Plant Science 242 (2016) 3–13

Contents lists available at ScienceDirect

Plant Science

journal homepage: www.elsevier.com/locate/plantsci

Next generation breeding

Delfina Barabaschi[a], Alessandro Tondelli[a], Francesca Desiderio[a], Andrea Volante[b], Patrizia Vaccino[c], Giampiero Valè[b], Luigi Cattivelli[a,*]

[a] Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria, Genomics Research Centre, Via San Protaso 302, 29017 Fiorenzuola d'Arda, Italy
[b] Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria, Rice Research Unit, SS 11 to Torino Km 2.5, 13100 Vercelli, Italy
[c] Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria, Research Unit for Cereal Selection in Continental areas, via R. Forlani, e, 26866 S. Angelo Lodigiano, Italy

ARTICLE INFO

Article history:
Received 17 April 2015
Received in revised form 10 July 2015
Accepted 11 July 2015
Available online 19 July 2015

Keywords:
Next generation sequencing
Molecular markers
Allele diversity
Genomic selection
Genome editing
Plant breeding

ABSTRACT

The genomic revolution of the past decade has greatly improved our understanding of the genetic make-up of living organisms. The sequencing of crop genomes has completely changed our vision and interpretation of genome organization and evolution. Re-sequencing allows the identification of an unlimited number of markers as well as the analysis of germplasm allelic diversity based on allele mining approaches. High throughput marker technologies coupled with advanced phenotyping platforms provide new opportunities for discovering marker-trait associations which can sustain genomic-assisted breeding. The availability of genome sequencing information is enabling genome editing (site-specific mutagenesis), to obtain gene sequences desired by breeders. This review illustrates how next generation sequencing-derived information can be used to tailor genomic tools for different breeders' needs to revolutionize crop improvement.

1. Introduction

The development of next generation sequencing (NGS) technologies has made DNA sequencing high throughput and very cost effective. Consequently, many opportunities are being opened to explore the relationships between genetic and phenotypic diversity with a resolution never reached before. Reference genome sequences have been published for many crop species [1] and many more genome sequencing projects are in progress (http://www.ncbi.nlm.nih.gov/genomes/leuks.cgi; http://plants.ensembl.org/index.html; http://phytozome.jgi.doe.gov/pz/portal.html). The sequences of crop genomes provide a useful starting point to explore genome organization and evolution and provide insight into genetic variation through partial or complete re-sequencing of different accessions [2]. Re-sequencing, leading to arrays of high-density single-nucleotide polymorphisms (SNPs), is allowing whole-genome scans to identify haplotype blocks that are significantly correlated with quantitative trait variation. The distribution of low cost sequencing technologies offers new opportunities to shape genetic diversity according to the needs of modern agriculture and, in turn, has a number of practical consequences for plant breeding: *i*) the analysis of genetic diversity can be based on genome re-sequencing; *ii*) genome wide association studies (GWAS) become an attractive approach for quantitative trait loci (QTLs) mapping in plants since broad genetic resources can be scanned for marker-trait association without any limitation of marker availability; *iii*) the great number of markers support genomic selection; and *iv*) the genome sequences allow the targeted modification of specific genes through genome editing technologies or identification of suitable mutations within mutagenized populations, resulting in the introduction of new allelic variants in the genome of cultivated varieties. Conversely, these achievements highlight new bottlenecks for breeding progress, particularly the phenotyping capacity (in terms of both precision and throughput [3]), and recombination frequency [4].

Over the last decades, plant breeding has moved from being a completely phenotyping-based process to having an increased reliance on some level of genotype-based selection [5]. This trend is

Abbreviations: CRISPR/Cas9, Clustered regularly interspaced short palindromic repeats/CRISPR-associated; GBS, genotyping by sequencing; GEBV, genomic estimated breeding values; GS, genomic selection; GWAS, genome-wide association studies; LD, linkage disequilibrium; MAGIC, multi-parent advanced generation intercross; MAS, marker-assisted selection; NAM, nested association mapping; NGS, next generation sequencing; RIL, recombinant inbreed line; QTL, quantitative trait locus; RAD, restriction-site associated DNA; SNP, single-nucleotide polymorphism; TALEN, transcription activator-like effector nucleases; TILLING, targeting induced local lesions in genomes; ZFN, zinc finger nucleases.

* Corresponding author.
E-mail address: luigi.cattivelli@entecra.it (L. Cattivelli).

http://dx.doi.org/10.1016/j.plantsci.2015.07.010
0168-9452/

expected to increase in the coming years as the NGS-based knowledge will be translated into "Next Generation Breeding". In this review, we consider current trends and future prospects for the application of genomic instruments in the improvement of plant breeding performance.

2. Genome sequencing and sequence-based markers

Molecular markers have been available for more than 25 years, nevertheless the advent of NGS represented a breakthrough in this field. Before NGS, a typical linkage map was based on few hundreds markers. In the age of NGS, thousands of markers can be easily included in any map, including in species with little a priori genome information available. With NGS technologies the DNA marker identification has shifted from fragment-based (RFLPs, AFLPs, microsatellites) to sequence-based polymorphisms (SNPs). Uniplex or multiplex SNP genotyping platforms that combine a variety of chemistries, detection methods, and reaction formats are available. Uniplex SNP genotyping platforms are more suitable for applications requiring small to moderate numbers of SNPs for a large number of samples. TaqMan™ (http://www.appliedbiosystems.com) and competitive allele-specific PCR (KASP™, http://www.lgcgenomics.com) are among the most popular techniques on the market. Multiplexed SNP analysis can be run on middle throughput platforms with a capacity of a few hundred SNPs per run (e.g. Illumina BeadXpress, Fluidigm EP1) or with high-throughput array-based technologies capable of generating between a few thousand to over one million SNPs per run (e.g. Illumina BeadArray™, Affymetrix GeneChip™ technology). The advent of high-density SNP arrays coupled with powerful computational pipelines has allowed the fast and easy scoring of large set of markers across many genotypes. Medium or high density arrays are available for many crop species, e.g. grapevine [6], maize [7], tomato [8], peach [9], soybean [10], barley [11], rice [12], wheat [13] and apple [14].

Nevertheless, the production of a high-quality array requires a substantial investment of resources, and the SNP panel, optimized for the population used to develop it, might be biased toward particular panels of germplasm. To circumvent these limitations, NGS technologies offer the possibility of shifting from array-based genotyping assays with pre-defined SNP panels to the direct sequencing of the populations of interest [15], producing a genome-wide and unbiased set of markers. These techniques employ a reduced genome representation achieved through restriction enzyme digestion and subsequent adaptor-mediated PCR amplification, and require no a priori knowledge of the SNPs being interrogated, making them useful for genetic analysis in species where no reference sequence is available. Among them, restriction-site associated DNA sequencing (RAD) [16] and genotyping by sequencing (GBS) [17] have been adopted in plants [18–22]. Furthermore, a strategy based on low coverage genome sequencing of all genotypes from a segregating population (POPSEQ) was recently employed for the development of high density genetic maps. POPSEQ was used to explore the organization of the gene space within the large, complex and highly repetitive barley genome [23], and contributed to the assembling of the hexaploid wheat genome [24].

3. Mining plant diversity: from genotype to phenotype

Many valuable genes and alleles are stored in seed bank collections, hidden in cultivars, landraces, mutagenized populations and wild species. The identification of these genes requires both genome information and phenotyping capacities. With the advent of NGS technologies a different dimension to the exploration of plant diversity arose. Extensive insights into plant genome composition and organization have been gained from the genome sequencing and new findings on plant origin and evolution (genome duplication, ancestral re-arrangements and polyploidization events) have been revealed [2]. An in silico paleogenomic study based on a deep comparison of monocot and eudicot genomes, allowed the reconstruction of ancestral protochromosome segments and a description of the evolutionary dynamics leading to the present-day genomes, their genome organization and regulation [25].

Since a single reference genome is not enough to represent the diversity within a species [26], the re-sequencing of different cultivars, landraces and wild accessions assumes an important role to reveal domestication events [27], identify gene diversification and variations [28] and explain heterosis mechanisms [29]. A most striking example is the "3,000 Rice Genomes Project", an initiative dedicated to the re-sequencing of 3000 rice accessions selected to represent the genetic and functional rice diversity available worldwide [30]. With re-sequencing information of many accessions, a strategy can be applied to search for allele diversity at candidate gene loci for which a clear association with specific phenotypic traits is known. This allele mining strategy can help trace the evolution of alleles, identify new useful haplotypes and guide the development of allele-specific markers for use in marker-assisted selection (MAS). For instance, following an allele mining approach several resistance gene homologues and functional resistance genes have been isolated in potato [31], wheat [32], rice [33,34], and barley [35].

When re-sequencing is applied to TILLING populations (TILLING-by-Sequencing), it allows a fast genome-wide identification of mutations [36,37]. Enhanced opportunities for functional genomics come from a genome-wide discovery pipeline of induced mutations based on multiplexed (10- to 30- fold) exome capture and sequencing. This method called MAPS (mutations and polymorphisms surveyor), was used to identify about 18,000 mutations in 72 independent M2 rice lines, of which >2,600 appeared to be detrimental to gene function [38].

Beside the description of allele diversity, genome re-sequencing coupled with de novo assembly of the sequences not matching the reference genome offers the possibility of harnessing the gene repertoire from wild relatives of crops leading to the description of their pan-genomes. Pan genome refers to the full complement of genes in a group of individuals (e.g. species) and consists of a core genome containing DNA sequences shared by all the genotypes and of a dispensable genome composed of partially shared genomic features (i.e. present in only some genotypes) [39]. For instance, the re-sequencing of seven accessions of *Glycine soja* led to the identification in the dispensable genome of many genes that had structural variant involved in the adaptation to the environment (R-genes, flowering time-related genes, genes involved in oil and fatty acid content), and which were therefore potentially useful in crop breeding [40].

Since the observed phenotypic diversity is not completely explained by variation at the genomic level, the establishment of the pan-genome should also be supported by the analysis of differences in gene expression and regulation. In maize the whole-seedling transcriptome variations among 503 diverse inbred lines have been described by RNA-sequencing. Besides the pan-genome, the analysis of the maize pan-transcriptome helps to explain some components of phenotypic variation (e.g. heterosis) in terms of variation in abundance of transcripts and their alternatively spliced forms in inbred lines [41]. In addition to gene regulation, epigenetic regulation represents another crucial factor to be considered for advanced crop breeding, especially for physiological phenotypes, having a fundamental role in regulating gene expression in response to developmental and environment changes [42].

