REVIEW ARTICLE

CrossMark

Accelerating crop genetic gains with genomic selection

Kai Peter Voss-Fels[1] · Mark Cooper[1] · Ben John Hayes[1]

Abstract

Key message Genomic prediction based on additive genetic effects can accelerate genetic gain. There are opportunities for further improvement by including non-additive effects that access untapped sources of genetic diversity.

Abstract Several studies have reported a worrying gap between the projected global future demand for plant-based products and the current annual rates of production increase, indicating that enhancing the rate of genetic gain might be critical for future food security. Therefore, new breeding technologies and strategies are required to significantly boost genetic improvement of future crop cultivars. Genomic selection (GS) has delivered considerable genetic gain in animal breeding and is becoming an essential component of many modern plant breeding programmes as well. In this paper, we review the lessons learned from implementing GS in livestock and the impact of GS on crop breeding, and discuss important features for the success of GS under different breeding scenarios. We highlight major challenges associated with GS including rapid genotyping, phenotyping, genotype-by-environment interaction and non-additivity and give examples for opportunities to overcome these issues. Finally, the potential of combining GS with other modern technologies in order to maximise the rate of crop genetic improvement is discussed, including the potential of increasing prediction accuracy by integration of crop growth models in GS frameworks.

The need for speed

The urgent need to increase crop productivity in the face of climatic fluctuations and constantly rising demands for plant-based products challenges all related fields of agricultural and environmental research. The genetic improvement of crop cultivars through plant breeding is likely to play a crucial role for global food security in the future, especially in marginal growing regions with unstable production conditions. Forecasts are that up to 70% more plant-based products are going to be required by the middle of this century in order to meet the rapidly growing demand (Tester and Langridge 2010). The current annual rates of yield improvement for all major staple crops are insufficient to meet this goal (Fischer et al. 2014). This situation is exacerbated for geographies facing an increasing number of weather extremes as a consequence of climate change, which

challenges plant breeders and producers around the world and ultimately constrains the rate of realised genetic gain. Particularly, extreme heat and drought events are likely to cause increasingly severe yield losses in both developed and developing countries (Boyer et al. 2013; Lesk et al. 2016). Hall and Richards (2013) see a worrying gap between projected potential yields under optimum and water-limited conditions for all major cereals, and they consider production improvements through superior farmer-ready cultivars, which have been developed using latest technologies, will take decades rather than years. According to them, continued investments into research that aims at genetic improvement of future crop varieties is fundamental in order to tackle the worrisome production gaps.

Current status of genomic selection in crops and livestock

Plant breeding programmes are typically time- and cost-intensive, complex and rigid endeavours, and depending on the crop, it can take up to two decades until a new variety is released. This reflects the fact that breeders have to test variety candidates in multi-year and multi-location trials in order

✉ Ben John Hayes
   b.hayes@uq.edu.au

[1] Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, St Lucia, QLD 4072, Australia

to select superior genotypes with a high agronomic performance across a range of different environmental conditions at the highest possible precision. This is a main driver of costs and time of the breeding cycle and ultimately limits the number of variety candidates that can be tested.

A breeding technology that can potentially accelerate the rate of genetic gain in crops is genomic selection (GS) (Heffner et al. 2009; Meuwissen et al. 2001). To implement GS, a training population is first established, of individuals with phenotypes for the target trait(s) and genome-wide DNA marker genotypes (Meuwissen et al. 2001). The genotype and phenotype information from the training set is used to derive a prediction equation, which predicts the effect of each marker on the trait, with all marker effects fitted simultaneously. If the markers are in sufficient linkage disequilibrium (LD) with the causal mutations affecting the trait, they will capture a large proportion of the genetic variance for the trait. Selection candidates (e.g. seedlings) are then genotyped to obtain genomic estimated breeding values (GEBVs) using the genetic model obtained from the training population and the prediction equation. Predicted superior genotypes are subsequently selected based on their GEBVs. Over the past two decades, several different statistical models and machine learning methods have been proposed for GS, including methods which assume a normal distribution of SNP effects (e.g. ridge regression best linear unbiased prediction (RR-BLUP, Genomic BLUP), methods which assume a prior distribution of effects with a higher probability of moderate to large effects (BayesA, weighted Bayesian shrinkage regression wBSR), methods which assume that some SNP effects are zero (BayesB, BayesCπ), and nonparametric methods (random forest, reproducing kernel Hilbert space (RKHS) or neural network approaches) (Heslot et al. 2012). When choosing a genomic prediction method, the application of the genomic predictions should be carefully considered. For example, because many implementations of RKHS incorporate non-additive interactions, RKHS could be very useful for predicting future phenotypes (e.g. how a variety will perform). However, RKHS is less suitable for predicting breeding values (which are additive by definition) for selection programmes unless an additive model is specified or additive breeding values are explicitly extracted from the genomic predictions, in which RKHS becomes similar to BLUP (Bennewitz et al. 2009).

