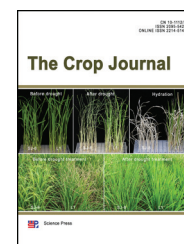


Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

ScienceDirect



# Harness the power of genomic selection and the potential of germplasm in crop breeding for global food security in the era with rapid climate change<sup>☆</sup>

Tianhua He<sup>a</sup>, Chengdao Li<sup>a,b,c,\*</sup><sup>a</sup>Western Barley Genetics Alliance, Western Australian State Agricultural Biotechnology Centre, College of Science, Health, Engineering and Education, Murdoch University, Murdoch, WA 6150, Australia<sup>b</sup>Agriculture and Food, Department of Primary Industries and Regional Development, South Perth, WA 6151, Australia<sup>c</sup>Hubei Collaborative Innovation Centre for Grain Industry, Yangtze University, Jingzhou 434025, Hubei, China

## ARTICLE INFO

### Article history:

Received 20 September 2019

Received in revised form 30

December 2019

Accepted 23 April 2020

Available online xxxx

## ABSTRACT

Crop genetic improvements catalysed population growth, which in turn has increased the pressure for food security. We need to produce 70% more food to meet the demands of 9.5 billion people by 2050. Climate changes have posed challenges for global food supply, while the narrow genetic base of elite crop cultivars has further limited our capacity to increase genetic gain through conventional breeding. The effective utilization of genetic resources in germplasm collections for crop improvement is crucial to increasing genetic gain to address challenges in the global food supply. Genomic selection (GS) uses genome-wide markers and phenotype information from observed populations to establish associations, followed by genome-wide markers to predict phenotypic values in test populations. Characterizing an extensive germplasm collection can serve a dual purpose in GS, as a reference population for predicting model, and mining desirable genetic variants for incorporation into elite cultivars. New technologies, such as high-throughput genotyping and phenotyping, machine learning, and gene editing, have great potential to contribute to genome-assisted breeding. Breeding programmes integrating germplasm characterization, GS and emerging technologies offer promise for accelerating the development of cultivars with improved yield and enhanced resistance and tolerance to biotic and abiotic stresses. Finally, scientifically informed regulations on new breeding technologies, and increased sharing of genetic resources, genomic data, and bioinformatics expertise between developed and developing economies will be the key to meeting the challenges of the rapidly changing climate and increased demand for food.

© 2020 Crop Science Society of China and Institute of Crop Science, CAAS. Production and hosting by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

<sup>☆</sup> Peer review under responsibility of Crop Science Society of China and Institute of Crop Science, CAAS.

\* Corresponding author at: Western Barley Genetics Alliance, Western Australian State Agricultural Biotechnology Centre, College of Science, Health, Engineering and Education, Murdoch University, Murdoch, WA 6150, Australia.

E-mail address: [c.li@murdoch.edu.au](mailto:c.li@murdoch.edu.au). (C. Li).

Peer review under responsibility of Crop Science Society of China and Institute of Crop Science, CAAS.

<https://doi.org/10.1016/j.cj.2020.04.005>

Please cite this article as: T. He and C. Li, Harness the power of genomic selection and the potential of germplasm in crop breeding for global fo..., The Crop Journal, <https://doi.org/10.1016/j.cj.2020.04.005>

## 1. Introduction

Crop breeding and genetic improvement were evident from the beginning of crop domestication 12,000–10,000 years ago [1,2], most likely by local farmers' selections based on the phenotype of key agronomic traits such as anti-shattering, high yields [2,3]. The advancements of crop breeding contributed significantly to meeting the needs of the human population, reaching three billion by 1960, before the beginning of the Green Revolution. Crop breeding equipped with modern genetic techniques led to the Green Revolution that produced varieties of staple cereal crops with several agronomic traits associated with significant increases in yield. Genetic techniques have also seen the stabilization in crop yield through the incorporation of genes for improved pest and disease resistance [4,5]. However, the increase in food productivity over time, while facilitating food security in the short run, has served as a catalyst for further population growth, which has again increased the pressure for food security [4]. Meeting the demand for more nutritious food around the globe, given climate change and increased competition for land and water resources, is a critical global challenge of our time.

Crop domestication and breeding transiting wild species to landraces, and further to elite genotypes were accompanied by a substantial narrowing of genetic diversity [6–9]. As a result, current crop varieties typically exhibit lower levels of genetic variation than their wild forms, especially in genomic regions that harbour agronomically relevant loci [10]. Narrow genetic variability in current cultivars is limiting our capacity to breed varieties with high yields, high grain quality, and resilient to environmental and climatic changes. For example, recent research in Australia found that early sowing combined with slower-developing wheat varieties could benefit from a longer growing season to increase wheat yields [11]. However, with the current understanding and breeding strategies that have delivered increasingly faster-developing cultivars over many decades, genotypes with slow development have been discarded by breeders, and genetic variation associated with slow development is currently lacking in Australian wheat cultivars. Because genetic diversity represents the fundamental key to breeding success and wide variation provides the basis for breeders to select varieties with continually improving crop performance, the consequences of direct intensive selection pose serious challenges for further crop improvement. Climate change is one of the significant challenges in producing sufficient food [12,13]. Warmer and hotter, with more frequent, extreme climate events have been projected for the 21st century [14,15]. Modelling crop growth and climate change have generated robust estimates of the potential impact of the climate under different emission scenarios on global crop production, with studies often concluding that dramatic climate changes could reduce yields in most crop species [16–19]. Indeed, a recent analysis examining the historical record of crop production suggested that changing climate has stalled wheat yields since 1990 in Australia [20].

The development and adoption of stress-resilient crop cultivars are imperative to ensure the sustainability of agricultural production and global food security. Uncertainties

in future climate change and the unprecedented rate of change require breeders to produce varieties that will perform well in diverse environmental settings and quick enough to keep pace with climate change [13]. To overcome the challenges, modern breeding will require: 1) significantly shorter breeding cycles; 2) development of varieties adapted to the current and rapid changing climate; 3) new genetic variability with potential to adapt future environments; 4) novel gene combinations for agronomic traits with improved grain yield, resistant to a wide range of biotic and abiotic stress into elite varieties; 5) crop varieties with improved nutrition and food quality but not at the expense of yield. Breeding strategies to combine the power of genomic selection (GS) and the potential of an extensive collection of germplasm, assisted by new technologies, will offer promise in crop breeding to contribute to global food security. In the following sections, we elaborate on how modern technologies meet these challenges. As several recent research and reviews have addressed the issue of using genomics to improve the utilization of global collection of germplasm [21–23]. In this review, we aim to explain how new technologies could empower the genomic selection.

## 2. The potential of germplasm collections

Germplasm collections have high levels of genetic diversity in many valuable agronomic traits. Many examples show that the genetic resources in germplasm collections can lead to substantial improvements in grain yield and agronomic performance, including the discovery and incorporation of beneficial genes such as the yellow leaf curl virus gene TY-1 in tomato [24], submerge tolerance gene *Sub1* in rice [25], and disease resistance gene *nMlo* in barley [26]. More than seven million crop plant accessions are currently held in gene banks worldwide, which represents one of the greatest, but mostly unexploited, opportunities for breaking crop productivity bottlenecks to accelerate yield gains [27]. Obstacles in exploiting germplasm include the overwhelming scale of germplasm collections, and our lack of an integrated methodology to use it, namely, how to: 1) accurately characterize germplasm accessions on a large scale, and 2) how to identify useful alleles and transfer them into elite cultivars for successful exploitation. Advances in high-throughput phenotyping and genotyping technologies, together with emerging biotechnologies, now provide opportunities to use exotic genetic variation in crop breeding [21–23].

There are challenges for sustainable use of germplasm and genetic resources. Current institutional policies and practice usually only allow resource sharing under restrictive licences, which may discourage cooperation in research and development activities [28]. Other challenges involve how benefits derived from the use of resource should be shared among those that have contributed to the collection and conservation of those resources [29]. Such challenges may be overcome through coordination to develop international obligations to share benefits derived from the use of genetic resource in exchange for the access to, and the use of, genetic resources in the development of new commercial varieties [29,30].

### 3. Genomic selection

Genomic selection assumes that the additive effects of many genes control the genetic variance for a trait, at a scale of hundreds to thousands of variants each with small effects [31]. Most of these variants may not be mappable or identifiable due to their minor effects. Instead of identifying associated loci for the trait, GS first uses genome-wide marker and phenotype (of the trait/s of interest) information from observed populations to establish associations between markers and phenotypes. It then uses genome-wide markers to predict the genetic value of test populations. Genomic selection has been successful in livestock breeding, with an estimated doubling of genetic gain in some species [32]. The success of the underlying methodology has influenced plant breeding [4].

The theoretical foundation for GS is Fisher's infinite model [33] that assumes that multiple genes could contribute to trait variation in the population. Genomic best linear unbiased prediction (GBLUP) models have been developed, based on Fisher's infinite model, and are widely used for GS analysis [34]. Recently, GBLUP modelling has been extended to incorporate  $G \times E$  interactions to improve prediction accuracy [35–39]. With GS gaining momentum, new prediction models, using computationally efficient Markov Chain Monte Carlo and Bayesian methods, have also been developed to consider non-additive genetic effects, the combined analysis of multiple traits in multiple environments, and conditional distributions of trait parameters [40,41].