Producing high throughput sequence information is only one component of next generation breeding. A major constraint is the ability to associate all these genomic data with systematic and robust characterization of phenotypes for a wide range of traits and conditions. Innovations in this field are expected from high-throughput phenotyping platforms that employ remote sensing and imaging techniques (based on visible/near-infrared and far-infrared radiation, reflected and emitted by the plants, respectively), and high-performance data recording and computing to evaluate plant performance in field or controlled environments. Automated greenhouse systems (LemnaTec system, http://www.lemnatec.com/) coupled with innovative image acquisition techniques (Phenoscope [43], RootReader3D [44], LemnaTec Scanalyzer 3D, http://www.lemnatec.de/scanalyzer_gh.htm), advanced image analysis pipelines (HTPheno [45]), and specific software applications (RootNav [46], Integrated Analysis Platform [47]) allow the non-destructive recording of a wide range of phenotypic traits over time (e.g. in barley [48] and tomato [49]). Nonetheless, efforts are still required to implement high-throughput and cost effective phenotyping in field conditions [50] and there is a great interest in drones or aircrafts as remote phenotyping sensor platforms able to monitor the field trials performance throughout the growing season (e.g. in maize [51]).

4. Collecting genetic information through meta-analysis

The statistical combination of a huge amount of molecular and phenotypic data, obtained from publications and omics databases, provides opportunities to unravel complex traits in crops through genome-wide meta-analysis, and is becoming a promising approach for crop breeding. Meta-analysis, with the support of dedicated statistical procedures, enables the evaluation of key genetic, genomic and environmental variables and their impact on crop agronomic performance, by exploiting and then integrating datasets from different studies using various multiple methodologies. Nonetheless, caution is needed in the implementation of meta-analysis to avoid biases in the interpretation of the results arising from inappropriate assumptions [52].

Of particular relevance for breeders are the QTL meta-analyses of QTL information, which allows QTL locations to be compared for a trait between populations and/or to prioritize candidate genes [53]. In rice, the genome sequence was used for a meta-analysis of QTLs involved in partial and complete blast resistance. This work involved an analysis of the co-localization of blast resistance genes and QTLs and indentification of candidate genes on the reference genome. A few partial-resistance QTLs active against most of the strains tested, and observable in several independent experiments, were identified, and represent worthy targets for future breeding programs [54]. In bread wheat, a meta-analysis for crop height variation loci was applied to four doubled populations with parents representing a wide diversity of European winter wheat. Out of the 16 meta-QTLs identified in the consensus linkage map, those having additive effects equivalent to height-reducing alleles (*Rht-D1* and *Rht8*) could be exploited in wheat breeding to modify crop stature and hence yield and biomass [55]. In another study, over 3100 maize individuals from 18 bi-parental populations were genotyped with the same SNP platform and evaluated in environments with different levels of available water. The data were then summarized into 68 meta-QTLs for grain yield and anthesis silking interval. Four meta-QTLs for grain yield detected under different environments and in 6 populations were identified as promising for pyramiding into breeding lines [56]. In durum wheat, when more than 80 QTLs and 51 resistance genes for powdery mildew from 62 different mapping populations were projected onto the same consensus map, they were summarized into 24 meta-QTLs [57]. This highlighted the most stable and relevant ones, a simplification of great relevance for breeding.

To ensure the accessibility and exploitation of all genome and phenotypic information from a broader number of researchers and breeders, it is of great value to have website portals to collect all information available for a given crop, as underway in rice [30] and wheat (http://wheat-urgi.versailles.inra.fr/Projects/Wheat-Initiative/Wheat-Information-System). These meta-data collections are expected to reduce the information redundancy and highlight what is still missing, thereby helping to decipher the genetic variation (SNPs, insertion or deletions, structural variants) between and within different populations, and to contribute in the association mapping effort by enabling the extraction of the haplotype structure and haplotype maps based on LD.

5. Marker-trait associations in large germplasm collections

The genetic bases of many traits have been conventionally dissected by linkage analysis in segregating mapping populations (e.g. Double Haploids -DHs, Recombinant Inbreed lines -RILs) or using nearly isogenic lines (NILs) developed using several backcrosses [58]. Nevertheless, the estimated effects are specific to the same or genetically related populations and are often not transferable to other genetic backgrounds, thus limiting their practical application for breeding purposes [58]. In the last decade, the availability of high resolution and cost effective genotyping platforms have opened the way to GWAS. By exploiting LD between markers and traits across all chromosomes, GWAS aims at genetically scrutinizing complex phenotypes in natural or ad hoc generated populations, and it has been widely adopted in different plant species to overcome some of the constraints inherent to bi-parental linkage mapping, such as the limited genetic diversity explored [59,60]. Moreover, the long history of recombination events captured in large germplasm collections, when combined with dense marker coverage, permit increased genetic resolution, sometimes to a level that allows a causative sequence variant to be identified [61,62]. Nevertheless, some drawbacks have to be considered: *i)* LD levels, and hence the mapping resolution, can vary not only among species (e.g. selfing vs. outcrossing), but also among populations within one species and among different regions within a given genome [63–65]; *ii)* population structure may lead to spurious associations; and *iii)* the effect of rare alleles (even if large) might not be detectable by GWAS analysis. The power of detecting significant marker-trait associations depends on the quality of the phenotypic data, sample size, the genetic architecture and heritability of the trait [61,66].

Successful examples of GWAS in crop species have been recently reviewed [60,62]. As a whole, these studies provide breeders with plenty of marker-traits associations that may be directly exploited for crop design since they are applicable to a much wider germplasm base, provided that high LD is maintained between the causal gene and the significant markers in the breeding materials. Despite the high number of GWAS done in crop plants, only in few cases has the effect of an underlying candidate gene been verified. In fact, several independent information pieces of evidence are often necessary to definitively assign SNP association signals to genes and identify the causal mutation. These pieces of evidence can include loss of function mutants [67], over-expression lines, expression data, proteome or metabolome data [68,69], sequencing of candidate genes in diverse germplasm collections, including crop wild relatives [11], and linkage mapping and map-based cloning in experimental populations [68].

New crossing schemes and types of experimental populations have been suggested to overcome some of the limitations (e.g. population structure) that are encountered when GWAS is run with

panels of natural germplasm,breeding lines or varieties, and to increase the statistical power and mapping resolution with respect to bi-parental populations, hence the ability to identify genes underlying phenotypic variation [70]. Multi-parent Advanced Generation Intercross (MAGIC) populations are created by intercrossing multiple parental lines and self-crossing the progeny to generate RILs. Multiple founders capture more allelic diversity (including rare alleles) whereas the multiple cycles of intercrossing give greater opportunities for recombination and hence, greater precision in QTL location. In rice 4 multi-parent populations named *indica*-MAGIC (8 *indica* parents), *japonica*-MAGIC (8 *japonica* parents), MAGIC-plus (extended *indica* MAGIC with two extra rounds of intermating) and Global-MAGIC (16 parents, 8 *indica* and 8 *japonica*) have been developed to explore many desirable traits (biotic/abiotic stress tolerance, yield, and grain quality) [71]. Similarly, a wheat MAGIC population was developed using eight winter wheat lines selected for yield capacity, quality and disease resistance. The genotypic data of the 1091 F_7 lines generated demonstrate that the population is highly recombined making it a powerful tool for the genetic dissection of the characters in wheat [72]. In Nested Association Mapping (NAM) populations a number of founder lines are crossed with the same reference line to develop sets of related (half-sib) mapping progenies. The advantage of this population derives from the the ability to incorporate a large number of alleles from the gene pool. In maize, a NAM population was developed crossing 25 inbred lines to the B73 reference line [73]. About 5000 RILs were obtained and investigated for QTLs controlling the developmental timing of the juvenile to adult transition [74] as well as kernel composition [75]. A similar resource is under development in barley from crosses of the elite barley cultivar Barke with 25 wild barley donors [76]. The genotyping of the 1420 BC_1S_3 lines will simultaneously enable study of the diversity of barley wild relatives, identification of genes controlling agronomic traits and the transfer of favorable exotic alleles into the elite barley gene pool.

6. Genome-wide prediction of breeding value and genomic selection

There are two main strategies to assist breeding with molecular selection: to use molecular markers that map near or within specific loci with known phenotypic effects (marker-assisted selection, MAS) or to exploit all available markers as predictors of breeding value (genomic selection, GS). MAS is used to drive the selection of a relative small set of genes having major phenotypic effects [77,78], and much information on these tools is available also through crop-dedicated websites (e.g. in wheat http://maswheat.ucdavis.edu/; in barley http://www.barleycap.org/). Nevertheless, frequently the success of new crop varieties is based on particular combinations of many small-effect loci (QTLs). In GS, all locus, haplotype and marker effects are estimated across the entire genome to calculate the genomic estimated breeding values (GEBVs) in a population of individuals representative of the breeding program in question (often referred as training populations) for which both phenotypic and genotypic data are known [79,80]. Accuracy of GEBV prediction is strongly affected by both the marker density and the rates of LD decay across the genome. The evaluation of the inter-marker coefficient of determination, r^2, can provide useful information about the marker density required to obtain sufficient GEBV predictions [79]. The training population is used to estimate model parameters that will be subsequently used to calculate GEBVs of breeding materials for which only genotypic data are available and to select the individuals for advancement in the breeding cycle.

Different statistical models are initially tested using the genotypic and phenotypic data from the training populations to find the one that predicts GEBV most accurately, as defined by the correlation between the GEBV and the true breeding value [79]. A number of statistical models have been proposed, mainly based on corrected linear regression, best linear unbiased prediction (BLUP) and Bayesian regression methods [81,82]. The prediction accuracy of the different methods is debated; while in some cases all the models gave similar accuracy in estimation [83], other studies evidenced how different population features (LD structure, presence of epistasis and relationship between the training and validation sets) as well as trait characteristics (genetic architecture, heritability) may influence the relative performance of the prediction methods [82,84]. Once the model providing the highest accuracy is identified, GS would allow selection of lines without utilization of phenotypic data through the model predicting the individual GEBVs [79,81,82,85].

Linkage disequilibrium is a parameter of great importance in the designing of a GS approach. Significant LD in outcrossed species extends for 0.1–1.5 kb in maize [63] and 15–20 kb for sorghum [86], while it extends for a considerably greater distances in self-crossed species like rice (from 75 kb in the *indica* background to 500 kb in temperate *japonica*), or barley where LD decays across 5–10 cM, representing approximately 20–40 Mb [87]. LD can also vary depending on the genomic region and the population structure. For species whose LD decays rapidly among unrelated individuals, a lower number of parental lines can be screened without lowering the detection power, and a increased number of markers can be employed. The size of the training and breeding population is also a critical issue. While larger training populations improve the accuracy of GEBV predictions [88], the training/breeding population size ratio is suggested to be more crucial. In general, a higher training/breeding population ratio is required for accurate GEBV prediction in case of greater genetic diversity, smaller-sized breeding populations, lower heritability of traits and larger numbers of existing QTL [78]. Co-dominant markers (e.g. SNPs, SSRs) provide a more accurate estimation of GEBV than dominant markers (e.g. DArT markers) due to the higher LD detection power and the accuracy can be further improved by considering haplotypes [89]. Bi-allelic markers (SNPs, DArTs, GBS and RAD markers) provide individually less information than multi-allelic ones (SSRs), and therefore require more data to achieve a similar accuracy. Nevertheless, the cost per data-point of the different marker types makes SNPs the marker of choice for GS.