GS was implemented in dairy cattle in 2008, as soon as a reasonably low-cost genotyping array with 50,000 single nucleotide polymorphisms (SNPs) was available. The great advantage of GS for dairy cattle breeding was that bulls could be accurately selected based on their GEBV and used to provide semen to the industry for artificial insemination at the age of only 12 months. This is in contrast to traditional progeny schemes, where bulls were up to 7 years old before they had sufficient daughters with milking records to derive

accurate enough estimated breeding values (EBV) for selection. The reduction in generation interval from 7 years to only 12 months has doubled the rate of genetic gain over the past decade, compared with the rate of gain under the progeny testing system (García-Ruiz et al. 2016). Interestingly, some of the greatest gains in dairy cattle since the implementation of GS have been for low heritability traits such as fertility and disease (mastitis) resistance, which are economically important but very difficult to genetically improve using classical selection principles. Gain for these traits has almost tripled since the introduction of GS (García-Ruiz et al. 2016). The training population for these traits in dairy cattle is sufficiently large that highly accurate GEBV can be derived.

In chickens, Wolc et al. (2015) described the implementation of GS in a real population of brown laying hens. The population was split into one sub-population undergoing conventional selection, with two generations every 3 years, and one sub-population undergoing GS, with four generations over the 3 years. Birds were selected for breeding based on an index of performance traits relevant for commercial egg production. The birds selected based on genomic predictions outperformed those in the conventional breeding scheme, for nearly all the 16 traits that were included in the index used for selection, in some cases by up to 50%. However, the realised inbreeding per year was higher in the genomic selected line than in the conventionally selected line.

GS has also been implemented on a very large scale in pigs, sheep and beef cattle, with over 4 million animals genotyped for this purpose to date (Georges et al. 2018). Gains from implementing GS have been largest where there is the opportunity to reduce generation intervals and early selection on GEBV can be done, for example, in the dairy and chicken egg layer industries (Wolc et al. 2015). Gains have also been realised from implementing GS for hard or expensive to measure traits, which are economically important but have previously been ignored due to costs. Examples are heat tolerance in dairy cattle (Garner et al. 2016) and feed efficiency in beef cattle (Lu et al. 2016).

In crops, numerous studies report the successful prediction of phenotypic performance using molecular markers for all major species, including maize (Riedelsheimer et al. 2012; Bernardo and Yu 2007; Zhao et al. 2012), rice (Spindel et al. 2015), wheat (Heffner et al. 2011; Poland et al. 2012; Rutkoski et al. 2012), sorghum (Hunt et al. 2018; Fernandes et al. 2018), barley (Zhong et al. 2009; Lorenz et al. 2012), rapeseed (Werner et al. 2018a) or cassava (Oliveira et al. 2012; Ly et al. 2013). The development of the most efficient strategies to incorporate GS in the breeding programme is, however, an active research field and depends on many factors, including the mating system of the crop, the heritability and genetic architecture of the

trait, the availability and costs of genotyping platforms and the financial budget of the breeding programme (Heffner et al. 2009). While a growing body of evidence suggests that GS is becoming a substantial component of modern crop breeding programmes, and detailed simulations of these breeding programmes have predicted large increases in rates of genetic gain as a result of implementing GS (e.g. Gaynor et al. 2017; Gorjanc et al. 2018; Lin et al. 2016; Voss-Fels et al. 2018a), there are limited reports on the actual impact of GS on realised performance improvement.