A 'reference population' is necessary for GS, no matter which model is used. The reference population (or training population) consists of individuals that have been genotyped and phenotyped for the trait of interest and is used to build the predictive regression model or to train prediction models using machine learning techniques. The performance of genotyped individuals that have not been phenotyped (the 'test population' or 'selection candidates') can, therefore, be predicted without conventional field or glasshouse trials. For high prediction accuracy, a large reference population with high-density genome-wide markers is needed [28,42]. For example, Xu et al. [43] used 575 hybrid rice accessions as the reference population and achieved a high prediction accuracy of 0.64. A reference population and the consequent prediction model should be adaptable and updatable with regard to the marker density and phenotypes. Muleta et al. [44] showed that an updatable training population could improve prediction accuracy. Noticeably, Wallace et al. [4] argued that even a moderate prediction accuracy might be enough to progress crop breeding and be more cost-effective than otherwise possible.

The best-recognized advantage of GS in crop breeding is its potential to increase genetic gains in less time and shorten breeding cycles with higher efficiency in resource usage, relative to the conventional approach. Conventional crossing-selection breeding can take 10–30 years to produce a cultivar with the desired agronomic performance [13]. The climate, environment, and best agronomic practice at the time of cultivar release may be very different from when it was developed; therefore, efforts to develop climate and

stress-resilient cultivars may lag behind the changing climate and environment. Genomic selection uses genome-wide high-density markers as predictors for assessing and selecting early generations in breeding programmes, thereby facilitating the rapid selection of superior genotypes and shortening the breeding cycle. Breeding research at International Maize and Wheat Improvement Center (CIMMYT) has shown that GS can reduce the breeding cycle by at least half and produce lines with significantly increased agronomic performance [45].

Soybean was among the first crops where GS was used to improve yield and agronomic traits using genotyping-by-sequencing (GBS) in a breeding program [46]. Further practices have demonstrated the potential of GS to develop superior lines quickly and enhance the rate of genetic gain in other legume crops, such as pea, soybean, chickpea, groundnut, and pigeon pea [47]. In rice, Xu et al. [43] showed genetic gains of 19–76% in multiple traits, including grain yield per plant, thousand-grain weight, panicle number per plant, plant height, secondary branch number, grain number per panicle, panicle length, and primary branch number. The GS-assisted breeding research is evident in a wide range of crop and horticultural plants, such as rice [43,48], wheat [49,50], maize [51–53], barley [54]; millets [55], banana [56], cacao [57], carrot [58], cassava [59], soybean [46,60], and strawberry [61]. Recently, GS was also used for selecting parents for generating new breeding population, for example in wheat [62], rice [63], and maize [64].

The simultaneous selection for several agronomically important traits has been a challenge, and observations often show that some of these traits are negatively correlated, such as grain yield and protein content in wheat [65]. Significant genetic improvement is now possible by simultaneously selecting multiple traits, even those that are negatively correlated, through GS to develop varieties that combine several superior agronomic traits. For example, Michel et al. [66] showed a substantial selection response for high protein content while maintaining high yield using GS, despite the negative correlation between the two traits. Zeng et al. [48] reported the successful development of new elite varieties by combining major genes that significantly contribute to grain quality and yield over five years using a GS breeding approach.

Genomics selection becomes increasingly attractive in crop breeding, it is, however, not without limits and challenges. Juliana et al. [67] detailed technical challenges of applying GS in breeding of bread wheat, ranging from poor prediction accuracies across nurseries and families to complication of genotype and environment interactions. Here we recognize an additional limit. It has long been recognized that the efficiency of MAS depends on the genetic architecture of the trait [68]. For phenotypic traits with low heritability, such as grain yield, genomic selection using genotype and phenotype can substantially improve selection accuracy [69]. However, for traits with high heritability, such as plant height [60], disease resistances [70], grain weight [71], GS may have less advantages over the phenotype-based selection, as the cost-to-genetic-gain ratio of marker profiling (DNA sequencing) may not as competitive as the direct phenotypic selection for such traits.

#### 4. Integrating germplasm mining and genomic selection

The key to the success of GS is the establishment of a large, adaptable and updatable reference population for the prediction model [44]. Ideally, the reference population should contain a high level of genetic diversity, particularly of the desired traits, as predictive models usually achieve high accuracy when the genotype and phenotype to be tested are within the variation spectrum of the reference population. In contrast, depleted genetic diversity in current cultivar gene pools of traits related to, for example, yield gain, product quality, and adaptive to altered biotic and abiotic stress, requires the introduction of new genetic variation into the current elite cultivars [11]. Conventional methods for using diversity in germplasm only select subsets of lines from a collection of some pre-defined criteria (e.g. phenotype or geographical origin) for gene discovery and introgression breeding and only a small amount of diversity can be captured because of the constraints in costs and logistics. Integrating the characterization of germplasm and genomic selection could offer unique potential and power in crop breeding when using a large germplasm collection that has been genotyped and phenotyped as a reference population in the GS [45].

A large germplasm population can serve a dual purpose: 1) to serve as a training population in the prediction model, and 2) to characterize the germplasm accessions at a large-scale to mine desirable genetic variants for introduction into elite cultivars through the crossing. A large germplasm collection as a reference population in GS promises high prediction accuracy. Early research suggested that maximizing the phenotypic variance in the training population is the key to increase the prediction ability in GS [72,73]. Research so far shows promising prediction accuracies for potential use in germplasm [49,53,58]. For example, Crossa et al. [45] achieved a prediction accuracy of 0.5 to 0.7 for several agronomic traits (e.g. maturity, quality traits, and grain yield and yield components) when examining 8416 Mexican wheat landrace/accessions and 2403 Iranian wheat landrace accessions from the CIMMYT gene bank. In soybean, Jarquín et al. [74] have revealed relatively high prediction accuracies when analyzing the USDA soybean collection. These preliminary results on the accuracy of GS using gene bank accessions demonstrate the enormous potential of applying GS to introduce genetic variation in landrace/accessions into elite germplasm to form gene pools and populations suitable for pre-breeding programs. In situations with extremely large collections—for example, tens to hundreds of thousands—a representative core collection at the manageable size of thousands, determined by clustering, can effectively accelerate the flow of novel genes from gene bank accessions to elite lines within reasonable cost and logistics requirements. The reference population built from the germplasm collection is updatable with additional genotypes and phenotypes of new accessions, landraces or even wild species into the model [44].

#### 5. Technologies can empower the genomic selection

One single tool or technology is unlikely to achieve the desired genetic gain to meet the challenges of crop breeding. The

rapid advancement and successful integration of techniques and concepts in genotyping, phenotyping, cutting-edge biotechniques and bioinformatics will provide an unprecedented opportunity to use the potential of germplasm and power of GS to accelerate crop breeding. Below we discuss some key techniques that can be used in GS.

##### 5.1. High-throughput DNA sequencing technology

High-throughput DNA sequencing technology fundamentally enables the power of genomic selection. When Meuwissen and colleague proposed the framework for predicting breeding values using genome-wide marker information [75], the technology to generate genome-wide markers did not exist (simulated data were used in their paper). With genotyping technologies becoming more advanced and more affordable, high-density genome-wide markers generated from sequence data are now available for all major crop species. Genotyping-by-sequencing was among the first in crop genotyping and provided powerful tools for rapid, high-resolution mapping of genetic variation underlying agronomic traits of interest [4]. Commercially available SNP arrays target alleles of interest and allow rapid genotyping of large numbers of markers across the genome without the need for sequencing [76]. With the advancement of sequencing methodologies, and its cost to be reduced to <\$10 per Gb data in the late 2010s to \$1 per Gb soon, it is possible to produce genome-wide high-density genotyping of an entire germplasm collection with thousands of accessions through whole-genome sequencing in a short time frame. High-throughput genotyping platforms also open the way for targeted variant enrichment, allowing rapid genotyping of large numbers of hybrids or inbred lines for phenotype prediction.

##### 5.2. Targeted enrichment of genetic variants associated with important agronomic traits

Whole-genome sequencing could provide a complete genomic characterization of a species with up to 50 million common variable sites in the genome, while most of them probably do not imply phenotypic variation [4]. For traits of agronomic importance, significant DNA variants (SNPs) can be determined through targeted enrichment, such as reduced-representation sequencing, which would significantly reduce the cost of sequencing, facilitating the screening of a large number of individuals for the desired phenotype in GS. Reduced-representation sequencing includes exome capture or transcriptome sequencing that enables targeted identification of SNPs in protein-coding genome regions that is even more efficient when a reference genome is available. Capture probes can be custom designed to target specific chromosome regions that harbour candidate genes for traits of interest across a given species of interest [77–81]. Current methods allow the rapid discovery of a large number of DNA variants at low cost, which enables flexible scaling of resequencing work to deal with many genes at low cost in large test populations using GS. However, it is worth noting that reduced-representation sequencing strategies such as exome capture and transcriptome sequencing may miss important genetic variation as these techniques ignore intergenic regions that control a significant fraction of trait variation [82].

### 5.3. High-throughput phenotyping

Conventional methods in crop breeding require large-scale field trials and subsequent phenotyping of multiple agronomic traits, which leads to the ratio of selected individuals: screened individuals to close to one in one million. The power of GS is its capacity to accelerate genetic gains by shortening the breeding cycle, as the phenotype of tested hybrids of inbred lines is predicted with a computer, rather than measured in the field. Genomic selection, however, requires a large training population with precisely measured phenotypes across multiple environments to enhance prediction accuracy. Phenotyping of multiple agronomic traits simultaneously with high precision has been a significant bottleneck that limits the power and scale of many breeding programmes. High-throughput phenotyping platforms (HTPPs) now allow precise measurements of phenotypic variation in individuals, mostly in a non-destructive way allowing continuous measurements through developmental stages [83], representing a technology breakthrough in crop breeding and precision agriculture [84].