GS has been applied to several traits in a variety of species including maize, barley, bread wheat and rice. (Table 1). Overall, these studies suggest a wide applicability of GS even for species with large and/or complex genomes such as bread wheat, sugarcane or maize. In many studies, GS has provided reliable information to an extent comparable or even higher than more traditional selection approaches; indeed the correlation between true breeding value and the GEBV has reached levels of 0.85 even for polygenic low heritability traits [79]. When the response to genome-wide selection was assessed in comparison with MAS in maize, the response to GS was significantly better, with a value depending on the heritability and number of considered QTLs [90]. Similarly, Rutkosky et al. [91] compared GS vs. selection based on a set of QTL associated markers or vs. phenotypic selection for traits related to Fusarium head blight resistance in wheat. In all comparisons the prediction accuracy was higher for GS.

Interestingly, GS provides a higher accuracy in the estimation of GEBV in plants than in animals, although the number of molecular markers used is generally lower. This is probably due to the narrower genetic base (and consequently a lower genetic diversity) of plant genetic materials, which are derived in many cases from a small number of parental varieties and a greater bottleneck in the breeding materials [80].

Table 1
–Examples of genomic selection in plant breeding

Species	Trait	Reference
Maize (*Zea mays* L.)	Grain yield, anthesis date, and anthesis-silking interval	[136]
	Anthesis date, grain yield, plant height under normal and water stressed conditions	[22]
Barley (*Hordeum vulgare* L.)	Fusarium head blight resistance, DON accumulation	[137]
	Yield, plant height, fusarium headblight resistance, DON accumulation, morphological traits, virus resistance	[138,139]
Wheat (*Triticum aestivum* L.)	Yield, thousand-kernel weight, number of kernels per spike, heading date	[140]
	Resistance to leaf rust, stem rust, stripe rust	[141]
Rice (*Oryza sativa* L.)	Days to heading, culm length, panicle number and length, grain length and width, brown rice length and width	[83]
	Yield, number of tillers per plant, number of grains per panicle, 1000 grain weight	[142]
	Grain yield, flowering time, plant height	[21]
Oat (*Avena sativa* L.)	β-Glucan percentage, yield, heading date, groat percentage, plant height	[143]
Soybean (*Glycine max* L.)	Seed weight	[144]
Sugarcane (*Saccharum officinarum* L.)	Sugar contents, digestibility and composition of the bagasse, plant morphology and disease resistance	[83]
Sugarbeet (*Beta vulgaris* L.)	Sugar yield and content, root yield, potassium and sodium content, α-amino nitrogen content	[145]
Apple (*Malus x domestica* Borkh.)	Fruit quality traits	[146]

7. Plant improvement through genome editing

Genome editing, i.e. the targeted modification of a gene, allows generation of new allelic variants in the genome of cultivated species; it represents an alternative to standard breeding processes based on recombination and, to some extent, to genetic transformation. Genome editing relies on the induction of double strand breaks in DNA in a targeted part of the genome using an engineered DNA-binding protein. Sequence-specific nucleases, including zinc finger nucleases (ZFN), and transcription activator like effector nucleases (TALEN) have been initially proposed for targeted genome editing in eukaryotic organisms [92]. More recently, another double strand breaks-based technology for genome editing, the CRISPR/Cas9 system, has been developed based on the bacterial and archaeal clustered regularly interspaced short palindromic repeats (CRISPR) adaptive immune system. The system exploits the endonuclease activity of CRISPR-associated (Cas) proteins, with sequence specificity directed by CRISPR RNAs (crRNAs) [92,93]. Double-strand breaks at specific genomic sites can introduce a mutation at the DNA break site via the error-prone non-homologous end-joining pathway. This is the most common system acting in plants and frequently induces small insertions/deletions which result in an array of mutations at the targeted gene. Double strand breaks in multiple sites can also result in homologous recombination between chromosomal DNA and foreign donor DNA through the homologous recombination pathway. In this way more significant modifications of the target sequence are possible (gene stacking, allele substitutions) [94]. To ensure specificity and a low rate of off-target cleaving for this system, the most crucial point appears to be the careful selection of the gRNA sequence [95]. Being based on nucleotide-nucleotide interactions, the target sequence for CRISPR/Cas9 system can be designed in a more predictable way compared to TALEN or ZFN. In this system un-specific mutations can almost completely be avoided in plants. Other possible strategies to increase specificity are based on the use of dimeric, partially inactivated CRISPR/Cas9 complexes, which require two precisely disposed recognition sites on the genome. These methods have been reported to increase specificity from tens to hundreds of times compared with fully-active cleavage complexes. The regulation of Cas9 and gRNA expression is also a crucial point; indeed, overexpression of these molecules is often related to an increased rate of non-target mutations [95]. Transgenic plant lines transiently or stably carrying the sequence-specific nucleases (able to induce the desired mutations) can be generated using biolistics [96], *Agrobacterium* or protoplast based methods of transformation [93].

Genome editing relies on very accurate genome sequence information for the precise determination of the target site particularly if the target gene is part of a multigene family or if duplications or homeologus copies are present. Availability of genome sequence for many crops will facilitate genome editing approaches for plant improvement, although few successful examples are yet known. ZFN-mediated mutagenesis was employed to engineer tobacco plants in the acetolactate synthase (ALS) genes (*SuRA* and *SuRB*) inducing herbicide resistance in transformed plants [97]. ZFNs were also used for the stacking of multiple herbicide resistance genes (*pat, aad1*) in maize: the accurate targeting of the ZFN activity allowed the insertion of the two genes in close proximity, and the entire array of transgenes segregated as a single locus in subsequent generations [97]. A variety of applications are reported for TALEN [98]. Shan et al. [99] targeted 4 rice genes related to morphological and quality traits, and 8 *Brachypodium* genes involved in hormone balance and gene regulation were targeted, for TALEN-directed mutagenesis. Short deletions were most frequently obtained, but the use of multiple TALEN constructs targeting different sites of the same gene also allowed large deletions to be obtained. TALEN was also used to knockout the *PROCERA* gene involved in gibberellin signaling in tomato. In the progeny plants, individual carrying the mutations (e. g. with modified leaf shape, growth rate and internode distance) but which were missing the TALEN construct by segregation, were selected [100].

The CRISPR/Cas9 system was tested for inducing targeted deletions in the *inositol oxygenase* and *phytoene desaturase* genes in tobacco and wheat suspension cell cultures [101]. The construct was targeted to one or more sites of the two genes in different combinations and the plants obtained by cell culture regeneration showed typically 20–40 bp deletions or insertions. The deletion of a complete gene was also possible by inducing simultaneous cleavage in two different sites within the gene sequence, and the method also allowed mutations to be introduced in multiple genes using the same expression cassette. Knockout mutations in the barley MLO gene is known to confer broad resistance to powdery mildew [102], therefore equivalent mutations accumulated in the three homeologous MLO genes of the wheat genome were expected to result in a similar phenotype. Indeed Wang et al. [96] used CRISPR-Cas9 for knocking-out the three wheat MLO genes and regenerated powdery mildew resistant plants.

When the TALENs, ZFNs or CRISPR/Cas9 systems are engineered to inactivate one of the two *FokI* complexes responsible for the double strand break, the enzymatic complex can act rather as a nickase to producea single strand cut. Besides being a possible way for improving specificity [95], this strategy promotes repair by homologous recombination as opposed to non homologous end joining, thus allowing much more complex and accurate modifications in the target sequence (i.e. gene/allele substitutions) [98,103,104].

The capacity to induce specific mutations by means of sequence specific nucleases would allows the direct modification/introduction of relevant agronomic traits into elite lines for breeding. Nevertheless, a critical question is whether these tar-

geted DNA modifications might be regarded as genetic modified organism (GMO) events for regulation. Although this is still being debated [105–108], one common sense indication, discussed in detail by Araki and Ishii [109], suggests that genome editing events not based on the integration of a heterologous DNA (i.e. small insertions/deletions induced by non homologous end joining mechanisms), should not be regulated as GMOs. Instead, events generated through homologous recombination with a foreign template sequence, especially in the case of a whole gene, are expected to result in an organism containing recombinant DNA, and therefore they should be considered as GMOs.

When deep and accurate knowledge is available concerning metabolic pathways of primary importance in crops (including information about the molecular mechanisms responsible for and regulating these processes), all the technologies grouped under the definition of "genome editing" can give a significant contribution in the design and development of "ad hoc" metabolically optimized crop lines, thanks to synthetic biology approaches [110].

8. The control of genetic recombination

Even with all the new available technologies, plant breeding still depends on recombination. New genes/alleles are required to be recombined into advanced lines and despite the great number of markers available, recombination is still required to give new allele combinations for tightly linked loci. It is therefore essential to develop tools capable of increasing crossover incidence to break negative allele associations.

To ensure proper segregation at metaphase I, each pair of chromosomes have at least one crossover known as an obligatory crossover. Nevertheless, the presence of a crossover inhibits the formation of a new crossover nearby on the chromosome and this inhibition decreases with the distance (crossover interference) [111]. Furthermore, the crossovers are not uniformly distributed along chromosome: some regions, termed as hot spots, show much higher frequencies of crossover than cold spot areas. In species such as barley, wheat or maize, the crossovers occur much more in the distal part of chromosome than in the pericentromeric regions [112–114]. Analysis of the crossover distribution along chromosome 3B of bread wheat showed a meiotic recombination gradient from the centromere to the telomeres on both arms. The two distal regions were characterized by an elevated meiotic recombination rate (of 0.60 and 0.96 cM/Mb) which was clearly distinct from the low meiotic recombination rate (of 0.05 cM/Mb) observed in the large proximal region spanning the centromeric–pericentromeric area [115]. Different studies are underway to understand the mechanisms (and the genes) regulating the formation of crossovers to allow the control of their frequency and position [116]. For example, it was shown that a mutation of the Arabidopsis *FANCM* gene results in a substantial increase of meiotic crossover formation, without negative impacts on chromosome stability [117].

Crossover localization is partially determined by the presence of particular chromosome structures, such as tandem repeats, short terminal deletions and translocations. For example, when a chromosome carries a distal translocation, this causes a strong dissimilarity at the chromosome end resulting in a shift of crossovers to interstitial chromosome segments [118]. Other internal and external factors can impact crossover incidence. For instance, barley cultivars with different genetic background (internal factor) showed up to 30% difference in crossover frequencies, while chemical/physical treatments (external factor) with actinomycin D, diepoxybutane and/or radiations gave a large increase (from two- to seven fold) in recombination frequencies in Arabidopsis [119,120]. The authors proposed to apply a chemical/physical treatment to a F_1 hybrid that results in DNA modifications or damage. The treated F_1 hybrid is then self-pollinated or backcrossed and the resultant F_2 progenies are searched for the desired recombination events [120]. A further complication in manipulating homoeologous recombination is caused by polyploidization [121]. Indeed in polyploid species there are factors that ensure that chromosome pairing occurs only between homologous chromosomes and not homoeologous chromosomes. This diploidization mechanism also affects the ability to exploit interesting characteristics from wild species in breeding programs [122]. A well known gene affecting homoeologous recombination in wheat is *Pairing homoeologous 1* (*Ph1*), which inhibits pairing between homoeologous chromosomes [123–125]. In the absence of *Ph1*, pairing and recombination between homoeologous chromosomes is frequent, facilitating introgressive hybridization. If the constitutive deletion of *Ph1* can over time lead to rearranged chromosomes in the genome, the knock-down of these genes, for example using RNA interference (RNAi) would bring great benefits for plant breeding [126].