An exception is maize in the private sector (Cooper et al. 2014b), for which there are industry-scale evaluations of the impact of drought-tolerant hybrids generated by GS and integration of other technologies from precision agriculture (Gaffney et al. 2015; Cooper et al. 2014b). Those varieties, referred to as the "AQUAmax" hybrids, were developed by integrating enhanced phenotyping in managed environments and information from crop growth models in genomic prediction frameworks through deployment of intermediate, yield-related traits that jointly determine genotype performance under drought (Cooper et al. 2014a). A large-scale evaluation of on-farm production data showed that AQUAmax maize hybrids were able to sustain significantly higher yields under both favourable and drought stress conditions in the USA, thereby significantly improving yield stability under water-limitation and reducing risks for producers (Gaffney et al. 2015).

While both plant and animal breeding have historically been built on quantitative genetics principles, theoretical concepts and applied breeding methods have diverged between the two main fields (Schön and Simianer 2015). Hickey et al. (2017) proposed the idea that GS has the potential to serve as a unifier between both branches, mainly because GS requires similar tools and concepts in both fields. They see a huge potential for significantly increasing genetic gain by establishing overarching platforms across both plant and animal kingdoms that integrate joint resources, data and multidisciplinary expert skillsets.

There are at least three key learnings of practical importance from implementing genomic selection in crop and livestock breeding programmes to date.

- The training population should include individuals (lines/varieties) that are closely related to the selection candidates (Daetwyler et al. 2012).
- The training populations must be very large. This reflects the large number of loci and very small effect size of these loci affecting a typical quantitative trait (e.g. yield). Estimates of the number of loci affecting quantitative traits likely range from 2000 to 4000 (e.g. MacLeod et al. 2016).
- To ensure the accuracy of the GEBV is maintained over time, the reference population must be frequently updated with new genotyped and phenotyped individuals (e.g. Podlich et al. 2004).

An important prerequisite for high prediction accuracies that persist across time is that markers have to be in strong linkage disequilibrium (LD) with QTL influencing the trait of interest (Meuwissen et al. 2001; Jannink et al. 2010). Accordingly, GS is most accurate in situations where the training and prediction populations are closely related and share long-range haplotypes (Meuwissen et al. 2016; Lorenz and Smith 2015; Cooper et al. 2014b), making the composition of the training population an essential basis for the success of GS.

The quality of phenotypes in the training population also turns out to be a key driver of the success of a GS programme. This can be understood through the equation that can be used to determine the size of the training population necessary to achieve a desired accuracy of GEBV. As described by Goddard and Hayes (2009), the accuracy of GEBV with a training population of size N, heritability of phenotypes h^2 and M_e independent loci affecting the trait is $\sqrt{Nh^2/(Nh^2 + M_e)}$ (Daetwyler et al. 2008; Hayes et al. 2009a). Under the infinitesimal model, M_e corresponds to the effective number of independent chromosome segments in the population. Various estimates for M_e are available, the simplest being $M_e = 2N_eL$, where N_e is the effective population size and L is the length of the genome in Morgans (Hayes et al. 2009b). For a wheat breeding programme, this could be $2 \times 50 \times 30 = 3000$. Taking yield as an example, and assuming that phenotyping has been done accurately (high level of plot replication, taking account of spatial variation, etc.) narrow-sense heritability might be $h^2 = 0.2$. If an accuracy of GEBV of 0.5 is desired (to enable rapid genetic gains), the above equation says that 5000 individuals are required in the training population to achieve this level of accuracy of GEBV. In contrast, if phenotyping has been done poorly (e.g. large spatial variation that is not accounted for, poor replication), and heritability is low (e.g. $h^2 = 0.1$), then the training population needs to include 10,000 genotyped and phenotyped individuals. Conversely, if improved experimental methods and phenotyping technologies are available and can be deployed such that heritability could be increased to 0.4, then the training population for genotyping and phenotyping would be 2500 individuals, which should be within the reach of the early stages of plant breeding programmes. Thus, to take advantage of genomic selection there are strong motivations for continuous efforts to enhance the quality of phenotyping data. Another consideration is that training populations should be large enough to capture reasonably rare alleles at frequencies which are sufficient to obtain reliable estimates of their effects (MacLeod et al. 2016).

