence of the pulse crop chickpea (*Cicer arietinum* L.). Plant J 74:715–729

Jha UC, Chaturvedi SK, Bohra A et al (2014) Abiotic stresses, constraints and improvement strategies in chickpea. Plant Breed 133:163–178

Jiang GL (2015) Molecular marker-assisted breeding: a plant breeder's review. In: Advances in plant breeding strategies: breeding, biotechnology and molecular tools. Springer, Cham, pp 431–472

Johnson R (2004) Marker assisted selection. Plant Breed Rev 24:293–309

Kärki L, Tigerstedt PMA (1985) Definition and exploitation of forest tree ideotypes in Finland. In: Cannell MGR, Jackson JE (eds) Attributes of trees as crop plants. Institute of Terrestrial Ecology, Penicuik, Scotland, pp 102–109

Khan AW, Garg V, Roorkiwal M et al (2020) Super-pangenome by integrating the wild side of a species for accelerated crop improvement. Trends Plant Sci 25:148–158

Khanday I, Skinner D, Yang B et al (2019) A maleexpressed rice embryogenic trigger redirected for asexual propagation through seeds. Nature 565:91–95

Kim MY, Lee S, Van K et al (2010) Whole-genome sequencing and intensive analysis of the undomesticated soybean (*Glycine soja* Sieb. and Zucc.) genome. Proc Natl Acad Sci U S A 107: 2032–22037

Kulwal PL (2016) Association mapping and genomic selection—where does sorghum stand? In: Rakshit S, Wang Y-H (eds) The sorghum genome. Springer, Cham, pp 137–148

Kulwal PL (2018) Trait mapping approaches through linkage mapping in plants. Adv Biochem Eng Biotechnol 164:53–82. https://doi.org/10.1007/10_2017_49

Kulwal PL, Mhase LB (2017) Protein content exhibits a significant positive correlation with seed weight in chickpea germplasm collection. Plant Genet Resour 15:283

Kulwal PL, Singh R (2021) Association mapping in plants. In: Tripodi P (ed) Crop breeding. Methods in molecular biology, vol 2264. Humana Press, New York. https://doi.org/10.1007/978-1-0716-1201-9_8

Kulwal P, Kumar N, Kumar A et al (2005) Gene networks in hexaploid wheat: interacting quantitative trait loci for grain protein content. Funct Integr Genomics 5:254–259

Kulwal PL, Mir RR, Kumar S et al (2010) QTL analysis and molecular breeding for seed dormancy and pre-harvest sprouting tolerance in bread wheat. J Plant Biol 37:59–74

Kulwal PL, Mahendar T, Varshney RK (2012) Genomics interventions in crop breeding for sustainable agriculture. Encycl Sustain Sci Technol 1:2527–2540

Kumar S, Kumar M, Mir RR et al (2021) Advances in molecular markers and their use in genetic improvement of wheat. In: Wani SH, Mohan A, Singh GP (eds) Physiological, molecular, and genetic perspectives of wheat improvement, pp 139–174

Kuriakose SV, Pushker R, Hyde EM (2020) Data-driven decisions for accelerated plant breeding. In: Gosal SS, Wani SH (eds) Accelerated plant breeding, vol vol 1. Springer, Cham, pp 89–119

Kushwaha UKS, Deo I, Jaiswal JP et al (2017) Role of bioinformatics in crop improvement. GJSFR-D 17:2249–4626

Li JF, Norville JE, Aach J et al (2013) Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. Nat Biotechnol 31:688–691

Lloyd A, Plaisier CL, Carroll D et al (2005) Targeted mutagenesis using zinc-finger nucleases in *Arabidopsis*. Proc Natl Acad Sci U S A 102:2232–2237

Lu Q, Li H, Hong Y et al (2018) Genome sequencing and analysis of the peanut B-genome progenitor (*Arachis ipaensis*). Front Plant Sci 3:604

Mahfouz MM, Li LX, Shamimuzzaman M et al (2011) De novo engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. Proc Natl Acad Sci U S A 108:2623–2628

Mascher M, Gundlach H, Himmelbach A et al (2017) A chromosome conformation capture ordered sequence of the barley genome. Nature 544:427–433

McCouch S, Navabi ZK, Abberton M et al (2020) Mobilizing crop biodiversity. Mol Plant 13: 1341–1344

Meena HP, Bainsla NK, Yadav DK (2017) Breeding for abiotic stress tolerance in crop plants. In: Yadav P, Kumar S, Jain V (eds) Recent advances in plant stress physiology, pp 329–377

Meuwissen T (2007) Genomic selection: marker assisted selection on a genomewide scale. J Anim Breed Genet 124:321–322

Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genomewide dense marker maps. Genetics 157:1819–1829

Mir RR, Varshney RK (2013) Future prospects of molecular markers in plants. In: Henry RJ (ed) Molecular markers in plants. Blackwell, Oxford, pp 169–190. ISBN 9781118473023