Although breeding is based on meiotic recombination, in specific cases there is an interest in blocking recombination to fix aheterozygous state. Many crops are cultivated as heterozygous F_1 hybrids, that carry a unique combination of alleles and outperform their parents due to hybrid vigour. The allele combinations that make F_1 unique are broken by recombination when the F_1 is selfed. The application of a new strategy, known as reverse breeding, allows genetic preservation of any selected fertile plant through seeds even if its genetic composition is unknown and if vegetative propagation is not applicable [127]. The method is based on reducing homologous recombination in the selected heterozygote by eliminating meiotic crossing-over. This is done using RNAi constructs targeting genes for proteins involved in the formation of crossovers, such as *DISRUPTED MEIOTIC CDNA1* (DMC1) [128]. The elite heterozygote is transformed using the RNAi construct and the resulting plant is expected to produce low numbers of viable balanced haploid spores. These are regenerated into doubled haploid, perfectly homozygous, plants. The resulting doubled haploids differ genetically solely as a consequence of the independent parental chromosome assortment which occurred during meiosis. Therefore, it is sufficient to make use of one co-dominant, polymorphic marker per chromosome to determine which of the lines should be combined through crossing to reconstruct the genetic composition of the original starting plant. The technique however is limited to crops in which spores can be regenerated into double haploids. In polyploids or species with high chromosome numbers, another hybrid reconstruction method based on plants regenerated from unreduced spores, named Near Reverse breeding, has been proposed [129].

A reproductive mechanism that can fix favorable allelic combinations obtained through meiotic recombination is represented by the apomixes, a process producing genetically identical offspring *via* seed without the intervention of a second parent. This process offers some advantages for fixing desirable complex genotypes, such as high yielding F_1 hybrids [130]. In crops apomixis is rare and cross pollination is not suitable for transferring the apomixis trait [131]. Nevertheless the engineering of apomictic crops, through targeted manipulations of reproduction, although challenging, could have a great value as a breeding technology [132]. Even though the genes controlling apomixis have not yet been identified and many studies are focusing on a possible epigenetic regulation, some genes identified from natural apomicts may switch the sexual pathway of crops to apomixis [130].

9. Conclusions

Current breeding programs rely on integrating phenotypic selection in standard breeding schemes (e.g. pedigree, backcross,

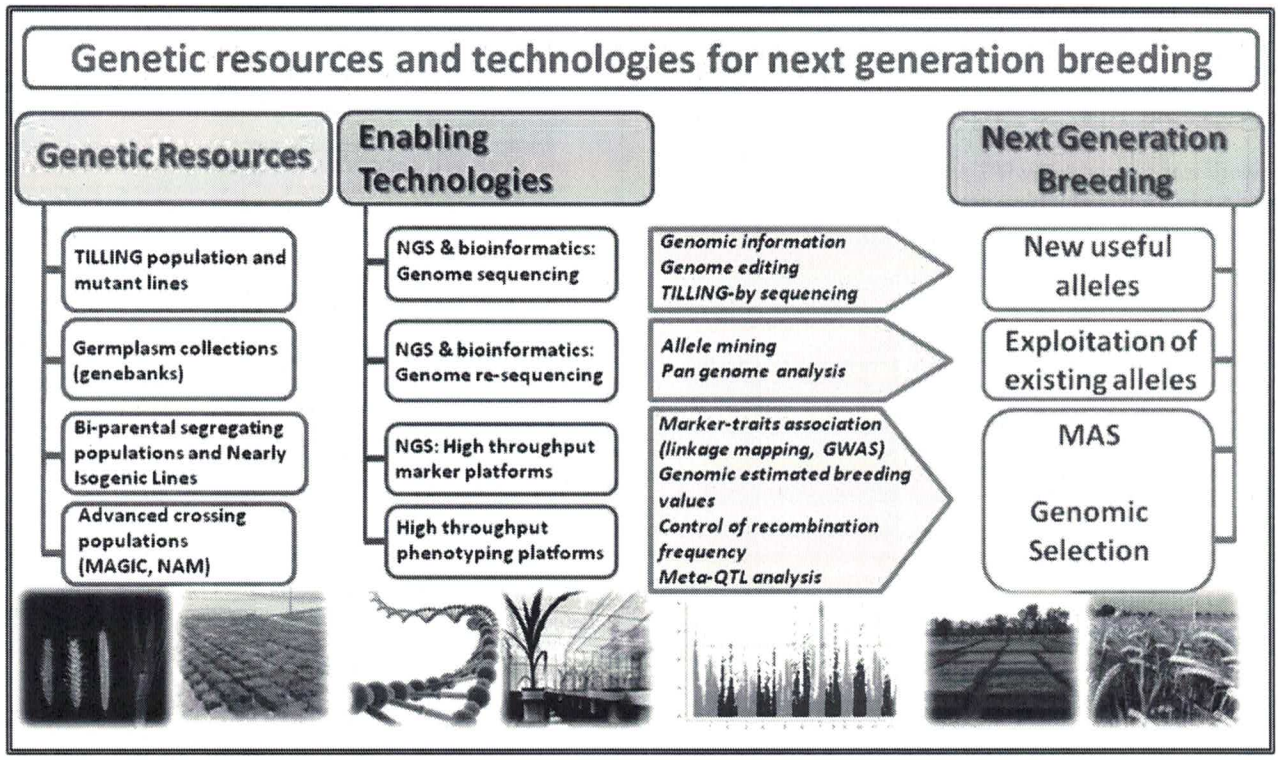

Fig. 1. Schematic representation of the key elements of a "next generation breeding". The combination of genetic resources, NGS technologies, bioinformatics capacities and automatic phenotyping facilities will revolutionize the traditional breeding strategies moving to a more efficient genetic improvement.

progeny test for combinatory efficiency) with molecular inputs (e.g. MAS and genetic transformation for GM plants). The availability of NGS, bio-informatics resources and phenotyping platforms is moving plant breeding a step forward and a next generation breeding strategies resulting from combining of genetic resources with advanced technologies can be foreseen for the near future (Fig. 1).

The first effect of the NGS revolution is to drop the marker cost per data point. SSRs or CAPS markers run on agarose gels or capillary sequencer are much more expensive than SNPs run on high-throughput platforms. As a result, while in the past only markers in critical genomic regions were employed to follow really important traits, nowadays markers are used to assess the inheritance of as many loci as possible across the entire genome and with nucleotide-level precision [5].

NGS technologies are giving a relevant contribution for the characterization of plant genetic resources. A worldwide effort is currently focused on understanding the genetic bases of agronomic traits, analyzing allelic variants at the corresponding loci and providing catalogs of allele series for the most important loci, thus enabling the breeders to select the most appropriate allele combinations. The massive accumulation of QTL information is paving the way for more accurate and powerful meta-analyses, allowing consistent genetic determinants of quantitative traits to be identified. Meanwhile, the availability of low cost markers is facilitating the introgression into elite cultivars of specific loci/QTLs from landraces or wild accessions while limiting the negative effects of linkage drag [5].

The exploitation of a genome-wide approach such as GS is becoming feasible and would help to design the new plant not only for a few selected traits but for virtually all loci in the genome, in a cost effective and relatively fast way, promising to give a remarkable contribution to the concept of next generation breeding. Application of the GS procedure would eliminate the need for extensive multi-location field trials at each generation, and would only require some phenotyping to maintain and increase the accuracy of the prediction models.

Once genes and alleles responsible for traits are identified, molecular and computational tools can be applied to potentially gain an understanding of the evolutionary processes that have shaped their current diversity in the genepool. The exploitable value of these genes/allele from unadapted germplasm for prebreeding purposes could also be determined. For example, knowledge of haplotype of favourable alleles present in elite cultivars will help to identify other (including superior) alleles from diverse landraces and wild relatives. With the help of agronomists and crop (eco)-physiologists, the optimal combinations of traits, or so-called ideotype, might be now defined by gene (network) modelling. The approach of molecular crop design has the opportunity to improve the speed and efficiency of breeding programs for several species of agricultural interest. In rice, for example, a number of genes have been identified and characterized in detail, which influence flowering time, reviewed by Matsubara et al., 2014 [133]. It has been suggested and observed that different allelic combinations of this kind of genes may influence the geographical distribution of this crop [133,134]. Therefore this information might be exploited to breed new early flowering rice varieties, in order to further facilitate the expansion northward of the cultivation region.

In disease resistance breeding, mechanisms of resistance-suppression were frequently observed in several crop species when different resistance alleles were stacked in pyramided lines, resulting in a significant limitation of the gene-pyramiding approach [135].

Although genetic resources, from elite cultivars to landraces and wild relatives, will remain the foundation of any breeding program, it is expected that in the near future the application of NGS technologies, bioinformatics and automatic phenotyping tools for the characterization and subsequent exploitation of genetic diver-

sity will revolutionize breeding strategies to achieve more efficient genetic improvement.

Acknowledgement

This work was supported by FP7 project WHEALBI (Wheat and barley Legacy for Breeding Improvement).