In addition to the composition and size of the training population, GS models need to be updated regularly in order to maintain accuracy in the respective selection stages (Podlich et al. 2004; Heffner et al. 2011; Yabe et al. 2017). The main reason for that is that LD between markers and QTL decreases with increasing numbers of recombination events (particularly if low-density markers are used). For example, Auinger et al. (2016) showed in a data set from a rye breeding programme that training the prediction model using multiple breeding cycles significantly increased prediction accuracy as opposed to situations where predictions were made from the initial breeding cycle to the subsequent cycles when the breeding programme progressed. Podlich et al. (2004) also considered the need to update training data set composition when epistatic non-additive effects were sufficiently strong to result in changes in the average effects of alleles as selection changed allele frequencies at the different loci involved in the interacting networks.

Genotyping for genomic selection

Depending on the size of the breeding programme, implementing GS may require genotyping of thousands or tens of thousands of individuals throughout the breeding cycle, which poses a significant financial load for the breeder who is typically operating on a fixed budget. Sequencing-based genotyping approaches like genotyping-by-sequencing (GBS) have become very popular for genotyping large plant populations at moderate costs (Baird et al. 2008; Elshire et al. 2011). Today, there are several whole-genome and reduced-representation GBS approaches in which either the whole genome or only a fraction of it (e.g. the exome or transcriptome) is used for SNP marker identification (Scheben et al. 2017). Especially, reduced-representation GBS approaches seem to be a straightforward and cost-efficient strategy for breeding purposes. Another recent GBS strategy is skim-based GBS (skimGBS), which enables high-resolution genotyping via low-coverage sequencing (Bayer et al. 2015). This approach which uses low-coverage genomic reads, typically $< 1x$, is particularly powerful when populations of homozygous individuals, such as recombinant inbred lines (RIL) or doubled-haploid (DH) lines, are used and a high-quality reference sequence of the parents is available. For species with very complex genomes, such as wheat, however, it seems unlikely that skimGBS will surpass reduced-representation GBS approaches in terms of practicability and cost efficiency (Scheben et al. 2017; Trick et al. 2012). Gorjanc et al. (2017) investigated the potential of low-coverage GBS and imputation for GS in bi-parental plant populations in a comprehensive simulation study and found that non-imputed GBS data at a $1x$ coverage yielded comparable prediction accuracies to those

obtained from SNP array data, but at significantly higher returns of investments. When considering 100,000 markers and sequence coverage of only $0.01x$, their measurement for return of investments was even 5.63 times higher, implying a great potential of approaches that use low-coverage GBS and imputation for genotyping of large bi-parental breeding populations. An almost identical method which enables genotype imputation based on sequencing read data only without additional array or reference panel data has been shown to be extremely accurate under very low sequencing coverages ($0.15x$–$1.7x$ for mice and humans, respectively), provided a very large number of individuals are sequenced (Davies et al. 2016). Considering the anticipated further cost decreases for DNA sequencing, sequencing-based genotyping technologies could provide very flexible and cost-efficient solutions for plant breeding programmes in the future. Buckler et al. (2016) proposed a different genotyping platform specifically tailored to GS approaches called "rAmpSeq" in maize. This approach focuses on conserved genome regions to design PCR primers for amplification of thousands of middle repetitive regions, yielding in hundreds to thousands of markers, which can be scored for less than 5 USD per sample.