Remote sensing is one the core technologies in high-throughput phenotyping, and it can be used for a wide range of crop breeding objectives, including growth assessment [85], biomass and yield potential [86,87], adaptation to abiotic, e.g. drought [88], and biotic (susceptibility to pests and diseases). Remote sensing relies on sensory cameras, with multispectral cameras among the most widely used for crop monitoring and phenotyping [84]. Sensory cameras are usually carried by ground-based or aerial vehicles. Several remoting sensing platforms for high throughput phenotyping have been developed and commercially available (Table 1). These platforms range from ground-based to aerial systems and are usually equipped with multispectral sensory cameras and global positioning system (GPS) navigation device.

Among the current available platforms, remote sensing with unmanned aerial vehicles (UAV) offers unprecedented spectral, spatial and temporal resolution for a range of crop vegetation characteristics [89]. It can provide increased power to generate a large scale of phenotype data to be used in genomics-assisted plant breeding. The massive amounts of data collected can then be processed and analyzed for modelling through high-performance computing approaches. Watanabe et al. [90] evaluated the potential of remote sensing with UAV in phenotyping sorghum for training data in a GS model. They concluded that GS models generated with remote sensing perform similar to models with data collected with traditional measurements. Hassan et al. [91] drew a similar conclusion when phenotyping plant height for GS model in bread wheat. They emphasized that remote sensing with UAV provides a much faster way to obtain time-series estimates of phenotypes than traditional methods. The integration of high-throughput phenotyping technologies to measure multiple agronomic traits with high spatiotemporal resolution in GS offers great potential to increase selection intensity and efficiency, and thereby accelerates genetic gains in crop breeding.

Despite the promising perspective of high throughput phenotyping, significant challenges need to be overcome for it to have a real benefit to breeding programmes. For some

**Table 1 – An incomplete list of commercially available phenotyping platforms.**

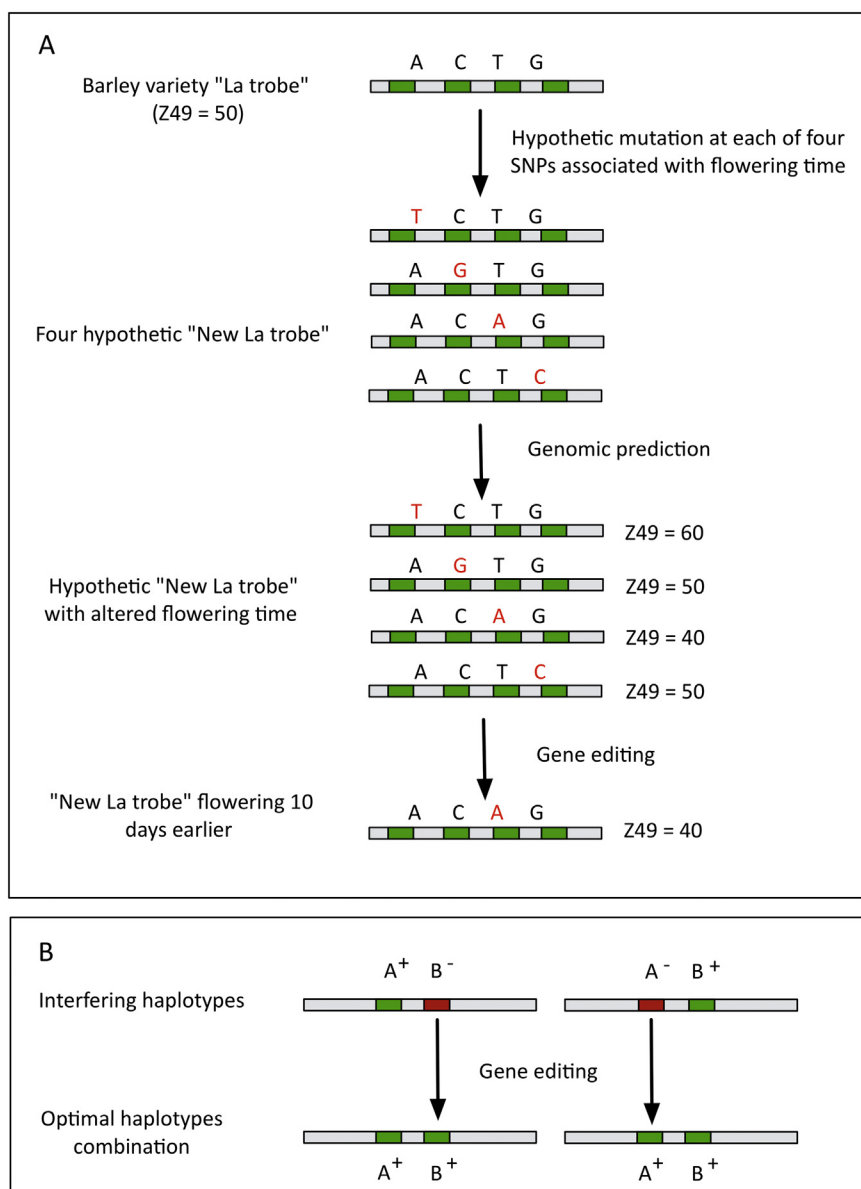
System	Platform	Company	Website
Blimps	Aerial (hot air balloon)	Not available	<a href="https://www.plant-phenotyping.org/">https://www.plant-phenotyping.org/</a>
Polycopters	Unmanned aerial	Ascending Technologies	<a href="http://www.asctec.de">http://www.asctec.de</a>
MikroKopter	Unmanned aerial	HiSystems GmbH	<a href="https://www.mikrokoetter.de/">https://www.mikrokoetter.de/</a> <a href="https://www.csiro.au/en/Do-business/Commercialisation/Marketplace/phenoMobile">https://www.csiro.au/en/Do-business/Commercialisation/Marketplace/phenoMobile</a>
Phenomobiles	Ground-based	CSIRO	<a href="http://www.phenospex.com">www.phenospex.com</a>
PlantEye	Ground-based	Phenospex	<a href="http://www.cropscience.bayer.com">http://www.cropscience.bayer.com</a>
Phenotower	Ground-based	Bayer Crop Science	

traits, such as root morphology and architecture, phenotyping technologies need to be refined to provide non-invasive image acquisition with appropriate spatial and time resolution [92]. While a range of automated and semi-automated phenotyping platforms are currently available for quantitative and dynamic analysis of phenotyping [93,94], the practicability of these methods under various field conditions also remains a challenge. Third, adopting high-technology phenotyping tools by ground or aerial platforms requires a large initial investment, which imposes a significant difficulty for high-throughput phenotyping as low-cost tools that can be applied across geographic locations with varying environmental conditions, especially in low- or middle-income [95].

### 5.4. The potential power of gene editing

The critical genetic technique to further increase the strength of GS-assisted breeding is genomic editing. Wallace et al. [4] predicted that direct genome editing would likely replace crosses as the most efficient way to alter specific alleles and pack desirable genetic variation into optimal combinations. Genome editing allows targeted mutations in crop genomes with base-pair precision (see Fig. 1 for a schematic illustration). Recently, CRISPR/Cas9 has emerged as one of the most promising systems to edit the crop genome with reported applications in rice, wheat, maize, banana, cassava, and tomato [96]. The surge of research in genome editing using the CRISPR/Cas9 system has seen these techniques increase in precision and efficiency. Several reviews are available on the potential of genome editing as a crop improvement tool [97–99].

As gene editing targeting DNA at base-pair precision, the bottleneck in realizing the potential of the gene-editing technique is the discovery and prioritisation of agronomically important genes [100]. The integration of gene editing and mining superior alleles from germplasm collections and wild relatives should be proven to be promising given the genetic impoverishment of many crops [99]. The successful use of gene-editing technology in breeding largely depends on our understanding of the mechanism of the genes controlling the



**Fig. 1 – Schematic illustration of the use of genome editing and targeted mutation for crop improvement. (A) Schematic illustration of the use of gene editing in genomic selection. The hypothetical scenario was to improve the elite Australian barley cultivar “La Trobe” to flower ten days earlier to avoid increased terminal heat stress as a consequence of climate change. Z49 is the number of days from sowing to the first awn emergence above the flag leaf in barley, as an indication of flowering. Note that genomic prediction is implemented with genome-wide SNP profiles; (B) Gene editing to break down interfering haplotypes in low recombination regions to form optimal haplotype combinations. <sup>+</sup>, beneficial allele; <sup>-</sup> deleterious allele. The selection acts weakly on the ‘interfering’ haplotypes (A<sup>+</sup> B<sup>-</sup> and A<sup>-</sup> B<sup>+</sup>), as the positive selection on one allele is counterbalanced by negative selection on the other.**

agronomic trait. Meanwhile, important agronomic traits (e.g. yield) are usually quantitative and controlled by many loci, editing in parallel on many genes could be a challenge in current practice. However, even editing a limited number of genes would likely progress breeding and lead to an improvement, albeit a small one [4]. Recently, Rodríguez-Leal et al. [101] propose to edit the regulatory sequences to generate new transcriptional alleles for fine-tuning of gene expression, so that most desirable transcriptional allele can be generated directly and selected for breeding.

### 5.5. Rapid generation advance

For GS to have a practical benefit on crop breeding, the target species should be able to achieve several life cycles per year [12,102], making it possible to genetically screen a large number of individuals of multiple generations in a short time. Rapid generation advancement techniques have been developed to accelerate breeding cycles and breeding progress by minimizing crop life cycles and have contributed to crop improvement [103–105]. Varieties or germplasm accessions

with extremely long or short life cycles have been observed in many crop species, which provide a useful platform for achieving rapid generation advancement. Currently, four to seven generations per year can be achieved in crop species such as barley, canola, chickpea, durum wheat, and wheat [106]. It may be possible to use controlled environmental conditions, extended photoperiods, and tissue culture to achieve even more generations within a year, which can further speed up genomic selection and breeding.