References

[1] T.P. Michael, S. Jackson, The first 5 plant genomes, The Plant Genome 6 (2013), http://dx.doi.org/10.3835/plantgenome2013.3.1
[2] D. Barabaschi, D. Guerra, K. Lacrima, P. Laino, V. Michelotti, S. Urso, G. Valè, L. Cattivelli, Emerging Knowledge from genome sequencing of crop species, Mol. Biotechnol. 50 (2011) 250–266.
[3] F. Fiorani, U. Schurr, Future scenarios for plant phenotyping, Ann. Rev. Plant Biol. 64 (2013) 267–291.
[4] B. McClosky, S.D. Tanksley, The impact of recombination on short-term selection gain in plant breeding experiments, Theor. Appl. Genet. 126 (2013) 2299–2312.
[5] R.K. Varshney, R. Terauchi, S.R. McCouch, Harvesting the promising fruits of genomics: applying genome sequencing technologies to crop breeding, PLoS Biol. 12 (2014) e1001883.
[6] S. Myles, J.-M. Chia, B. Hurwitz, C. Simon, G.Y. Zhong, E. Buckler, D. Ware, Rapid genomic characterization of the genus *vitis*, PLoS One 5 (2010) e8219.
[7] M.W. Ganal, G. Durstewitz, A. Polley, A. Berard, E.S. Buckler, A. Charcosset, J.D. Clarke, E.M. Graner, M. Hansen, J. Joets, M.C. Le Paslier, M.D. McMullen, P. Montalent, M. Rose, C.C. Schön, Q. Sun, H. Walter, O.C. Martin, M. Falque, A large maize (*Zea mays* L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome, PLoS One 6 (2011) e28334.
[8] S.-C. Sim, G. Durstewitz, J. Plieske, R. Wieseke, M.W. Ganal, A. Van Deynze, J.P. Hamilton, C.R. Buell, M. Causse, S. Wijeratne, D.M. Francis, Development of a large SNP genotyping array and generation of high-density genetic maps in tomato, PLoS One 7 (2012) e40563.
[9] I. Verde, N. Bassil, S. Scalabrin, B. Gilmore, C.T. Lawley, K. Gasic, D. Micheletti, U.R. Rosyara, F. Cattonaro, E. Vendramin, D. Main, V. Aramini, A.L. Blas, T.C. Mockler, D.W. Bryant, L. Wilhelm, M. Troggio, B. Sosinski, M.J. Aranzana, P. Arús, A. Iezzoni, M. Morgante, C. Peace, Development and evaluation of a 9K SNP array for peach by internationally coordinated SNP detection and validation in breeding germplasm, PLoS One 7 (2012) e35668.
[10] Q. Song, D.L. Hyten, G. Jia, C.V. Quigley, E.W. Fickus, R.L. Nelson, P.B. Cregan, Development and evaluation of SoySNP50K, a high-density genotyping array for soybean, PLoS One 8 (2013) e54985.
[11] J. Comadran, B. Kilian, J. Russell, L. Ramsay, N. Stein, M. Ganal, P. Shaw, M. Bayer, W. Thomas, D. Marshall, P. Hedley, A. Tondelli, N. Pecchioni, E. Francia, V. Korzun, A. Walther, R. Waugh, Natural variation in a homolog of *Antirrhinum* CENTRORADIALIS contributed to spring growth habit and environmental adaptation in cultivated barley, Nat. Genet. 44 (2012) 1388–1392.
[12] H. Chen, W. Xie, H. He, H. Yu, W. Chen, J. Li, R. Yu, Y. Yao, W. Zhang, Y. He, X. Tang, F. Zhou, X. Wang Deng, Q. Zhang, A high-density SNP genotyping array for rice biology and molecular breeding, Mol. Plant 7 (2014) 541–553.
[13] S. Wang, D. Wong, K. Forrest, A. Allen, S. Chao, B.E. Huang, M. Maccaferri, S. Salvi, S.G. Milner, L. Cattivelli, A.M. Mastrangelo, A. Whan, S. Stephen, G. Barker, R. Wieseke, J. Plieske, International Wheat Genome Sequencing Consortium, M. Lillemo, D. Mather, R. Appels, R. Dolferus, G. Brown-Guedira, A. Korol, A.R. Akhunova, C. Feuillet, J. Salse, M. Morgante, C. Pozniak, M.-C. Luo, J. Dvorak, M. Morell, J. Dubcovsky, M. Ganal, R. Tuberosa, C. Lawley, I. Mikoulitch, C. Cavanagh, K.J. Edwards, M. Hayden, E. Akhunov, Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array, Plant Biotechnol. J. 12 (2014) 787–796.
[14] L. Bianco, A. Cestaro, D.J. Sargent, E. Banchi, S. Derdak, M. Di Guardo, S. Salvi, J. Jansen, R. Viola, I. Gut, F. Laurens, D. Chagné, R. Velasco, E. van de Weg, M. Troggio, Development and validation of a 20K single nucleotide polymorphism (SNP) whole genome genotyping array for apple (*Malus* × *domestica* Borkh), PloS One 9 (2014) e110377.
[15] J.P. Hamilton, C.R. Buell, Advances in plant genome sequencing, Plant J. 70 (2012) 177–190.
[16] M.R. Miller, J.P. Dunham, A. Amores, W.A. Cresko, E.A. Johnson, Rapid and cost effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers, Genome Res. 17 (2007) 240–248.
[17] R.J. Elshire, J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler, S.E. Mitchell, A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species, PLoS One 6 (2011) e19379.
[18] H. Liu, M. Bayer, A. Druka, J.R. Russell, C.A. Hackett, J. Poland, L. Ramsay, P.E. Hedley, R. Waugh, An evaluation of genotyping by sequencing (GBS) to map the *Breviaristatum-e* (*ari-e*) locus in cultivated barley, BMC Genomics 15 (2014) 104.
[19] E. Portis, L. Barchi, L. Toppino, S. Lanteri, N. Acciarri, N. Felicioni, F. Fusari, V. Barbierato, F. Cericola, G. Vale', G.L. Rotino, QTL mapping in eggplant reveals clusters of yield-related loci and orthology with the tomato genome, PLoS One 9 (2014).
[20] K. Wu, H. Liu, M. Yang, Y. Tao, H. Ma, W. Wu, Y. Zuo, Y. Zhao, High-density genetic map construction and QTLs analysis of grain yield-related traits in Sesame (*Sesamum indicum* L.) based on RAD-Seq techonology, BMC Plant Biol. 14 (2014) 274.
[21] J. Spindel, H. Begum, D. Akdemir, P. Virk, B. Collard, E. Redoña, G. Atlin, J.-L. Jannink, S.R. McCouch, Genomic selection and association mapping in rice (*Oryza sativa*): effect of trait genetic architecture, training population composition, marker number and statistical model on accuracy of rice genomic selection in elite, tropical rice breeding lines, PLoS Genet. 11 (2015) e1004982.
[22] X. Zhang, P. Pérez-Rodríguez, K. Semagn, Y. Beyene, R. Babu, M.A. López-Cruz, F. San Vicente, M. Olsen, E. Buckler, J.-L. Jannink, B.M. Prasanna, J. Crossa, Genomic prediction in biparental tropical maize populations in water-stressed and well-watered environments using low-density and GBS SNPs, Heredity 114 (2014) 291–299.
[23] M. Mascher, T.A. Richmond, D.J. Gerhardt, A. Himmelbach, L. Clissold, D. Sampath, S. Ayling, B. Steuernagel, M. Pfeifer, M. D'Ascenzo, E.D. Akhunov, P.E. Hedley, A.M. Gonzales, P.L. Morrell, B. Kilian, F.R. Blattner, U. Scholz, K.F. Mayer, A.J. Flavell, G.J. Muehlbauer, R. Waugh, J.A. Jeddeloh, N. Stein, Barley whole exome capture: a tool for genomic research in the genus *Hordeum* and beyond, Plant J. 76 (2013) 494–505.
[24] J.A. Chapman, M. Mascher, A. Buluç, K. Barry, E. Georganas, A. Session, V. Strnadova, Je. Jenkins, S. Sehgal, L. Oliker, J. Schmutz, K.A. Yelick, U. Scholz, R. Waugh, J.A. Poland, G.J. Muehlbauer, N. Stein, D.S. Rokhsar, A whole-genome shotgun approach for assembling and anchoring the hexaploid bread wheat genome, Genome Biol. 16 (2015), 26.
[25] J. Salse, In silico archeogenomics unveils modern plant genome organization, regulation and evolution, Curr. Opin. Plant Biol. 15 (2012) 122–130.
[26] M.E. Bolger, B. Weisshaar, U. Scholz, N. Stein, B. Usadel, K.F.X. Mayer, Plant genome sequencing – applications for crop improvement, Curr. Opin. Biotechnol. 26 (2014) 31–37.
[27] X. Huang, N. Kurata, X. Wei, Z.X. Wang, A. Wang, Q. Zhao, Y. Zhao, K. Liu, H. Lu, W. Li, Y. Guo, Y. Lu, C. Zhou, D. Fan, Q. Weng, C. Zhu, T. Huang, L. Zhang, Y. Wang, L. Feng, H. Furuumi, T. Kubo, T. Miyabayashi, X. Yuan, Q. Xu, G. Dong, Q. Zhan, C. Li, A. Fujiyama, A. Toyoda, T. Lu, Q. Feng, Q. Qian, J. Li, B. Ha, A map of rice genome variation reveals the origin of cultivated rice, Nature 490 (2012) 497–501.
[28] X. Xu, X. Liu, S. Ge, J.D. Jensen, F. Hu, X. Li, Y. Dong, R.N. Gutenkunst, L. Fang, L. Huang, J. Li, W. He, G. Zhang, X. Zheng, F. Zhang, Y. Li, C. Yu, K. Kristiansen, X. Zhang, J. Wang, M. Wright, S. McCouch, R. Nielsen, J. Wang, W. Wang, Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes, Nat. Biotechnol. 30 (2012) 105–111.
[29] P.S. Schnable, N.M. Springer, Progress toward understanding heterosis in crop plants, Annu. Rev. Plant Biol. 64 (2013) 71–88.
[30] J. J-Y. Li, R.S. Zeigler Wang, The 3000 rice genomes project: new opportunities and challenges for future rice research, GigaScience 3 (2014) 8.
[31] M. Wang, S. Allefs, R.G. van den Berg, V.G. Vleeshouwers, E.A. van der Vossen, B. Vosman, Allele mining in *Solanum*: conserved homologues of *Rpi-blb1* are identified in *Solanum stoloniferum*, Theor. Appl. Genet. 116 (2008) 933–943.
[32] N.K. Bhullar, Z. Zhang, T. Wicker, B. Keller, Wheat gene bank accessions as a source of new alleles of the powdery mildew resistance gene Pm3: a large scale allele mining project, BMC Plant Biol. 10 (2010) 88.
[33] G. Ramkumar, K. Srinivasarao, K.M. Mohan, I. Sudarshan, A. Sivaranjani, K. Gopalakrishna, C. Neeraja, S. Balachandran, R. Sundaram, M. Prasad, Development and validation of functional marker targeting an InDel in the major rice blast disease resistance gene *Pi54* (*Pik h*), Mol. Breed. 27 (2011) 129–135.
[34] D. Wang, C. Guo, J. Huang, S. Yang, D. Tian, X. Zhang, Allele-mining of rice blast resistance genes at AC134922 locus, Biochem. Biophys. Res. Commun. 446 (2014) 1085–1090.
[35] C. Biselli, S. Urso, G. Tacconi, B. Steuernagel, D. Schulte, A. Gianinetti, P. Bagnaresi, N. Stein, L. Cattivelli, G. Valè, Haplotype variability and identification of new functional alleles at the Rdg2a leaf stripe resistance gene locus, Theor. Appl. Genet. 126 (2013) 1575–1586.
[36] H. Tsai, T. Howell, R. Nitcher, V. Missirian, B. Watson, K.J. Ngo, M. Lieberman, J. Fass, C. Uauy, R.K. Tran, A.A. Khan, V. Filkov, T.H. Tai, J. Dubcovsky, L. Comai, Discovery of rare mutations in populations: TILLING by sequencing, Plant Physiol. 156 (2011) 1257–1268.
[37] S.-I. Kim, H.T. Tai, Identification of novel rice low phytic acid mutations via TILLING by sequencing, Mol. Breed. 34 (2014) 1717–1729.
[38] I.M. Henry, U. Nagalakshmi, M.C. Lieberman, K.J. Ngo, K.V. Krasileva, H. Vasquez-Gross, A. Akhunova, E. Akhunov, J. Dubcovsky, T.H. Tai, L. Comai, Efficient genome-wide detection and cataloging of EMS-induced mutations using exome capture and next-generation sequencing, Plant Cell 26 (2014) 1382–1397.
[39] M. Morgante, E. De Paoli, S. Radovic, Transposable elements and the plant pan-genomes, Curr. Opin. Plant Biol. 10 (2007) 149–155.
[40] Y.H. Li, G. Zhou, J.W. Jiang, L.G. Jin, Z. Zhang, Y. Guo, J. Zhang, Y. Sui, L. Zheng, S.S. Zhang, Q. Zuo, X.H. Shi, Y.F. Li, W.K. Zhang, Y. Hu, G. Kong, H.L. Hong, B. Tan, J. Song, Z.X. Liu, Y. Wang, H. Ruan, C.K. Yeung, J. Liu, H. Wang, L.J. Zhang, R.X. Guan, K.J. Wang, W.B. Li, S.Y. Chen, R.Z. Chang, Z. Jiang, S.A. Jackson, R. Li, L.J. Qiu, De novo assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits, Nat. Biotechnol. 32 (2014) 1045–1052.