Since GS relies on LD between markers and QTL, GS would theoretically be most successful if every causal polymorphism in the genome was actually genotyped and considered in the model. The highest possible density is achieved if the whole genome is sequenced, which is becoming increasingly feasible and in fact genomic prediction accuracies were slightly increased in dairy cattle when whole-genome sequencing (WGS) data were used instead of low-coverage SNP data from a conventional genotyping array (Brøndum et al. 2015). Simulation studies have shown that nonlinear prediction methods (e.g. BayesR, BayesCpi) can even yield substantially higher accuracies if used in combination with WGS data, likely because causal molecular polymorphisms are prioritised while most polymorphisms with very small or zero effects are neglected (Meuwissen and Goddard 2010). For most crop species, however, the small effective population sizes leading to strong prevalence of genome-wide LD (Flint-Garcia et al. 2003) makes it questionable whether deployment of whole-genome sequences would improve prediction accuracies when in fact each LD block only needs to be tagged by one reliable marker. Therefore, another strategy to decrease genotyping costs is to reduce the number of markers used for genotyping. It was recently shown in rapeseed (Werner et al. 2018b) and barley (Abed et al. 2018) that several complex traits could be predicted at high accuracies when only a small subset of 1000–2000 polymorphic markers, which tag pronounced LD blocks, were chosen from the initial 20,000–35,000 genome-wide markers. This demonstrates a potential for significant genotyping cost reduction for breeders via customised genotyping

platforms with selected sets of markers that tag important QTL within key haplotype blocks (Cooper et al. 2014b; Qian et al. 2017). Especially in early stages of breeding programmes, which typically involve full- and half-sib family populations from bi-parental crosses, only very few markers (e.g. a few hundred) would in theory be required to track important haplotypes and eliminate unfavourable individuals early in the breeding cycle. Flexible genotyping platforms could enable breeders to fingerprint many more genotypes in their breeding programmes which would ultimately help to turn over much larger germplasm numbers. Highly multiplexed amplicon sequencing technologies could provide platforms to enable the use of targeted marker subsets in genome-wide genotyping (Yang et al. 2016).

## Phenotyping for genomic selection

As described above, high-quality phenotype data are essential for the success of any GS-supported breeding strategy (Desta and Ortiz 2014). At the same time, breeders have to balance the cost of high-quality phenotypic information for large numbers of individuals in their breeding programme with limited financial budgets. High-throughput phenotyping (HTP) technologies and remote sensing platforms are gaining popularity in plant breeding because they are typically non-invasive and provide the opportunity to generate useful trait information at an unprecedented scale. There are numerous examples for accurate modelling of yield-related traits, such as biomass, inflorescence density or plant height based on HTP platform data (Busemeyer et al. 2013; Araus and Cairns 2014; Wang et al. 2018), and it was shown that considering such traits in genomic prediction could significantly increase prediction accuracy (Araus et al. 2018). Rutkoski et al. (2016) used canopy temperature and vegetation index from aerial measurements in multi-variate models to predict grain yield in wheat and could show that incorporating these secondary traits resulted in significantly improved predictions compared to univariate models that considered only grain yield in the training population. Hayes et al. (2017) demonstrated that near-infrared (NIR) and nuclear magnetic resonance (NMR) spectra could be used to accurately predict end use quality traits for wheat from very small quantities of flour, which enabled large reference populations for predicting quality traits to be assembled. Krause et al. (2018) derived a relationship matrix among varieties from hyperspectral reflectance phenotyping data to improve prediction accuracies for yield in CIMMYT wheat data sets. While a huge variety of different HTP methods is available, data handling, management and processing is becoming a serious challenge (Araus and Cairns 2014; Araus et al. 2018). Tardieu et al. (2017) see a conceptual challenge in the field of modern plant phenotyping. According to them,

research to date was mainly focussed on the development of sensor and imaging technology, whereas novel methods to translate HTP information into useful knowledge (i.e. phenotypes) that can be explored in crop improvement are urgently required. Modern tools, such as machine learning or deep learning algorithms, have been shown to be useful for analysing HTP data, e.g. from stress phenotyping experiments (Singh et al. 2018). A promising concept that could significantly benefit from recent advances in HTP technologies is "envirotyping" which aims at capturing and accounting for sources of variation in agronomically important traits that are associated with quantifiable environmental variables (Cooper et al. 2014b). Envirotyping involves collecting environmental factors through multi-environment trials, geographic and soil information systems and empirical evaluations and has various applications, including environmental characterisation (Chapman et al. 2000b, c; Chenu et al. 2011), genotype × environment interaction analysis (Chapman et al. 2000a, b, c) and phenotypic prediction (Xu 2016). Accurate envirotyping can increase the performance of crop growth models, which can be integrated in modern crop breeding programmes (Fig. 1B). Envirotype parameters could also be included in the linear mixed models used for genomic prediction to increase heritability and improve prediction accuracy (van Eeuwijk et al. 2018).