### 5.6. Bioinformatic analysis of genomic data: tools and resource

Bioinformatic analysis of genomic data is a critical component of crop genomics. The cost for the analysis of genomic data has now surpassed that of genomic sequencing, and bioinformatic analysis increasingly becomes a bottleneck in genomic studies [107]. The rapid development of analytic tools in genomic studies and the availability of powerful personal computers now allow researchers without advanced programming skills to be able to handle genomic data much easier than ever before. Traditionally, many crop genomic studies have used the user-friendly software package TASSEL [108] to evaluate trait associations, evolutionary patterns, and linkage disequilibrium in crop genomics. Recently, other tools developed for human genetics, e.g., principal component analysis software [109], PLINK [110], FaST-LMM [111], and GCTA [112], have been actively used in crop genomics. However, there has been limited interest in crop genomics with regard to the use of summary statistics [113], gene-set genome-wide association studies [114,115], and epistatic analysis [12,116], and multi-traits GWAS [117], which could be a fruitful area for future research. For GS and prediction models, several software packages are available (Table 2). Roorkiwal et al. [118] tested several models and relevant software packages (e.g. RR-BLUP, Kinship GAUSS, Bayes Cp, Bayes B, Bayesian LASSO, and random forest regression) in chickpea, and concluded that most of these models could produce high and consistent prediction accuracies for the traits of interest. In the meantime, web and cloud-based bioinformatics platforms, e.g. Galaxy (<https://usegalaxy.org/>) and GenomeSpace ([www.genomespace.org/](http://www.genomespace.org/)), provide accessible, reproducible, and transparent bioinformatic analysis, where complete computational studies can be built and shared through easy-to-use web interfaces. These platforms provide access to a diverse range of bioinformatics tools and bridge the gap between the tools, making it easy to integrate the available analyses, which will significantly facilitate crop genomic analysis for researchers who may not have the advanced computing skills needed for bioinformatics analysis.

Recently, publicly accessible database systems designed to manage and share information for crop improvement have become available. For example, the Triticeae toolbox (<https://triticeaetoolbox.org/>), incorporating the Triticeae toolbox (T3), Hordeum toolbox and T3 Oat, provides information on germplasm lines, pedigrees, and genotype and phenotypic data from breeding programs and core germplasm collections maintained by the USDA. Others include the integrated breeding platform (<https://www.integratedbreeding.net/>), wheat specific information consortia (<http://wheatis.org/>)

**Table 2 – An incomplete list of software for genomic selection.**

Package	Models	Reference	Website
BESSIE	Bayesian genomic linear regression	Boerner and Tier [119]	<a href="http://turing.une.edu.au/~agbu-admin/BESSIE/">http://turing.une.edu.au/~agbu-admin/BESSIE/</a>
BGGE	Linear mixed models	Granato et al. [120]	<a href="https://cran.r-project.org/web/packages/BGGE/index.html">https://cran.r-project.org/web/packages/BGGE/index.html</a>
BGLR	Bayesian linear regression	Pérez et al. [121]	<a href="https://cran.r-project.org/web/packages/BGLR/index.html">https://cran.r-project.org/web/packages/BGLR/index.html</a>
BLR	Bayesian genomic linear regression and Bayesian LASSO algorithm	Pérez et al. [121]	<a href="https://cran.r-project.org/web/packages/BLR/index.html">https://cran.r-project.org/web/packages/BLR/index.html</a>
DeepGS	Deep learning of convolutional neural network	Ma et al. [122]	<a href="https://github.com/cma2015/DeepGS">https://github.com/cma2015/DeepGS</a>
FaST-LMM	Machine learning	Lippert et al. [111]	<a href="https://github.com/MicrosoftGenomics/FaST-LMM">https://github.com/MicrosoftGenomics/FaST-LMM</a>
GAPIT	Extended genomic best linear unbiased prediction	Wang et al. [123]	<a href="http://zzlab.net/GAPIT">http://zzlab.net/GAPIT</a>
MTG2	Genomic best linear unbiased prediction	Maier et al. [124]	<a href="https://github.com/uqrmaie1/mtgblup">https://github.com/uqrmaie1/mtgblup</a>
RR-BLUP	Ridge regression BLUP	Endelman et al. [125]	<a href="https://cran.r-project.org/package=rrBLUP">https://cran.r-project.org/package=rrBLUP</a>
VIGoR	Variational Bayesian inference	Onogi et al. [126]	<a href="https://cran.r-project.org/web/packages/VIGoR/index.html">https://cran.r-project.org/web/packages/VIGoR/index.html</a>

and rice informatics consortium (<http://iric.irri.org>) that are designed to collate and translate genomic information for crop improvement. The significant challenges are the integration of such extensive, diverse data, and presentation to end-users in a user-friendly way with high efficiency.

### 5.7. Machine learning for genomic selection

Genomic selection requires adaptive prediction models that can provide accurate predictions of agronomic performance from large-scale genomic and phenotypic data that are usually complicated by multiple trait-environment combinations. The challenge for GS is the sheer and ever-increasing volumes of genomic and phenotypic data and complexity. Machine learning in genomic research offers a powerful tool to meet this challenge. Machine learning uses statistical techniques to allow computer systems to progressively analyze the dependencies of data and improve the performance of a program [127]. Machine learning can dissect complex relationships from massive biological datasets without a pre-defined genetic/biological hypothesis, which is particularly useful in GS, as the pathways from genes to phenotype are usually more complicated than simple additive effects. For example, machine learning with neural network approach builds multi-layered neural networks containing a

large number of neurons (data points) with non-linear transformations to model the complex dependent relationships within the extensive data (training population), further to predict a response variable (predicted phenotype value) in the input dataset (test populations) with improved performance [128].

Machine learning procedures can employ either a linear or a non-linear algorithm for prediction model [129–131]. Current comparative research suggests that no single method performs best in all cases [132–134]. Elements such as the size of the training dataset, the number of markers, the heritability and genetic architecture of the target trait, can all affect the performance of a particular algorithm [135–137]. Deep learning through such as neural network, are gaining popularity in genomic selection models, researches so far, however, did not find it outperforming other linear or non-linear models [132,133], though Azodi et al. believe that strategies, such as feature selection, seeded starting weights, may boost their performance [134]. It is therefore wise for breeders to test the performance of multiple models on the training population to identify the algorithm or a combination of algorithms performing best for target traits in their breeding program.

## 6. The role of agronomic practices and public concerns on genetic technology

It should be noted that agronomic inputs, such as labour, machinery, fertilizer and irrigation [138,139], and agronomic practices, such as early sowing [140], optimization of life duration [141], precision fertilization and irrigation, and disease management, are also critical for increasing and stabilizing crop production for global food security. Furthermore, the potential of genomic selection for increasing food security by improving crop yield, drought tolerance, or nutrition level can only be realized if the public accepts the products. Despite the promise of new biotechnologies, such as gene editing, and strong support from the scientific literature regarding safety and sustainability [142,143], many countries have strict regulations on its application in food production because of the varying acceptance from the public. Successful development and dissemination of high-yielding and climate-resilient crop varieties could serve as a clear example to demonstrate the capacities of using new biotechnologies in crop breeding to build the public's trust.

## 7. Concluding remarks

By selecting superior crop phenotypes in the past 12,000 years, humankind can produce one billion tons of food a year to feed 6.5 billion people at the beginning of this millennium [144]. Advances in crop breeding and improved agronomic practices facilitate food security, but also catalyse further population growth, which feeds back to increased pressure for food security [4]. Seventy percent more food needs to be produced by 2050, when the global population will exceed 9.5 billion as projected [145]. Added to this, dramatic climate changes will impact and destabilise the yields of most major crops, and the narrow genetic base of elite cultivars of major staple crops has limited our capacity to increase genetic gain through conventional breeding. Genomic selection assisted breeding can

rapidly produce genetically improved elite varieties, which is crucial if we are to sustain crop production in environments with rapid climate change. Current constraints in crop breeding can be partially addressed by the recent emergence of gene-editing technology that provides practical applications and tools to efficiently modify the relevant genes to form a desirable combination of beneficial genes to introduce into elite cultivars through genomic-assisted breeding methods [142]. The integration of germplasm characterization, high-throughput genotyping and phenotyping, machine learning in genomic analysis, and new biotechnologies will accelerate the development of climate-resilient cultivars with improved yields and enhanced resistance/tolerance to anticipated biotic and abiotic stresses to increase and stabilize crop production in farmer's fields, especially in the developing world and areas with stalling production as a consequence of rapid climate and environmental changes.

Finally, most undernourished people live in rural areas in developing countries, and agricultural outputs are the primary source of their food and income. Increased and sustained improvements in agricultural productivity, particularly crop production, are the key to their future food security and central to their socio-economic development. Genomic selection has excellent potential for improving and stabilizing crop production, but in most cases that potential is yet to be realized, which is particularly true for species outside of the major staple crops, such as the so-called orphan crops [146], and in countries outside the developed world. GS requires high-throughput genotyping systems and bioinformatics expertise that is inaccessible to most breeding programs in developing countries. Increased sharing of genomic resources, genomic data, and bioinformatics expertise between developed countries, developing regions and emerging economies will be the key to global food security in the era of rapid climate and environmental change [29,147].

## Declaration of competing interest

Authors declare that there are no conflicts of interest.

## Acknowledgments

The research is supported by the Australian Grain Research and Development Corporation (UMU00049 and UMU00050).

## Author contributions

Tianhua He and Chengdao Li conceived the idea, reviewed the literature and wrote the paper.

## REFERENCES

- [1] K. Khoshbakht, K. Hammer, How many plant species are cultivated? *Genet. Resour. Crop Evol.* 55 (2008) 925–928.