[41] C.N. Hirsch, J.M. Foerster, J.M. Johnson, R.S. Sekhon, G. Muttoni, B. Vaillancourta, F. Peñagaricano, E. Lindquist, M.A. Pedraza, K. Barry, N. de Leon, S.M. Kaeppler, C.R. Buell, Insights into the maize pan-genome and pan-transcriptome, Plant Cell 26 (2014) 121–135.
[42] M. Moshelion, A. Altman, Current challenges and future perspectives of plant and agricultural biotechnology, Trends Biotechnol. 33 (2015) 337–342.
[43] S. Tisné, Y. Serrand, L. Bach, E. Gilbault, R. Ben Ameur, H. Balasse, R. Voisin, D. Bouchez, M. Durand-Tardif, P. Guerche, G. Chareyron, J. Da Rugna, C. Camilleri, O. Loudet, Phenoscope: an automated large-scale phenotyping platform offering high spatial homogeneity, Plant J. 74 (2013) 534–544.
[44] R.T. Clark, R.B. MacCurdy, J.K. Jung, J.E. Shaff, S.R. McCouch, D.J. Aneshansley, L.V. Kochian, Three-dimensional root phenotyping with a novel imaging and software platform, Plant Physiol. 156 (2011) 455–465.
[45] A. Hartmann, T. Czauderna, R. Hoffmann, N. Stein, F.S. Hartmann, HTPheno An image analysis pipeline for high-throughput plant phenotyping, BMC Bioinf. 12 (2011) 148.
[46] M.P. Pound, A.P. French, J.A. Atkinson, D.M. Wells, M.J. Bennett, T. Pridmore, RootNav: navigating images of complex root architectures, Plant Physiol. 162 (2013) 1802–1814.
[47] C. Klukas, D. Chen, J.-M. Pape, Integrated analysis platform: an open-source information system for high-throughput plant phenotyping, Plant Physiol. 165 (2014) 506–518.
[48] D. Chen, K. Neumann, S. Friedel, B. Kilian, M. Chen, T. Altmann, C. Klukasa, Dissecting the phenotypic components of crop plant growth and drought responses based on high-throughput image analysis, Plant Cell 26 (2014) 4636–4655.
[49] A. Petrozza, A. Santaniello, S. Summerer, G. Di Tommaso, D. Di Tommaso, E. Paparelli, A. Piaggesi, P. Perata, F. Cellini, Physiological responses to Megafol® treatments in tomato plants under drought stress: a phenomic and molecular approach, Sci. Hortic. 174 (2014) 185–192.
[50] J.L. Araus, J.E. Cairns, Field high-throughput phenotyping: the new crop breeding frontier, Trends Plant Sci. 1 (2014) 52–61.
[51] F. Liebisch, N. Kirchgessner, D. Schneider, A. Walter, A. Hund, Remote, aerial phenotyping of maize traits with a mobile multi-sensor approach, Plant Methods 11 (2015) 9.
[52] A. Philibert, C. Loyce, D. Makowski, Assessment of the quality of meta-analysis in agronomy, Agric. Ecosyst. Environ. 148 (2012) 72–82.
[53] X.L. Wu, Z.L. Hu, Meta-analysis of QTL mapping experiments, Methods Mol. Biol. 871 (2012) 145–171.
[54] E. Ballini, J.-B. Morel, G. Droc, A. Price, B. Courtois, J.-L. Notteghem, D. Tharreau, A genome-wide meta-analysis of rice blast resistance genes and quantitative trait loci provides new insights into partial and complete resistance, Mol. Plant Microbe Interact. (2008) 859–868, 7.
[55] S. Griffiths, J. Simmonds, M. Leverington, Y.K. Wang, L. Fish, L. Sayers, L. Alibert, S. Orford, L. Wingen, J. Snape, Meta-QTL analysis of the genetic control of crop height in elite European winter wheat germplasm, Mol. Breed 29 (2010) 159–171.
[56] K. Semagn, Y. Beyene, M. Warburton, A. Tarekegne, S. Mugo, B. Meisel, P. Sehabiague, B.M. Prasanna, Meta-analyses of QTL for grain yield and anthesis silking interval in 18 maize populations evaluated under water stressed and well-watered environments, BMC Genomics 14 (2013) 313.
[57] D. Marone, M.A. Russo, G. Laidò, P. De Vita, R. Papa, A. Blanco, A. Gadaleta, D. Rubiales, A.M. Mastrangelo, Genetic basis of qualitative and quantitative resistance to powdery mildew in wheat: from consensus regions to candidate genes, BMC Genomics 14 (2013) 562.
[58] A. Lehmensiek, W. Bovill, P. Wenzl, P. Langridge, R. Appels, Genetic mapping in the triticeae, in: C. Feuillet, G.J. Muehlbauer (Eds.), Genetics and Genomics of the Triticeae, Springer, Heidelberg, 2009, pp. 201–236.
[59] A. Rafalski, Applications of single nucleotide polymorphisms in crop genetics, Curr. Opin. Plant Biol. 5 (2002) 94–100.
[60] X. Huang, B. Han, Natural variations and genome-wide association studies in crop plants, Annu. Rev. Plant Biol. 65 (2014) 531–551.
[61] A. Korte, A. Farlow, The advantages and limitations of trait analysis with GWAS: a review, Plant Methods 9 (2013) 29.
[62] T. Ogura, W. Busch, From phenotypes to causal sequences: using genome wide association studies to dissect the sequence basis for variation of plant development, Curr. Opin. Plant Biol. 23 (2015) 98–108.
[63] M.L Tenaillon, M.C. Sawkins, A.D. Long, R.L. Gaut, J.F. Doebley, B.S. Gaut, Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. *mays* L.), Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 9161–9166.
[64] K.S. Caldwell, J. Russell, P. Langridge, W. Powell, Extreme population dependent linkage disequilibrium detected in an inbreeding plant species, Hordeum vulgare, Genetics 172 (2006) 557–567.
[65] B. Courtois, A. Audebert, A. Dardou, S. Roques, T. Ghneim-Herrera, G. Droc, J. Frouin, L. Rouan, E. Gozé, A. Kilian, N. Ahmadi, M. Dingkuhn, Genome-wide association mapping of root traits in a Japonica rice panel, PLoS One 8 (2013) e78037.
[66] M.T. Hamblin, E.S. Buckler, J. Jannink, Population genetics of genomics-based crop improvement methods, Trends Genet. 27 (2011) 98–106.
[67] L. Ramsay, J. Comadran, A. Druka, D.F. Marshall, W.T.B. Thomas, M. Macaulay, K. MacKenzie, C. Simpson, J. Fuller, N. Bonar, P.M. Hayes, U. Lundqvist, J.D. Franckowiak, T.J. Close, G.J. Muehlbauer, R. Waugh, INTERMEDIUM-C, a modifier of lateral spikelet fertility in barley, is an ortholog of the maize domestication gene TEOSINTE BRANCHED 1, Nat. Genet. 43 (2011) 169–172.
[68] Q. Yang, Z. Li, W. Li, L. Ku, C. Wang, J. Ye, K. Li, N. Yang, Y. Li, T. Zhong, J. Li, Y. Chen, Ji. Yan, X. Yang, M. Xu, CACTA-like transposable element in ZmCCT attenuated photoperiod sensitivity and accelerated the postdomestication spread of maize, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 16969–16974.
[69] W. Chen, Y. Gao, W. Xie, L. Gong, K. Lu, W. Wang, Y. Li, X. Liu, H. Zhang, H. Dong, W. Zhang, L. Zhang, S. Yu, G. Wang, X. Lian, J. Luo, Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism, Nat. Genet. 46 (2014) 714–721.
[70] P.K. Gupta, P.L. Kulwal, R.R. Mir, QTL mapping: methodology and applications in cereal breeding, in: P.K. Gupta, R.K. Varshney (Eds.), Cereal Genomics II, Springer, Heidelberg, 2013, pp. 275–318.
[71] N. Bandillo, C. Raghavan, P.A. Muyco, M.A. Sevilla, I.T. Lobina, C.J. Dilla-Ermita, C.W. Tung, S. McCouch, M. Thomson, R. Mauleon, R.K. Singh, G. Gregorio, H. Leung, Multi-parent advanced generation inter-cross (MAGIC) populations in rice: progress and potential for genetics research and breeding, Rice 6 (2013) 11.
[72] I.J. Mackay, P. Bansept-Basler, T. Barber, A.R. Bentley, J. Cockram, N. Gosman, A.J. Greenland, R. Horsnell, R. Howells, D.M. O'sullivan, G.A. Rose, P.J. Howell, An eight-parent multiparent advanced generation inter-cross population for winter-sown wheat: creation, properties, and validation, G3 4 (2014) 1603–1610.
[73] E.S. Buckler, J.B. Holland, P.J. Bradbury, C.B. Acharya, P.J. Brown, C. Browne, E. Ersoz, S. Flint-Garcia, A. Arturo Garcia, J.C. Glaubitz, M.M. Goodman, C. Harjes, K. Guill, D.E. Kroon, S. Larsson, N.K. Lepak, H. Li, S.E. Mitchell, G. Pressoir, J.A. Peiffer, T.R. Rocheford, M.C. Romay, S. Romero, S. Salvo, H. Sanchez Villeda, H.S. da Silva, Q. Sun, F. Feng Tian, N. Upadyayula, D. Ware, H. Yates, J. Yu, Z. Zhang, S. Kresovich, M.D. McMullen, The genetic architecture of maize flowering time, Science 325 (2009) 714–718.
[74] J.M. Foerster, T. Beissinger, N. de Leon, S. Kaeppler, Large effect QTL explain natural phenotypic variation for the developmental timing of vegetative phase change in maize (*Zea mays* L.), Theor. Appl. Genet. 128 (2015) 529–538.
[75] J.P. Cook, M.D. McMullen, J.B. Holland, F. Tian, P. Bradbury, J. Ross-Ibarra, E.S. Buckler, S.A. Flint-Garcia, Genetic architecture of maize kernel composition in the nested association mapping and inbred association panels, Plant Physiol. 158 (2011) 824–834.
[76] F. Schnaithmann, D. Kopahnke, K. Pillen, A first step toward the development of a barley NAM population and its utilization to detect QTLs conferring leaf rust seedling resistance, Theor. Appl. Genet. 127 (2014) 1513–1525.
[77] E. Francia, G. Tacconi, C. Crosatti, D. Bulgarelli, D. Barabaschi, E. Dall'Aglio, G. Valè, Marker-assisted selection in crop plants, Plant Cell Tissue Organ Cult. 82 (2005) 317–342.
[78] Y. Xu, J.H. Crouch, Marker-assisted selection in plant breeding: from publications to practice, Crop Sci. 48 (2008) 391–407.
[79] E.L. Heffner, M.E. Sorrells, J.-L. Jannink, Genomic selection for crop improvement, Crop Sci. 49 (2009) 1–12.
[80] **A.A. Nakaya, S.N. Isobe, Will genomic selection be a practical method for plant breeding, Ann. Bot. 110 (2012) 1303–1316,** http://dx.doi.org/10.1093/aob/mcs109
[81] Z.A. Desta, R. Ortiz, Genomic selection: genome-wide prediction in plant improvement, Trends Plant Sci. 19 (2014) 592–601.
[82] A. Onogi, O. Ideta, Y. Inoshita, K. Ebana, T. Yoshioka, M. Yamasaki, H. Iwata, Exploring the areas of applicability of whole-genome prediction methods for Asian rice (Oryza sativa L.), Theor. Appl. Genet. 128 (2014) 41–53.
[83] M. Gouy, Y. Rousselle, D. Bastianelli, P. Lecomte, L. Bonnal, D. Roques, J.C. Efile, S. Rocher, J. Daugrois, L. Toubi, S. Nabeneza, C. Hervouet, H. Telismart, M. Denis, A. Thong-Chane, J.C. Glaszmann, J.Y. Hoarau, S. Nibouche, L. Costet, Experimental assessment of the accuracy of genomic selection in sugarcane, Theor. Appl. Genet. 126 (2013) 2575–2586.
[84] T. Guo, H. Li, J. Yan, J. Tang, J. Li, Z. Zhang, L. Zhang, J. Wang, Performance prediction of F1 hybrids between recombinant inbred lines derived from two elite maize inbred lines, Theor. Appl. Genet. 126 (2012) 189–201.
[85] T.H.E. Meuwissen, B.J. Hayes, M.E. Goddard, Prediction of total genetic value using genome-wide dense marker maps, Genetics 157 (2001) 1819–1829.
[86] M.T. Hamblin, M.G. Salas Fernandez, A.M. Casa, S.E. Mitchell, A.H. Paterson, S. Kresovich, Equilibrium processes cannot explain high levels of short-and medium-range linkage disequilibrium in the domesticated grass Sorghum bicolor, Genetics 171 (2005) 1247–1256.
[87] R.K. Pasam, R. Sharma, M. Malosetti, F.A. van Eeuwijk, G. Haseneyer, B. Kilian, A. Graner, Genome-wide association studies for agronomical traits in a word wide spring barley collection, BMC Plant Biol. 12 (2012) 16.
[88] **E.L. Heffner, J.-L. Jannink, H. Iwata, E. Souza, M.E. Sorrells, Genomic selection accuracy for grain quality traits in biparental wheat populations, Crop Sci. 51 (2011) 2597–2606,** http://dx.doi.org/10.2135/cropsci2011.05.0253
[89] Y. Li, Y. Li, S. Wu, K. Han, Z. Wang, W. Hou, Y. Zeng, R. Wu, Estimation of multilocus linkage disequilibria in diploid populations with dominant markers, Genetics 176 (2007) 1811–1821.
[90] R. Bernardo, J. Yu, Prospects for genome-wide selection for quantitative traits in maize, Crop Sci. 47 (2007) 1082–1090.
[91] **J. Rutkoski, J. Benson, Y. Jia, G. Brown-Guedira, J.-L. Jannink, M. Sorrells, Evaluation of genomic prediction methods for Fusarium Head Blight resistance in wheat, Plant Genome 5 (2012),** http://dx.doi.org/10.3835/plantgenome2012.02.0001
[92] W.M. Ainley, L. Sastry-Dent, M.E. Welter, M.G. Murray, B. Zeitler, R. Amora, D.R. Corbin, R.R. Miles, N.L. Arnold, T.L. Strange, M.A. Simpson, Z. Cao, C. Carroll, K.S. Pawelczak, R. Blue, K. West, L.M. Rowland, D. Perkins, P. Samuel, C.M. Dewes, L. Shen, S. Sriram, S.L. Evans, E.J. Rebar, L. Zhang, P.D. Gregory,