Mir RR, Kumar N, Jaiswal V et al (2012a) Genetic dissection of grain weight (GW) in bread wheat through QTL interval and association mapping. Mol Breed 29:963–972

Mir RR, Zaman-Allah M, Sreenivasulu N et al (2012b) Integrated genomics, physiology and breeding approaches for improving drought tolerance in crops. Theor Appl Genet 125:625–645

Mir RR, Hiremath PJ, Riera-Lizarazu O et al (2013) Evolving molecular marker technologies in plants: from RFLPs to GBS. In: Lübberstedt T, Varshney RK (eds) Diagnostics in plant breeding. Springer Science+Business, New York, pp 229–247. ISBN 978–94–007-5686-1; 978–94–007-5687-8

Mir RR, Choudhary N, Singh B et al (2015) Harnessing genomics through phenomics. In: Jitendar K et al (eds) Phenomics in crop plants: trends, options and limitations. Springer, New Delhi

Mir RR, Reynolds M, Pinto F et al (2019) High-throughput phenotyping for crop improvement in the genomics era. Plant Sci 282:60–72

Mir AH, Bhat MA, Dar SA et al (2021a) Assessment of cold tolerance in chickpea (*Cicer* spp.) grown under cold/freezing weather conditions of north-western Himalayas of Jammu and Kashmir, India. Physiol Mol Biol Plants 27:1105–1118

Mir RR, Kumar S, Shafi S (2021b) Genetic dissection for yield and yield-related traits in bread wheat (*Triticum aestivum* L.). In: Wani SH, Mohan A, Singh GP (eds) Physiological, molecular, and genetic perspectives of wheat improvement, pp 209–227

Munck L (2009) Breeding for quality traits in cereals: a revised outlook on old and new tools for integrated breeding. In: Carena MJ (ed) Cereals: handbook of plant breeding. Springer Publishers, New York, pp 333–366

Murphy KP (2012) Machine learning: a probabilistic perspective. MIT Press, Cambridge

Nekrasov V, Staskawicz B, Weigel D et al (2013) Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. Nat Biotechnol 31:691–693

Niazian M, Niedbała G (2020) Machine learning for plant breeding and biotechnology. Agriculture 10:436

Pandey MK, Chaudhari S, Jarquin D et al (2020) Genome-based trait prediction in multi-environment breeding trials in groundnut. Theor Appl Genet 133:3101–3117

Parmley KA, Higgins RH, Ganapathysubramanian B et al (2019) Machine learning approach for prescriptive plant breeding. Sci Rep 9:17132

Rahim MS, Bhandawat A, Rana N et al (2020) Genomic selection in cereal crops: methods and applications. In: Gosal SS, Wani SH (eds) Accelerated plant breeding, vol 1. Springer, Cham, pp 51–88

Rahman M, Davies P, Bansal U et al (2020) Marker-assisted recurrent selection improves the crown rot resistance of bread wheat. Mol Breed 40:1–4

Ramstein GP, Jensen SE, Buckler ES (2019) Breaking the curse of dimensionality to identify causal variants in breeding 4. Theor Appl Genet 132:559–567

Rani A, Devi P, Jha UC et al (2020) Developing climate-resilient chickpea involving physiological and molecular approaches with a focus on temperature and drought stresses. Front Plant Sci 10: 1759

Rasmusson DC (1987) An evaluation of ideotype breeding. Crop Sci 27:1140–1146

Ribaut JM, Vicente MC, Delannay X (2010) Molecular breeding in developing countries: challenges and perspectives. Curr Opin Plant Biol 13:1–6

Richardson KL, Vales MI, Kling JG et al (2006) Pyramiding and dissecting disease resistance QTL to barley stripe rust. Theor Appl Genet 113:485–495

Rodriguez-Leal D, Lemmon ZH, Man J et al (2017) Engineering quantitative trait variation for crop improvement by genome editing. Cell 171:470–480

Roorkiwal M, Bharadwaj C, Barmukh R et al (2020) Integrating genomics for chickpea improvement: achievements and opportunities. Theor Appl Genet 133:1703–1720

Schmutz J, Cannon SB, Schlueter J et al (2010) Genome sequence of the palaeopolyploid soybean. Nature 463:178–183

Schmutz J, McClean PE, Mamidi S et al (2014) A reference genome for common bean and genome-wide analysis of dual domestications. Nat Genet 46:707–713

Schnable PS, Ware D, Fulton RS et al (2009) The B73 maize genome: complexity, diversity, and dynamics. J Sci 326:1112–1115

Shan QW, Wang YP, Li J et al (2013) Targeted genome modification of crop plants using a CRISPR-Cas system. Nat Biotechnol 31:686–688

Siddique KHM, Sedgley RH (1987) Canopy development modifies the water economy of chickpea (*Cicer arietinum* L.). Aust J Agr Res 37:599–610

Singh NK, Gupta DK, Jayaswal PK et al (2012) The first draft of the pigeonpea genome sequence. J Plant Biochem Biotechnol 21:98–11