Challenges to the success of genomic selection in crops

Genotype × environment interactions

The presence of genotype-by-environment interaction (GEI) for target traits presents potential challenges for the implementation of GS in crop breeding. In particular, there is concern that in the presence of GEI the substitution effects of QTL alleles will change among the target environments of the crop breeding programmes. The most challenging case is where there are changes in ranks of allele effects associated with the changes in substitution effects with environment (Boer et al. 2007), leading to crossover GEI (Haldane 1946; Baker 1988; Cooper 1999). In livestock, where genomic selection was first developed, the extent of GEI is relatively limited in many situations (e.g. Hayes et al. 2016). Consequently, GS has largely been implemented ignoring GEI. This is in sharp contrast to the situation in many crops where GEI can be very extensive (Basford and Cooper 1998; Elias et al. 2016; Chenu et al. 2011; Chapman et al. 2000a). Many studies have reported the presence of GEI for various traits at magnitudes that can negatively affect genetic gain in breeding. DeLacy et al. (1996) gave an historical perspective for plant breeding. GEI have been studied at many levels, extending from genome to phenome, including the primary
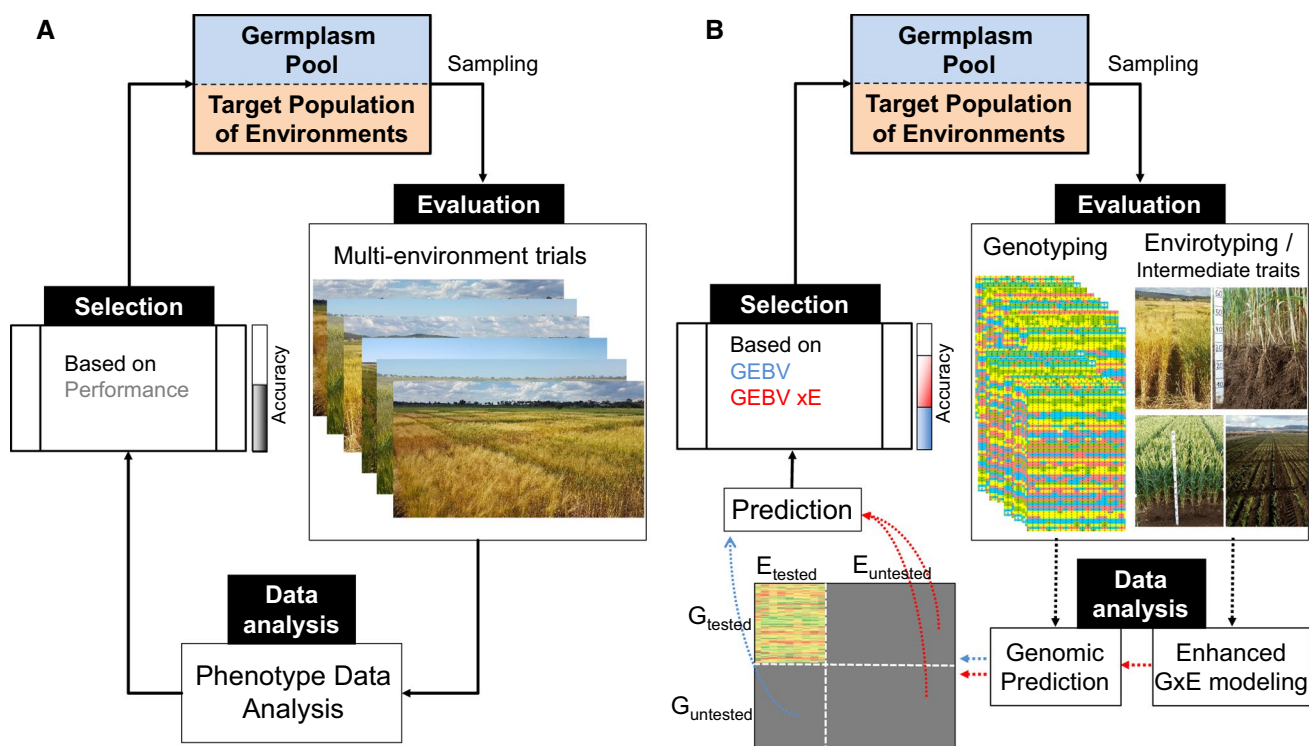
Fig. 1 Conceptual difference between representative conventional and modern crop breeding programme cycles. **a** A conventional crop breeding programme. To initiate the breeding cycle, breeders sample genotypes from their germplasm pool and choose field trial locations that represent their target population of environments (TPE) for multi-environment trials (MET). Genotypes are tested in multi-year and multi-location METs, and phenotype data are analysed. Selection decisions are made based on the genotypes' performance across the environments sampled in the METs. If genotype-by-environment interaction (GEI) is large, breeders might have to operate in separate breeding programmes. **b** A modern crop breeding programme including information from biophysical environmental characterisation (envirotyping), variation for important adaptive traits and integration of environmental and trait information based on a crop growth model (CGM). Breeders initiate the breeding cycle similar to **a**, and