F.D. Urnov, S.R. Webb, J.F. Petolino, Trait stacking via targeted genome editing, Plant Biotechnol. J. 11 (2013) 1126–1134.
[93] K. Belhaj, A. Chaparro-Garcia, S. Kamoun, V. Nekrasov, Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system, Plant Methods 9 (2013) 39.
[94] A. Knoll, F. Fauser, H. Puchta, DNA recombination in somatic plant cells: mechanisms and evolutionary consequences, Chromosome Res. 22 (2014) 191–201.
[95] L. Bortesi, R. Fischer, The CRISPR/Cas9 system for plant genome editing and beyond, Biotechnol. Adv. 33 (1) (2015) 41–52.
[96] Y. Wang, X. Cheng, Q. Shan, Y. Zhang, J. Liu, C. Gao, J.-L. Qiu, Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew, Nat. Biotechnol. 32 (2014) 947–951.
[97] J.A. Townsend, D.A. Wright, R.J. Winfrey, F. Fu, M.L. Maeder, J.K. Joung, D.F. Voytas, High-frequency modification of plant genes using engineered zinc-finger nucleases, Nature 459 (2009) 442–445.
[98] T. Sprink, J. Metje, F. Hartung, Plant genome editing by novel tools: TALEN and other sequence specific nucleases, Curr. Opin. Biotechnol. 32 (2015) 47–53.
[99] Q. Shan, Y. Wang, J. Li, Y. Zhang, K. Chen, Z. Liang, K. Zhang, J. Liu, J.J. Xi, J.-L. Qiu, C. Gao, Targeted genome modification of crop plants using a CRISPR-Cas system, Nat. Biotechnol. 7 (2013) 686–688.
[100] V.S. Lor, C.G. Starker, D.F. Voytas, D. Weiss, N.E. Olszewski, Targeted mutagenesis of the tomato PROCERA gene using transcription activator-like effector nucleases, Plant Physiol. 166 (2014) 1288–1291.
[101] S.K. Upadhyay, J. Kumar, A. Alok, R. Tuli, RNA-guided genome editing for target gene mutations in wheat, G3 3 (2013) 2233–2238.
[102] P. Piffanelli, L. Ramsay, R. Waugh, A. Benabdelmouna, A. D'Hont, K. Hollricher, J.H. Jørgensen, P. Schulze-Lefert, R. Panstruga, A barley cultivation-associated polymorphism conveys resistance to powdery mildew, Nature 430 (2004) 887–891.
[103] J.P. Guilinger, D.B. Thompson, D.R. Liu, Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification, Nat Biotechnol. 32 (2014) 577–582.
[104] S.Q. Tsai, N. Wyvekens, C. Khayter, J.A. Foden, V. Thapar, D. Reyon, M.J. Goodwin, M.J. Aryee, J.K. Joung, Dimeric CRISPR RNA guided FokI nucleases for highly specific genome editing, Nat. Biotechnol. 32 (32) (2014) 569–576.
[105] **EFSA Panel on Genetically Modified Organisms: Scientific opinion addressing the safety assessment of plants developed using zinc finger nuclease 3 and other site-directed nucleases with similar function, EFSA J., 10, (2012), 2943.**
[106] M. Lusser, H.V. Davies, Comparative regulatory approaches for groups of new plant breeding techniques, New Biotechnol. 30 (2013) 437–446.
[107] K. Pauwels, N. Podevin, D. Breyer, D. Carroll, P. Herman, Engineering nucleases for gene targeting: safety and regulatory considerations, New Biotechnol. 31 (2014) 18–27.
[108] F. Hartung, J. Schiemann, Precise plant breeding using new genome editing techniques: opportunities, safety and regulation in the EU, Plant J. 78 (2014) 742–752.
[109] M. Araki, T. Ishii, Towards social acceptance of plant breeding by genome editing, Trends Plant Sci. 20 (2015) 145–149.
[110] R. Kelwick, J.T. MacDonald, A.J. Webb, P. Freemont, Developments in the tools and methodologies of synthetic biology, Front. Bioeng. Biotechnol. 2 (2014), 60.
[111] S. Wang, D. Zickler, N. Kleckner, L. Liangran Zhang, Meiotic crossover patterns: obligatory crossover, interference and homeostasis in a single process, Cell Cycle 14 (2015) 305–314.
[112] **C. Mezard, Meiotic recombination hotspots in plants, Biochem. Soc. Trans. 34 (2006) 531–534,** http://dx.doi.org/10.1042/BST0340531
[113] C. Mezard, J. Vignard, J. Drouaud, R. Mercier, The road to crossovers: plants have their say, Trends Genet. 23 (2007) 91–99.
[114] L.K. Anderson, S.M. Stack, Recombination nodules in plants, Cytogenet. Genome Res. 109 (2005) 198–204.
[115] F. Choulet, A. Alberti, S. Theil, N. Glover, V. Barbe, J. Daron, L. Pingault, P. Sourdille, A. Couloux, E. Paux, P. Leroy, S. Mangenot, N. Guilhot, J. Le Gouis, F. Balfourier, M. Alaux, V. Jamilloux, J. Poulain, C. Durand, A. Bellec, C. Gaspin, J. Safar, J. Dolezel, J. Rogers, K. Vandepoele, J.M. Aury, K. Mayer, H. Berges, H. Quesneville, P. Wincker, C. Feuillet, Structural and functional partitioning of bread wheat chromosome 3B, Science 345 (2014) 1249721.
[116] J.D. Higgins, K. Osman, G.H. Jones, F.C.H. Franklin, Factors underlying restricted crossover localization in barley meiosis, Ann. Rev. Genet. 48 (2014) 29–47.
[117] A. Knoll, J.D. Higgins, K. Seeliger, S.J. Reha, N.J. Dangel, M. Bauknecht, S. Schröpfer, F.C.H. Franklin, H. Puchtaa, The Fanconi anemia ortholog FANCM ensures ordered homologous recombination in both somatic and meiotic cells in *Arabidopsis*, Plant Cell 24 (2012) 1448–1464.
[118] E. Wijnker, H. de Jong, Managing meiotic recombination in plant breeding, Trends Plant Sci. 13 (2008) 640–646.
[119] R.P. Sinha, S.B. Helgason, The action of actinomycin D and diepoxybutane on recombination of two closely linked loci in *Hordeum*, Can. J. Genet. Cytol. 11 (1969) 745–751.
[120] **G. Copenhaver, D. Preuss, Chemical and physical treatment that stimulate recombination (2000) WO Patent App. PCT/US2000/006,994.**
[121] R.T. Gaeta, J. Chris Pires, Homoeologous recombination in allopolyploids: the polyploid ratchet, New Phytol. 186 (2009) 18–28.
[122] G. Moore, Early stages of meiosis in wheat and the role of *Ph1*, in: C. Feuillet, G.J. Muehlbauer (Eds.), Genetics and Genomics of the Triticeae, Springer, Heidelberg, 2009, pp. 237–252.
[123] E.R. Sears, M. Okamoto, Intergenomic chromosome relationships in hexaploid wheat, in: Proceedings of the Tenth International Congress of Genetics, Univ. of Toronto Press, Toronto, 1958, pp. 258–259.
[124] S. Griffiths, R. Sharp, T.N. Foote, I. Bertin, M. Wanous, S. Reader, I. Colas, G. Moore, Molecular characterization of Ph1 as a major chromosome pairing locus in polyploid wheat, Nature 439 (2006) 749–752.
[125] N. Al-Kaff, E. Knight, I. Bertin, T. Foote, N. Hart, S. Griffiths, G. Moore, Detailed dissection of the chromosomal region containing the Ph1 locus in wheat *Triticum aestivum*: With deletion mutants and expression profiling, Ann. Bot. 101 (2008) 863–872.
[126] R. Bhullar, R. Nagarajan, H. Bennypaul, G.K. Sidhu, G. Sidhu, S. Rustgi, D. von Wettstein, K.S. Gill, Silencing of a metaphase I-specific gene results in a phenotype similar to that of the *pairing homeologous 1* (*Ph1*) gene mutations, Proc. Natl. Acad. Sci. U. S. A. 111 (2014) 14187–14192.
[127] R. Dirks, K. van Dun, C.B. de Snoo, M. van den Berg, C.L.C. Lelivelt, W. Voermans, L. Woudenberg, J.P.C. de Wit, K. Reinink, J.W. Shut, E. van der Zeeuw, A. Vogelaar, G. Freymark, E.W. Gutteling, M.N. Keppel, P. van Drongelen, M. Kieny, P. Ellul, A. Touraev, H. Ma, H. de Jong, E. Wijnker, Reverse breeding: a novel breeding approach based on engineered meiosis, Plant Biotechnol. J. 7 (2009) 837–845.
[128] N. Siaud, E. Dray, I. Gy, E. Gérard, N. Takvorian, M.P. Doutriaux, Brca2 is involved in meiosis in *Arabidopsis thaliana* as suggested by its interaction with Dmc1, EMBO J. 23 (2004) 1392–1401.
[129] **K. Van Dun, R. Dirks, Near reverse breeding. Patent WO 2 006 0 94773 (2006).**
[130] M.L. Hand, A.M.G. Koltunow, The genetic control of apomixis: asexual seed formation, Genetics 197 (2014) 441–450.
[131] Y. Savidan, Apomixis: Genetics and Breeding, in: J. Janick (Ed.), Plant Breeding Reviews, Wiley, New York, 2000, pp. 13–86.
[132] A. Schmidt, M.W. Schmid, U. Grossniklaus, Plant germline formation: common concepts and developmental flexibility in sexual and asexual reproduction, Development 142 (2015) 229–241.
[133] K. Matsubara, K. Hori, E. Ogiso-Tanaka, M. Yano, Cloning of quantitative trait genes from rice reveals conservation and divergence of photoperiod flowering pathways in Arabidopsis and rice, Front. Plant Sci. 5 (2014), 193.
[134] **X. Li, H. Liu, M. Wang, H. Liu, X. Tian, W. Zhou, T. Lü, Z. Wang, C. Chu, J. Fang, Q. Bu, Combinations of Hd2 and Hd4 genes determine rice adaptability to Heilongjiang Province, northern limit of China, J. Integr. Plant Biol. (2015),** http://dx.doi.org/10.1111/jipb.12326
[135] D. Stirnweis, S.D. Milani, S. Brunner, G. Herren, G. Buchmann, D. Peditto, T. Jordan, B. Keller, Suppression among alleles encoding nucleotide-binding–leucine-rich repeat resistance proteins interferes with resistance in F_1 hybrid and allele-pyramided wheat plants, Plant J. 79 (2014) 893–903.
[136] V.S. Windhausen, C.N. Atlin, J.M. Hickey, J. Crossa, J.L. Jannink, M.E. Sorrells, B. Raman, J.E. Cairns, A. Tarekegne, K. Semagn, Y. Beyene, P. Grudloyma, F. Technow, C. Riedelsheimer, A.E. Melchinger, Effectiveness of genomic predictions of maize hybrid performance in different breeding populations and environments, G3 2 (2012) 1427–1436.
[137] A.J. Lorenz, K.P. Smith, J.L. Jannink, Potential and optimization of genomic selection for *Fusarium* Head Blight resistance in six-row barley, Crop Sci. 52 (2012) 1609–1621.
[138] **A.H. Sallam, J.B. Endelman, J.-L. Jannink, K.P. Smith, Assessing genomic selection prediction accuracy in a dynamic barley breeding population, Plant Genome 8 (2015) 1–15,** http://dx.doi.org/10.3835/plantgenome2014.05.0020
[139] K.J. Schmid, P. Thorwarth, Genomic selection in barley breeding, in: J. Kumlehn, N. Stein (Eds.), Biotechnological Approaches to Barley Improvement, vol. 69, Springer, Berlin, 2014, pp. 367–378.
[140] **B. Lado, I. Matus, A. Rodríguez, L. Inostroza, J. Poland, F. Belzile, A. del Pozo, M. Quincke, M. Castro, J. von Zitzewitz, Increased genomic prediction accuracy in wheat breeding through spatial adjustment of field trial data, G3 3 (2013) 2105–2114,** http://dx.doi.org/10.1534/g3.113.007807
[141] H.D. Daetwyler, U.K. Bansal, H.S. Bariana, M.J. Hayden, B.J. Hayes, Genomic prediction for rust resistance in diverse wheat landraces, Theor. Appl. Genet. 127 (2014) 1795–1803.
[142] S. Xu, D. Zhu, Q. Zhang Predicting hybrid performance in rice using genomic best linear unbiased prediction, Proc. Natl. Acad. Sci. U. S. A. 111 (2014) 12456–12461.
[143] F.G. Asoro, M.A. Newell, W.D. Beavis, M.P. Scott, J.-L. Jannink, Accuracy and training population design for genomic selection on quantitative traits in elite north American oats, Plant Genome 4 (2011) 132–144.
[144] Y.J. Shu, D.S. Yu, D. Wang, X. Bai, Y.M. Zhu, C.H. Guo, Genomic selection of seed weight based on low-density SCAR markers in soybean, Genet. Mol. Res. 12 (2013) 2178–2188.
[145] T. Würschum, J.C. Reif, T. Kraft, G. Janssen, Y. Zhao, Genomic selection in sugar beet breeding populations, BMC Genet. 14 (2013) 85.
[146] S. Kumar, D. Chagné, M.C. Bink, R.K. Volz, C. Whitworth, C. Carlisle, Genomic selection for fruit quality traits in apple (*Malus × domestica* Borkh.), PLoS One 7 (2012) e36674.