Tan Y-y, Du H, Wu X et al (2020) Gene editing: an instrument for practical application of gene biology to plant breeding. J Zhejiang Univ Sci B (Biomed Biotechnol) 21:460–473

Tester M, Langridge P (2010) Breeding technologies to increase crop production in a changing world. Science 327:818–822

Thudi M, Palakurthi R, Schnable JC et al (2021) Genomic resources in plant breeding for sustainable agriculture. J Plant Physiol 257:153351

Tyagi S, Sharma S, Ganie SA et al (2019) Plant microRNAs: biogenesis, gene silencing, web-based analysis tools and their use as molecular markers. 3 Biotech 9:413

Tyagi S, Kumar A, Gautam T et al (2021) Development and use of miRNA-derived SSR markers for the study of genetic diversity, population structure, and characterization of genotypes for breeding heat tolerant wheat. PLoS One 16:e0231063

Varshney RK, Glaszmann JC, Leung H, Ribaut JM (2010) More genomic resources for less-studied crops. Trends Biotechnol 28:452–460

Varshney RK, Chen W, Li Y et al (2012) Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. Nat Biotechnol 30:83–89

Varshney RK, Mohan SM, Gaur PM et al (2013a) Achievements and prospects of genomics-assisted breeding in three legume crops of the semi-arid tropics. Biotechnol Adv 31(8): 1120–1134

Varshney RK, Song C, Saxena RK et al (2013b) Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement. Nat Biotechnol 31:240–246

Varshney RK, Thudi M, Nayak SN, Gaur PM, Kashiwagi J, Krishnamurthy L, Jaganathan D, Koppolu J, Bohra A, Tripathi S, Rathore A, Jukanti AK, Jayalakshmi V, Vemula A, Singh SJ, Yasin M, Sheshshayee MS, Viswanatha KP (2014) Genetic dissection of drought tolerance in chickpea (*Cicer arietinum* L.). Theor Appl Genet 127:445–462

Varshney RK, Godwin ID, Mohapatra T et al (2019) A SWEET solution to rice blight. Nat Biotechnol 37:1280–1282

Varshney RK, Bohra A, Yu J et al (2021) Designing future crops: genomics-assisted breeding comes of age. Trends Plant Sci 26(6):631–649. https://doi.org/10.1016/j.tplants.2021.03.010

Vassilev D, Leunissen J, Atanassov A et al (2005) Application of bioinformatics in plant breeding. Biotechnol Biotechnol Equip 19(sup3):139–152

Wallace JG, Rodgers-Melnick E, Buckler ES (2018) On the road to breeding 4.0: unraveling the good, the bad, and the boring of crop quantitative genomics. Annu Rev Genet 52:421–444

Wang C et al (2019) Clonal seeds from hybrid rice by simultaneous genome engineering of meiosis and fertilization genes. Nat Biotechnol 37:283–286

Watson A, Ghosh S, Williams MJ et al (2018) Speed breeding is a powerful tool to accelerate crop research and breeding. Nat Plants 4:23–29

Wong CK, Bernardo R (2008) Genome-wide selection in oil palm: increasing selection gain per unit time and cost with small populations. Theor Appl Genet 116:815–824

Xu Y, Liu X, Fu J et al (2020) Enhancing genetic gain through genomic selection: from livestock to plants. Plant Commun 1:100005

Yin D, Ji C, Ma X et al (2018) Genome of an allotetraploid wild peanut *Arachis monticola*: a *de novo* assembly. Gigascience 7:giy66

Zeng X, Long H, Wang Z et al (2015) The draft genome of Tibetan hulless barley reveals adaptive patterns to the high stressful Tibetan plateau. Proc Natl Acad Sci U S A 112:1095–1100

Zeng X, Guo Y, Xu Q et al (2018) Origin and evolution of qingke barley in Tibet. Nat Commun 9: 1–11

Zhang Y, Massel K, Godwin ID, Gao C (2018) Applications and potential of genome editing in crop improvement. Genome Biol 19:210

Zhang YX, Malzahn AA, Sretenovic S et al (2019) The emerging and uncultivated potential of CRSIPR technology in plant science. Nat Plants 5:778–794

Zhang Y, Pribil M, Palmgren M et al (2020) A CRISPR way for accelerating improvement of food crops. Nat Food 1(4):200–205. https://doi.org/10.1038/s43016-020-0051-8

Zhuang W, Chen H, Yang M et al (2019) The genome of cultivated peanut provides insight into legume karyotypes, polyploid evolution and crop domestication. Nat Genet 51:865–876

Zimny T, Sowa S, Tyczewska A et al (2019) Certain new plant breeding techniques and their marketability in the context of EU GMO legislation—recent developments. N Biotechnol 51: 49–56

Zong Y et al (2018) Efficient C-to-T base editing in plants using a fusion of nCas9 and human APOBEC3A. Nat Biotechnol 36:950–953