- [2] R.S. Meyer, A.E. DuVal, H.R. Jensen, Patterns and processes in crop domestication: an historical review and quantitative analysis of 203 global food crops, *New Phytol.* 196 (2012) 29–48.
- [3] J.F. Doebley, B.S. Gaut, B.D. Smith, The molecular genetics of crop domestication, *Cell* 127 (2006) 1309–1321.
- [4] J.G. Wallace, E. Rodgers-Melnick, E.S. Buckler, On the road to Breeding 4.0: unraveling the good, the bad, and the boring of crop quantitative genomics, *Ann. Rev. Genet.* 52 (2018) 421–444.
- [5] E.S. Lander, D. Botstein, Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps, *Genetics* 121 (1989) 185–199.
- [6] M.B. Hufford, X. Xu, J. van Heerwaarden, T. Pyhajarvi, J.M. Chia, R.A. Cartwright, R.J. Elshire, J.C. Glaubitz, K.E. Guill, S. M. Kaeppeler, J. Lai, P.L. Morrell, L.M. Shannon, C. Song, N.M. Springer, R.A. Swanson-Wagner, P. Tiffin, J. Wang, G. Zhang, J. Doebley, M.D. McMullen, D. Ware, E.S. Buckler, S. Yang, J. Ross-Ibarra, Comparative genomics of maize domestication and improvement, *Nat. Genet.* 44 (2012) 808–811.
- [7] C.R. Cavanagh, S. Chao, S. Wang, B.E. Huang, S. Stephen, S. Kiani, K. Forrest, C. Sautenac, G.L. Brown-Guedira, A. Akhunova, D. See, G. Bai, M. Pumphrey, L. Tomar, D. Wong, S. Kong, M. Reynolds, M.L. da Silva, H. Bockelman, L. Talbert, J.A. Anderson, S. Dreisigacker, S. Baenziger, A. Carter, V. Korzun, P.L. Morrell, J. Dubcovsky, M.K. Morell, M.E. Sorrells, M.J. Hayden, E. Akhunov, Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars, *Proc. Natl. Acad. Sci. U. S. A.* 110 (2013) 8057–8062.
- [8] J. Shi, J. Lai, Patterns of genomic changes with crop domestication and breeding, *Curr. Opin. Plant Biol.* 24 (2015) 47–53.
- [9] M.J. Kovach, S.R. McCouch, Leveraging natural diversity: back through the bottleneck, *Curr. Opin. Plant Biol.* 11 (2008) 193–200.
- [10] E.S. Mace, S. Tai, E.K. Gilding, Y. Li, P.J. Prentis, L. Bian, B.C. Campbell, W. Hu, D.J. Innes, X. Han, A. Cruickshank, Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum, *Nat. Commun.* 4 (2013) 2320.
- [11] H.A. Eagles, K. Cane, N. Vallance, The flow of alleles of important photoperiod and vernalisation genes through Australian wheat, *Crop Pasture Sci* 60 (2009) 646–657.
- [12] X. Li, T. Guo, Q. Mu, X. Li, J. Yu, Genomic and environmental determinants and their interplay underlying phenotypic plasticity, *Proc. Natl. Acad. Sci. U. S. A.* 115 (2018) 6679–6684.
- [13] A.J. Challinor, A.K. Koehler, J. Ramirez-Villegas, S. Whitfield, B. Das, Current warming will reduce yields unless maize breeding and seed systems adapt immediately, *Nat. Climat. Chang.* 6 (2016) 954.
- [14] B. Mueller, S.I. Seneviratne, Hot days induced by precipitation deficits at the global scale, *Proc. Natl. Acad. Sci. U. S. A.* 109 (2012) 12398–12403.
- [15] T. Cowan, A. Purich, S. Perkins, A. Pezza, G. Boschat, K. Sadler, More frequent, longer, and hotter heat waves for Australia in the twenty-first century, *J. Climat.* 27 (2014) 5851–5871.
- [16] D.B. Lobell, W. Schlenker, J. Costa-Roberts, Climate trends and global crop production since 1980, *Science* 333 (2011) 616–620.
- [17] A.J. Challinor, J. Watson, D.B. Lobell, S.M. Howden, D.R. Smith, N. Chhetri, A meta-analysis of crop yield under climate change and adaptation, *Nat. Clim. Chang.* 4 (2014) 287.
- [18] W. Xie, W. Xiong, J. Pan, T. Ali, Q. Cui, D. Guan, J. Meng, N. D. Mueller, E. Lin, S.J. Davis, Decreases in global beer supply due to extreme drought and heat, *Nat. Plants* 4 (2018) 964.
- [19] B. Liu, P. Martre, F. Ewert, J.R. Porter, J.A.J. Challinor, C. Müller, A.C. Ruane, K. Waha, P.J. Thorburn, P.K. Aggarwal, M. Ahmed, J. Balkovič, B. Basso, C. Biernath, M. Bindi, D. Cammarano, G. De Sanctis, B. Dumont, M. Espadafor, E.E. Rezaei, R. Ferrise, M. Garcia-Vila, S. Gayler, Y. Gao, H. Horan, G. Hoogenboom, R.C. Izaauralde, C.D. Jones, B.T. Kassie, K.C. Kersebaum, C. Klein, A. Koehler, A. Maiorano, S. Minoli, M. Montesino, S. Martin, S.N. Kumar, C. Nendel, G.J. O'Leary, T. Palosuo, E. Priesack, D. Dominique Ripoche, R. Rötter, M. Semenov, C. Stöckle, T. Streck, I. Supit, F. Tao, M. Van der Velde, D. Wallach, E. Wang, H. Webber, J. Wolf, L. Xiao, Z. Zhang, Z. Zhao, Y. Zhu, M. Asseng, Global wheat production with 1.5 and 2.0° C above pre-industrial warming, *Glob. Chang. Biol.* 25 (2019) 1428–1444.
- [20] Z. Hochman, D.L. Gobbett, H. Horan, Climate trends account for stalled wheat yields in Australia since 1990, *Glob. Chang. Biol.* 23 (2017) 2071–2081.
- [21] P.W. Wambugu, M. Ndjiondjop, R.J. Henry, Role of genomics in promoting the utilization of plant genetic resources in genebanks, *Brief. Funct. Genomics* 17 (2018) 198–206.
- [22] M. Mascher, M. Schreiber, U. Scholz, A.G. Jochen, C. Reif, N. Stein, Genebank genomics bridges the gap between the conservation of crop diversity and plant breeding, *Nat. Genet.* 51 (2019) 1076–1081.
- [23] N.L. Anglin, A. Amri, Z. Kehel, D. Ellis, A case of need: linking traits to Genebank accessions, *Biopreserv. Biobanking* 16 (2018) 337–349.
- [24] M.G. Verlaan, S.F. Hutton, R.M. Ibrahim, R. Kormelink, R.G. Visser, J.W. Scott, J.D. Edwards, Y. Bai, The tomato yellow leaf curl virus resistance genes Ty-1 and Ty-3 are allelic and code for DFDGD-class RNA-dependent RNA polymerases, *PLoS Genet.* 9 (2013), e1003399.
- [25] K. Xu, X. Xu, T. Fukao, P. Canlas, R. Maghirang-Rodriguez, S. Heuer, A.M. Ismail, J. Bailey-Serres, P.C. Ronald, D.J. Mackill, Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice, *Nature* 442 (2006) 705.
- [26] R. Büschges, K. Hollricher, R. Panstruga, G. Simons, M. Wolter, A. Frijters, R. van Daelen, T. van der Lee, P. Diergaarde, J. Groenendijk, S. Töpsch, The barley Mlo gene: a novel control element of plant pathogen resistance, *Cell* 88 (1997) 695–705.
- [27] C. Wang, S. Hu, C. Gardner, T. Lübberstedt, Emerging avenues for utilization of exotic germplasm, *Trend. Plant Sci.* 22 (2017) 624–637.
- [28] T.H. Mulesa, O.T. Westengen, Against the Grain? A Historical Institutional Analysis of Access Governance of Plant Genetic Resources for Food and Agriculture in Ethiopia, *J. World Intellectual Property* (2020)<https://doi.org/10.1111/jwip.12142>.
- [29] M. Brink, T. van Hintum, Genebank operation in the arena of access and benefit-sharing policies, *Front. Plant Sci.* 10 (2020) 1712.
- [30] J.K. Haile, A. N'Diaye, F. Clarke, J. Clarke, R. Knox, J. Rutkoski, F.M. Bassi, C.J. Pozniak, Genomic selection for grain yield and quality traits in durum wheat, *Mol. Breed.* 38 (2018) 75.
- [31] J. Crossa, P. Pérez-Rodríguez, J. Cuevas, O. Montesinos-López, D. Jarquín, G. de los Campos, J. Burgueño, J.M. González-Camacho, S. Pérez-Elizalde, Y. Beyene, S. Dreisigacker, Genomic selection in plant breeding: methods, models, and perspectives, *Trend. Plant Sci.* 22 (2017) 961–975.
- [32] M. Georges, C. Charlier, B. Hayes, Harnessing genomic information for livestock improvement, *Nat. Rev. Genet.* 20 (2019) 135–156.
- [33] R.A. Fisher, XV—the correlation between relatives on the supposition of Mendelian inheritance, *Trans. Roy. Soc. Edinburgh.* 52 (1918) 399–433.
- [34] S.A. Clark, J. van der Werf, Genomic best linear unbiased prediction (gBLUP) for the estimation of genomic breeding