Glossary

CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats/Cas9(CRISPR-associated) is a tool based on the reprogramming of CRISPR/Cas9 endonuclease activity normally acting in prokaryote cells to target a specific sequencing, using short non-coding RNAs designed on the basis of the targeted sequence. Cas9 nuclease is used for genome engineering applications, and a single guide RNA (sgRNA) homologous to a target sequence is used to drive the CRISPR/Cas complex and to induce desired mutations (insertions/deletions). The target specificity is given by the sgRNA which forms the editing complex with the Cas enzyme, and whose sequence is engineered ad hoc.

GBS: Genotyping By Sequencing is genotyping method involving the digestion of genomic DNA with a frequent cutter restriction endonuclease and the sequencing of the ends of the resulting restriction fragments with a NGS platform. Adaptors containing barcodes and common adaptors (without barcodes) are mixed and used in the ligation reaction. Due to the specificity the NGS system only small fragments (between 100 and 250 bases) featuring a barcode-common adapter combination are yielded allowing simultaneous marker discovery and genotyping. SNPs are called by comparing DNA of different genotypes using dedicated bioinformatic pipeline.

GWAS: Genome-wide association studies aims at the detection and fine mapping of quantitative trait loci (QTLs) underlying complex agronomic traits thought linkage disequilibrium (LD) analysis. GWAS exploits ancestral recombination events that occurred in existing natural populations and takes into account all major alleles present in the population to identify significant marker-phenotype associations. In comparison to LD studies commonly performed on bi-parental populations, due to many more historical recombination events occurred in the population, a much higher mapping resolution is achieved.

POPSEQ: Population Sequencing is based on genome sequencing of a segregating population that allows de novo production of a genetically anchored linear assembly of the gene space. A high coverage whole-genome shotgun is generated for one parent, and used to construct a gene space assembly on which gene models are defined using RNA-seq. In parallel, a low coverage whole genome sequencing is performed on the whole population, and a medium-density framework genetic map is calculated. SNPs and associated sequencing contigs are then integrated into the framework map through nearest-neighbor search.

RAD: Restriction-site Associated DNA sequences short genomic DNA regions surrounding most restriction sites of a given restriction endonuclease. To achieve this, the restriction fragments are randomly sheared and fragments with a length suitable for the NGS platform of choice are selected after size fractionation and subjected to a PCR reaction designed to amplify for sequencing only those fragments containing the selected restriction site. SNPs are called by comparing DNA of different genotypes using dedicated bioinformatic pipeline.

TALEN: Transcription Activator-Like Effector Nucleases are a class of DNA-binding proteins produced by plant pathogens of the genus *Xanthomonas*, which during infection deliver the proteins to plant cells aimed at increasing the susceptibility to the pathogen infection. A central portion of TALE proteins that contains as many as 30 tandem repeats of a 33-35-amino-acid-sequence motif responsible for DNA binding, while the nuclease activity is executed by a *FokI* nuclease domain.

TILLING: Targeting Induced Local Lesions IN Genomes is a high-throughput reverse genetic technique for the discovery of mutations in a specific gene. Mutations are induced through chemical mutagenesis (usually by ethyl methane sulfonate - EMS) and detected through enzymatic cleavage of heteroduplex DNA molecules or through whole genome/exome re-sequencing (TILLING by sequencing).

ZFN: Zinc Finger Nucleases are engineered enzymatic complexes developed by fusing the non-specific cleavage domain from the *FokI* restriction endonuclease (responsible for the induction of DNA double strand cuts) with custom-designed Cys2-His2 zinc-finger proteins, able to guide the whole complex onto the desired position on the genome.