- values, in: C. Gondro, J. van der Werf, B. Hayes (Eds.), *Genome-Wide Association Studies and Genomic Prediction*, Humana Press, Totowa, NJ, USA 2013, pp. 321–330.
- [35] J. Burgueño, G. de los Campos, K. Weigel, J. Crossa, genomic prediction of breeding values when modeling genotype× environment interaction using pedigree and dense molecular markers, *Crop Sci.* 52 (2012) 707–719.
- [36] D. Jarquín, J. Crossa, X. Lacaze, P. Du Cheyron, J. Daucourt, J. Lorgeou, F. Piraux, L. Guerreiro, P. Pérez, M. Calus, J. Burgueño, A reaction norm model for genomic selection using high-dimensional genomic and environmental data, *Theor. Appl. Genet.* 127 (2014) 595–607.
- [37] N. Heslot, D. Akdemir, M.E. Sorrells, J.L. Jannink, Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions, *Theor. Appl. Genet.* 127 (2014) 463–480.
- [38] M. López-Cruz, J. Crossa, D. Bonnett, S. Dreisigacker, J. Poland, J.L. Jannink, R.P. Singh, E. Autrique, G. de los Campos, Increased prediction accuracy in wheat breeding trials using a marker× environment interaction genomic selection model, *G3-Genes Genomes Genet.* 5 (2015) 569–582.
- [39] J. Cuevas, J. Crossa, O.A. Montesinos-López, J. Burgueño, P. Pérez-Rodríguez, G. de los Campos, Bayesian genomic prediction with genotype× environment interaction kernel models, *G3-Gene Genomes Genet.* 7 (2017) 41–53.
- [40] O.A. Montesinos-López, J. Crossa, J. Burgueño, K.M. Eskridge, E. Falconi-Castillo, X. He, P. Singh, K. Cichy, Genomic bayesian prediction model for count data with genotype × environment interaction, *G3-Genes Genomes Genet.* 6 (2016) 1165–1177.
- [41] O.A. Montesinos-López, J. Crossa, F.H. Toledo, J.C. Montesinos-López, P. Singh, P. Juliana, J. Salinas-Ruiz, A Bayesian Poisson-lognormal model for count data for multiple-trait multiple-environment genomic-enabled prediction, *G3-Genes Genomes Genet.* 7 (2017) 1595–1606.
- [42] M.E. Goddard, B.J. Hayes, Mapping genes for complex traits in domestic animals and their use in breeding programmes, *Nat. Rev. Genet.* 10 (2009) 381.
- [43] Y. Xu, X. Wang, X. Ding, X. Zheng, Z. Yang, C. Xu, Z. Hu, Genomic selection of agronomic traits in hybrid rice using an NCII population, *Rice* 11 (2018) 32.
- [44] K.T. Muleta, G. Pressoir, G.P. Morris, Optimizing genomic selection for a sorghum breeding program in Haiti: a simulation study, *G3-Genes Genomes Genet.* 9 (2019) 391–401.
- [45] J. Crossa, D. Jarquín, J. Franco, P. Pérez-Rodríguez, J. Burgueño, C. Saint-Pierre, P. Vikram, C. Sansaloni, C. Petrolí, D. Akdemir, C. Sneller, Genomic prediction of gene bank wheat landraces, *G3-Genes Genomes Genet.* 6 (2016) 1819–1834.
- [46] D. Jarquín, K. Kocak, L. Posadas, K. Hyma, J. Jedlicka, G. Graef, A. Lorenz, Genotyping by sequencing for genomic prediction in a soybean breeding population, *BMC Genomics* 15 (2014) 740.
- [47] R.K. Varshney, M. Thudi, M.K. Pandey, F. Tardieu, C. Ojiewo, V. Vadez, A.M. Whitbread, K.H.M. Siddique, H.T. Nguyen, P.S. Carberry, B.D. Bergvinson, Accelerating genetic gains in legumes for the development of prosperous smallholder agriculture: integrating genomics, phenotyping, systems modelling and agronomy, *J. Exp. Bot.* 69 (2018) 3293–3312.
- [48] D. Zeng, Z. Tian, Y. Rao, G. Dong, Y. Yang, L. Huang, Y. Leng, J. Xu, C. Sun, G. Zhang, J. Hu, Rational design of high-yield and superior-quality rice, *Nat. Plants* 3 (2017) 17031.
- [49] J.K. Haile, A.N. Diaye, F. Clarke, J. Clarke, R. Knox, J. Rutkoski, F.M. Bassi, C.J. Pozniak, Genomic selection for grain yield and quality traits in durum wheat, *Mol. Breed.* 38 (2018) 75.
- [50] D.L. Larkin, D.N. Lozada, R.E. Mason, Genomic selection—considerations for successful implementation in wheat breeding programs, *Agronomy* 9 (2019) 479.
- [51] A.T. Mastrodomenico, M.O. Bohn, A.E. Lipka, F.E. Below, Genomic selection using maize ex-plant variety protection germplasm for the prediction of nitrogen-use traits, *Crop Sci.* 59 (2019) 212–220.
- [52] Y. Beyene, K. Semagn, S. Mugo, A. Tarekegne, R. Babu., B. Meisel, P. Sehabiague, D. Makumbi, C. Magorokosho, S. Oikeh, J. Gakunga, M. Vargas, Mi. Olsen, B.M. Prasanna, M. Banziger, J. Crossa, Genetic gains in grain yield through genomic selection in eight bi-parental maize populations under drought stress, *Crop Sci.* 55 (2015) 154.
- [53] A. Allier, S. Teyssède, C. Lehermeier, A. Charcosset, L. Moreau, Genomic prediction with a maize collaborative panel: identification of genetic resources to enrich elite breeding programs, *Theor. Appl. Genet.* 133 (2020) 201–215.
- [54] M. Bhatta, L. Gutierrez, L. Cammarota, F. Cardozo, S. Germán, B. Gómez-Guerrero, M.F. Pardo, V. Lanaro, M. Sayas, A.J. Castro, Multi-trait genomic prediction model increased the predictive ability for agronomic and malting quality traits in Barley (*Hordeum vulgare* L.), *G3-Genes Genom Genet.* 10 (2020) 1113–1124.
- [55] Z. Liang, S.K. Gupta, C.T. Yeh, Y. Zhang, D.W. Ngu, R. Kumar, H.T. Patil, K.D. Mungra, D.V. Yadav, A. Rathore, R.K. Srivastava, R. Gupta, J. Yang, R.K. Varshney, P.S. Schnable, J. C. Schnable, Phenotypic data from inbred parents can improve genomic prediction in pearl millet hybrids, *G3-Genes Genomes Genet.* 8 (2018) 2513–2522.
- [56] M. Nyine, B. Uwimana, N. Blavet, E. Hřibová, H. Vanrespaille, M. Batte, V. Akech, A. Brown, J. Lorenzen, R. Swennen, J. Doležal, Genomic prediction in a multiploid crop: genotype by environment interaction and allele dosage effects on predictive ability in banana, *Plant Genome* 11 (2018) 170090.
- [57] M.S. McElroy, A. Navarro Jr., G. Mustiga, C. Stack, S. Gezan, G. Peña, W. Sarabia, D. Saquicela, I. Sotomayor, G.M. Douglas, Z. Mígicovsky, F. Amores, O. Tarqui, S. Myles, J.C. Motamayor, Prediction of cacao (*Theobroma cacao*) resistance to Monilophthora spp. diseases via genome-wide association analysis and genomic selection, *Front. Plant Sci.* 9 (2018) 343.
- [58] K.E. Corak, S.L. Ellison, P.W. Simon, D.M. Spooner, J.C. Dawson, Comparison of representative and custom methods of generating core subsets of a carrot germplasm collection, *Crop Sci.* 59 (2019) 1107–1121.
- [59] E.J. de Oliveira, M.D. de Resende, V. da Silva Santos, C.F. Ferreira, G.A. Oliveira, M.S. da Silva, L.A. de Oliveira, C.I. Aguilar-Vildoso, Genome-wide selection in cassava, *Euphytica* 187 (2012) 263–276.
- [60] A. Xavier, W.M. Muir, K.M. Rainey, Assessing predictive properties of genome-wide selection in soybeans, *G3-Genes Genomes Genet.* 6 (2016) 2611–2616.
- [61] S.A. Gezan, L.F. Osorio, S. Verma, V.M. Whitaker, An experimental validation of genomic selection in octoploid strawberry, *Hort. Res.* 4 (2017) 16027.
- [62] J. Yao, D. Zhao, X. Chen, Y. Zhang, J. Wang, Use of genomic selection and breeding simulation in cross prediction for improvement of yield and quality in wheat (*Triticum aestivum* L.), *Crop J.* 6 (2018) 353–365.
- [63] Y. Cui, R. Li, G. Li, F. Zhang, T. Zhu, Q. Zhang, J. Ali, Z. Li, S. Xu, Hybrid breeding of rice via genomic selection, *Plant Biotechnol. J.* 18 (2020) 57–67.
- [64] Y. Hao, H. Wang, X. Yang, H. Zhang, C. He, D. Li, H. Li, G. Wang, J. Wang, J. Fu, Genomic prediction using existing historical data contributing to selection in biparental populations: a study of kernel oil in maize, *Plant Genome* 12 (2019) 180025.
- [65] N.W. Simmonds, Yields of cereal grain and protein, *Exp. Agric.* 32 (1996) 351–356.
- [66] S. Michel, F. Löschenberger, C. Ametz, B. Pachler, E. Sparry, Simultaneous selection for grain yield and protein content in genomics-assisted wheat breeding, *Theor. Appl. Genet.* 132 (2019) 1745–1760.

- [67] P. Juliana, R.P. Singh, J. Poland, S. Mondal, J. Crossa, O.A. Montesinos-López, S. Dreisigacker, P. Pérez-Rodríguez, J. Huerta-Espino, L. Crespo-Herrera, V. Govindan, Prospects and challenges of applied genomic selection—a new paradigm in breeding for grain yield in bread wheat, *Plant Genome* 11 (2018) 180017.
- [68] F. Hospital, L. Moreau, F. Lacoudre, A. Charcosset, A. Gallais, More on the efficiency of marker-assisted selection, *Theor. Appl. Genet.* 95 (1997) 1181–1189.
- [69] D.N. Lozada, R.E. Mason, J.M. Sarinelli, G. Brown-Guedira, Accuracy of genomic selection for grain yield and agronomic traits in soft red winter wheat, *BMC Genet.* 20 (2019) 82.
- [70] J. Poland, J. Rutkoski, Advances and challenges in genomic selection for disease resistance, *Ann. Rev. Phytopath.* 54 (2016) 79–98.
- [71] R.A. Fischer, G.J. Rebetzke, Indirect selection for potential yield in early-generation, spaced plantings of wheat and other small-grain cereals: a review, *Crop Pasture Sci.* 69 (2018) 439–459.
- [72] A.A. Boligon, N. Long, L.G. de Albuquerque, K.A. Weigel, D. Gianola, G.J.M. Rosa, Comparison of selective genotyping strategies for prediction of breeding values in a population undergoing selection, *J. Animal Sci.* 90 (2012) 4716–4722.
- [73] J. Isidro, J.L. Jannink, D. Akdemir, J. Poland, N. Heslot, M.E. Sorrells, Training set optimization under population structure in genomic selection, *Theor. Appl. Genet.* 128 (2015) 145–158.
- [74] D. Jarquín, J. Specht, A. Lorenz, Prospects of genomic prediction in the USDA Soybean Germplasm Collection: historical data creates robust models for enhancing selection of accessions, *G3-Genes Genomes Genet.* 6 (2016) 2329–2341.
- [75] T.H. Meuwissen, B.J. Hayes, M.E. Goddard, Prediction of total genetic value using genome-wide dense marker maps, *Genetics* 157 (2001) 1819–1829.
- [76] A.C. Syvänen, Toward genome-wide SNP genotyping, *Nat. Genet.* 37 (2005) S5–S10.
- [77] M. Gholami, W.A. Bekele, J. Schondelmaier, R.J. Snowdon, A tailed PCR procedure for cost-effective, two-order multiplex sequencing of candidate genes in polyploid plants, *Plant Biotechnol. J.* 10 (2012) 635–645.
- [78] L.J. Clarke, P. Czechowski, J. Soubrier, M.I. Stevens, A. Cooper, Modular tagging of amplicons using a single PCR for high-throughput sequencing, *Mol. Ecol. Resour.* 14 (2014) 117–121.
- [79] M. Mascher, G.J. Muehlbauer, D.S. Rokhsar, J. Chapman, J. Schmutz, K. Barry, M. Muñoz-Amatriain, T.J. Close, R.P. Wise, A.H. Schulman, A. Himmelbach, Anchoring and ordering NGS contig assemblies by population sequencing (POPSEQ), *Plant J.* 76 (2013) 718–727.
- [80] S. Schiessl, B. Samans, B. Hüttel, R. Reinhard, R.J. Snowdon, Capturing sequence variation among flowering-time regulatory gene homologs in the allopolyploid crop species *Brassica napus*, *Front. Plant Sci.* 5 (2014) 404.
- [81] T.W. Rife, S. Wu, R.L. Bowden, J.A. Poland, Spiked GBS: a unified, open platform for single marker genotyping and whole-genome profiling, *BMC Genomics* 16 (2015) 248.
- [82] E. Rodgers-Melnick, D.L. Vera, H.W. Bass, E.S. Buckler, Open chromatin reveals the functional maize genome, *Proc. Natl. Acad. Sci. U. S. A.* 113 (2016) E3177–E3184.
- [83] J.N. Cobb, G. DeClerck, A. Greenberg, R. Clark, S. McCouch, Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype–phenotype relationships and its relevance to crop improvement, *Theor. Appl. Genet.* 126 (2013) 867–887.
- [84] J.L. Araus, J.E. Cairns, Field high-throughput phenotyping: the new crop breeding frontier, *Trend. Plant Sci.* 19 (2014) 52–61.
- [85] L. McAusland, J.A. Atkinson, T. Lawson, E.H. Murchie, High throughput procedure utilising chlorophyll fluorescence imaging to phenotype dynamic photosynthesis and photoprotection in leaves under controlled gaseous conditions, *Plant Methods* 15 (2019) 109.
- [86] V.S. Weber, J.L. Araus, J.E. Cairns, C. Sanchez, A.E. Melchinger, E. Orsini, Prediction of grain yield using reflectance spectra of canopy and leaves in maize plants grown under different water regimes, *Field Crops Res.* 128 (2012) 82–90.
- [87] M. Kim, J. Ko, S. Jeong, J. Yeom, H. Kim, Monitoring canopy growth and grain yield of paddy rice in South Korea by using the GRAMI model and high spatial resolution imagery, *GIScience Remote Sens.* 54 (2017) 534–551.
- [88] P.J. Zarco-Tejada, V. González-Dugo, L.E. Williams, L. Suárez, J.A.J. Berni, D. Goldhamer, E. Fereres, A PRI-based water stress index combining structural and chlorophyll effects: assessment using diurnal narrow-band airborne imagery and the CWSI thermal index, *Remote Sens. Environ.* 138 (2013) 38–50.
- [89] W.H. Maes, K. Steppe, Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture, *Trends Plant Sci.* 24 (2019) 152–164.
- [90] K. Watanabe, W. Guo, K. Arai, H. Takanashi, H. Kajiyama-Kanegae, M. Kobayashi, K. Yano, T. Tokunaga, T. Fujiwara, N. Tsutsumi, H. Iwata, High-throughput phenotyping of sorghum plant height using an unmanned aerial vehicle and its application to genomic prediction modeling, *Front. Plant Sci.* 8 (2017) 421.
- [91] M.A. Hassan, M. Yang, L. Fu, A. Rasheed, B. Zheng, X. Xia, Y. Xiao, Z. He, Accuracy assessment of plant height using an unmanned aerial vehicle for quantitative genomic analysis in bread wheat, *Plant Methods* 15 (2019) 37.
- [92] Y. Rouphael, L. Spíchal, K. Panzarová, R. Casa, G. Colla, High-throughput plant phenotyping for developing novel biostimulants: from lab to field or from field to lab? *Front. Plant Sci.* 14 (2018) 1197.
- [93] A.M. Mutka, R.S. Bart, Image-based phenotyping of plant disease symptoms, *Front. Plant Sci.* 5 (2015) 734.
- [94] Y. Fang, R.P. Ramasamy, Current and prospective methods for plant disease detection, *Biosensors* 5 (2015) 537–561.
- [95] J.L. Araus, S.C. Kefauver, M. Zaman-Allah, M.S. Olsen, J.E. Cairns, Translating high-throughput phenotyping into genetic gain, *Trend. Plant Sci.* 23 (2018) 451–466.
- [96] T. Wang, H. Zhang, H. Zhu, CRISPR technology is revolutionizing the improvement of tomato and other fruit crops, *Hort. Res.* 6 (2019) 77.
- [97] G. Song, M. Jia, K. Chen, X. Kong, B. Khattak, C. Xie, A. Li, L. Mao, CRISPR/Cas9: a powerful tool for crop genome editing, *Crop J.* 4 (2016) 75–82.
- [98] A. Scheben, F. Wolter, J. Batley, H. Puchta, D. Edwards, Towards CRISPR/Cas crops—bringing together genomics and genome editing, *New Phytol.* 216 (2017) 682–698.
- [99] F. Wolter, P. Schindele, H. Puchta, Plant breeding at the speed of light: the power of CRISPR/Cas to generate directed genetic diversity at multiple sites, *BMC Plant Biol.* 19 (2019) 176.
- [100] A. Scheben, D. Edwards, Bottlenecks for genome-edited crops on the road from lab to farm, *Genome Biol.* 19 (2018) 178.
- [101] D. Rodríguez-Leal, Z.H. Lemmon, J. Man, M.E. Bartlett, Z.B. Lippman, Engineering quantitative trait variation for crop improvement by genome editing, *Cell* 171 (2017) 470–480.
- [102] K.P. Voss-Fels, M. Cooper, B.J. Hayes, Accelerating crop genetic gains with genomic selection, *Theor. Appl. Genet.* 132 (2019) 669–686.
- [103] G. Rizal, S. Karki, M. Alcasid, F. Montecillo, K. Acebron, N. Larazo, R. Garcia, I.H. Slamet-Loedin, W.P. Quick, Shortening

- the breeding cycle of sorghum, a model crop for research, *Crop Sci.* 54 (2014) 520–529.
- [104] J. Tanaka, T. Hayashi, H. Iwata, A practical, rapid generation-advancement system for rice breeding using simplified biotron breeding system, *Breed. Sci.* 66 (2016) 542–551.
- [105] A. Watson, S. Ghosh, M.J. Williams, W.S. Cuddy, J. Simmonds, M.D. Rey, M.A. Hatta, A. Hinchliffe, A. Steed, D. Reynolds, N.M. Adamski, Speed breeding is a powerful tool to accelerate crop research and breeding, *Nat. Plants* 4 (2018) 23–29.
- [106] H. Li, A. Rasheed, L.T. Hickey, Z. He, Fast-forwarding genetic gain, *Trends Plant Sci.* 23 (2018) 184–186.
- [107] M.B. Scholz, C.C. Lo, P.S. Chain, Next generation sequencing and bioinformatic bottlenecks: the current state of metagenomic data analysis, *Current Opin. Biotechnol.* 23 (2012) 9–15.
- [108] P.J. Bradbury, Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss, E.S. Buckler, TASSEL: software for association mapping of complex traits in diverse samples, *Bioinformatics* 23 (2007) 2633–2635.
- [109] A.L. Price, N.J. Patterson, R.M. Plenge, M.E. Weinblatt, N.A. Shadick, D. Reich, Principal components analysis corrects for stratification in genome-wide association studies, *Nat. Genet.* 38 (2006) 904.
- [110] C.C. Chang, C.C. Chow, L.C. Tellier, S. Vattikuti, S.M. Purcell, J.J. Lee, Second-generation PLINK: rising to the challenge of larger and richer datasets, *Gigascience* 4 (2015) 7.
- [111] C. Lippert, J. Listgarten, Y. Liu, C.M. Kadie, R.I. Davidson, D. Heckerman, FaST linear mixed models for genome-wide association studies, *Nat. Methods* 8 (2011) 833.
- [112] J. Yang, S.H. Lee, M.E. Goddard, P.M. Visscher, GCTA: a tool for genome-wide complex trait analysis, *Am. J. Hum. Genet.* 88 (2011) 76–82.
- [113] B. Pasaniuc, A.L. Price, Dissecting the genetics of complex traits using summary association statistics, *Nat. Rev. Genet.* 18 (2017) 117.
- [114] D. Nam, J. Kim, S.Y. Kim, S. Kim, GSA-SNP, a general approach for gene set analysis of polymorphisms, *Nucleic Acids Res.* 38 (2010) 749–754.
- [115] T. He, C.B. Hill, T.T. Angessa, X.Q. Zhang, D. Moody, P. Telfer, S. Westcott, C. Li, Gene-set association and epistatic analyses reveal complex gene interaction networks affecting flowering time in a worldwide barley collection, *J. Expt. Bot.* 20 (2019) 5603–5616.
- [116] A.N. Doust, L. Lukens, K.M. Olsen, M. Mauro-Herrera, A. Meyer, K. Rogers, Beyond the single gene, how epistasis and gene-by-environment effects influence crop domestication, *Proc. Natl. Acad. Sci. U. S. A.* 111 (2014) 6178–6183.
- [117] G. Pei, H. Sun, Y. Dai, X. Liu, Z. Zhao, P. Jia, Investigation of multi-trait associations using pathway-based analysis of GWAS summary statistics, *BMC Genomics* 20 (2019) 79.
- [118] M. Roorkiwal, A. Rathore, R.R. Das, M.K. Singh, A. Jain, S. Srinivasan, P.M. Gaur, B. Chellapilla, S. Tripathi, Y. Li, J.M. Hickey, Genome-enabled prediction models for yield related traits in chickpea, *Front. Plant Sci.* 7 (2016) 1666.
- [119] V. Boerner, B. Tier, BESSiE: a software for linear model BLUP and Bayesian MCMC analysis of large-scale genomic data, *Genet. Sel. Evol.* 48 (2016) 63.
- [120] I. Granato, J. Cuevas, F. Luna-Vázquez, J. Crossa, O. Montesinos-López, J. Burgueño, R. Fritsche-Neto, BGGE: a new package for genomic-enabled prediction incorporating genotype × environment interaction models, *G3-Genes Genomes Genet.* 8 (2018) 3039–3047.
- [121] P. Pérez, G. de los Campos, J. Crossa, D. Gianola, Genomic-enabled prediction based on molecular markers and pedigree using the Bayesian linear regression package in R, *Plant Genome* 3 (2010) 106–116.
- [122] W. Ma, Z. Qiu, J. Song, J. Li, Q. Cheng, J. Zhai, C. Ma, A deep convolutional neural network approach for predicting phenotypes from genotypes, *Planta* 248 (2018) 1307–1318.
- [123] J. Wang, Z. Zhou, Z. Zhang, H. Li, D. Liu, Q. Zhang, P.J. Bradbury, E.S. Buckler, Z. Zhang, Expanding the BLUP alphabet for genomic prediction adaptable to the genetic architectures of complex traits, *Heredity* 121 (2018) 648.
- [124] R. Maier, G. Moser, G.B. Chen, S. Ripke, D. Absher, I. Agartz, H. Akil, F. Amin, O.A. Andreassen, A. Anjorin, R. Anney, Joint analysis of psychiatric disorders increases accuracy of risk prediction for schizophrenia, bipolar disorder and major depression disorder, *Am. J. Hum. Genet.* 96 (2015) 283–294.
- [125] J.B. Endelman, Ridge regression and other kernels for genomic selection with R package rrBLUP, *Plant Genome* 4 (2011) 250–255.
- [126] A. Onogi, H. Iwata, VIGoR: variational Bayesian inference for genome-wide regression, *J. Open Res. Software* 4 (2016), e11.
- [127] A.L. Samuel, Some studies in machine learning using the game of checkers, *IBM J. Res. Dev.* 3 (1959) 211–229.
- [128] J. Zou, M. Huss, A. Abid, P. Mohammadi, A. Torkamani, A. Telenti, A primer on deep learning in genomics, *Nat. Genet.* 51 (2018) 12–18.
- [129] M. Blondel, A. Onogi, H. Iwata, N. Ueda, A ranking approach to genomic selection, *PLoS One* 10 (2015), e0128570.
- [130] G.P. Ramstein, J. Evans, S.M. Kaeppler, R.B. Mitchell, K.P. Vogel, C.R. Buell, M.D. Casler, Accuracy of genomic prediction in switchgrass (*Panicum virgatum* L.) improved by accounting for linkage disequilibrium, *G3-Genes Genomes Genet.* 6 (2016) 1049–1062.
- [131] M. Roorkiwal, A. Rathore, R.R. Das, M.K. Singh, A. Jain, S. Srinivasan, P.M. Gaur, B. Chellapilla, S. Tripathi, Y. Li, J.M. Hickey, A. Lorenz, T. Sutton, J. Crossa, J.L. Jannink, R.K. Varshney, Genome-enabled prediction models for yield related traits in chickpea, *Front. Plant Sci.* 7 (2016) 1666.
- [132] J.O. Ogutu, H.P. Piepho, T. Schulz-Streeck, A comparison of random forests, boosting and support vector machines for genomic selection, *BMC Proc.* 5 (2011) S11.
- [133] A. Montesinos-López, O.A. Montesinos-López, D. Gianola, J. Crossa, C.M. Hernández-Suárez, Multi-environment genomic prediction of plant traits using deep learners with dense architecture, *G3-Genes Genomes Genet.* 8 (2018) 3813–3828.
- [134] C.B. Azodi, E. Bolger, A. McCarren, M. Roantree, G. de los Campos, S.H. Shiu, Benchmarking parametric and machine learning models for genomic prediction of complex traits, *G3-Genes Genomes Genet.* 9 (2019) 3691–3702.
- [135] C. Riedelsheimer, F. Technow, A.E. Melchinger, Comparison of whole-genome prediction models for traits with contrasting genetic architecture in a diversity panel of maize inbred lines, *BMC Genomics* 13 (2013) 452.
- [136] 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.
- [137] A. Norman, J. Taylor, J. Edwards, H. Kuchel, Optimising genomic selection in wheat: effect of marker density, population size and population structure on prediction accuracy, *G3-Genes Genomes Genet.* 8 (2018) 2889–2899.
- [138] A. Iglesias, L. Garrote, S. Quiroga, M. Moneo, A regional comparison of the effects of climate change on agricultural crops in Europe, *Climat. Chang.* 112 (2012) 29–46.
- [139] G.C. Nelson, H. Valin, R.D. Sands, P. Havlík, H. Ahammad, D. Deryng, J. Elliott, S. Fujimori, T. Hasegawa, E. Heyhoe, P. Kyle, Climate change effects on agriculture: economic responses to biophysical shocks, *Proc. Natl. Acad. Sci. U. S. A.* 111 (2014) 3274–3279.
- [140] J.R. Hunt, J.M. Lilley, B. Trevaskis, B.M. Flohr, A. Peake, A. Fletcher, A.B. Zwart, D. Gobbett, J.A. Kirkegaard, Early sowing systems can boost Australian wheat yields despite recent climate change, *Nat. Climat. Chang.* 9 (2019) 244.

- [141] B. Parent, M. Leclere, S. Lacube, M.A. Semenov, C. Welcker, P. Martre, F. Tardieu, Maize yields over Europe may increase in spite of climate change, with an appropriate use of the genetic variability of flowering time, *Proc. Natl. Acad. Sci. U. S. A.* 115 (2018) 10642–10647.
- [142] S.S. Zaidi, H. Vanderschuren, M. Qaim, M.M. Mahfouz, A. Kohli, S. Mansoor, M. Tester, New plant breeding technologies for food security, *Science* 363 (2019) 1390–1391.
- [143] A. Nicolia, A. Manzo, F. Veronesi, D. Rosellini, An overview of the last 10 years of genetically engineered crop safety research, *Crit. Rev. Biotechnol.* 34 (2014) 77–88.
- [144] R.L. Phillips, Mobilizing science to break yield barriers, *Crop Sci.* 50 (2010) S99–S108.
- [145] P.R. Ehrlich, J. Harte, Opinion: to feed the world in 2050 will require a global revolution, *Proc. Natl. Acad. Sci. U. S. A.* 112 (2015) 14743–14744.
- [146] R.K. Varshney, J.M. Ribaut, E.S. Buckler, R. Tuberosa, J.A. Rafalski, P. Langridge, Can genomics boost productivity of orphan crops? *Nat. Biotechnol.* 30 (2012) 1172.
- [147] M. Halewood, T. Chiurugwi, H.R. Sackville, B. Kurtz, E. Marden, E. Welch, F. Michiels, J. Mozafari, M. Sabran, N. Patron, P. Kersey, Plant genetic resources for food and agriculture: opportunities and challenges emerging from the science and information technology revolution, *New Phytol.* 217 (2018) 1407–1419.