genotypes are genotyped using high-throughput genotyping platforms and phenotyped in targeted METs which focus on generating information on traits that jointly determine the end-point trait (e.g. grain yield). Additionally, environments are characterised, and information can be used to compare environment frequencies as sampled in the MET with their expected frequencies in the TPE, to make weightings or adjustments to the environment sampling strategy. Conventional genomic prediction based on additive models is used, and where non-additive effects are large, CGM-based genomic prediction models are used to (1) predict the performance of unphenotyped genotypes in tested environments or (2) predict the performance of unphenotyped genotypes in and untested environments. Predictions are used to obtain genomic estimated breeding values for genotypes and to inform selection decisions

end-point traits such as grain yield of crops, biomass yield of forages, but also at the level of physiological traits that contribute to differences in these yield traits (Fukai and Cooper 1995; Chapman et al. 2003; Messina et al. 2011; Chenu et al. 2011; Bustos-Korts et al. 2016). Due to significant advances in molecular biology, GEI have recently also been investigated at the DNA sequence level to identify variable regions of the genome that contribute to the genetic variation for the physiological and yield traits associated with the GEI. It has also been investigated at the cellular level by analysing gene expression and pathways and how they are connected with the physiological traits they regulate (Chapman et al. 2003; Boer et al. 2007; Chenu et al. 2009; van Eeuwijk et al. 2018). The rapidly expanding body of experimental results is beginning to reveal and demonstrate the many possible biophysical connections between variable environmental

conditions, trait genetic variation and the GEI observed at the trait phenotype level.

From a breeding perspective, GEI should be studied in context with trait genetic variation representing the target germplasm pool and environmental variation representing the target population of environments (TPE) of the breeding programme (Comstock 1977). The presence of GEI is recognised as a significant factor that can limit genetic gain in breeding. Within the traditional breeder's equation $\Delta G_{\mathrm{P}} = \frac{i * h^2 * \sigma_{\mathrm{P}}}{t}$, where ΔG_{P} is response or gain in phenotypic performance from selection, i is the standardised selection differential applied to the selection units, (e.g. plant, plot or mean across plots), h^2 is heritability of the selection unit, σ_{P} is the phenotypic standard deviation of the target trait for the selection unit in the reference population, and t is the time to complete one cycle of the breeding programme,

the potential for GEI to limit gain is quantified through the downward impact on the heritability. In general for any trait, as the magnitude of the GEI variance increases relative to the magnitude of the genetic variance, the effective heritability of the trait will decrease and the breeder's equation will predict reduced gain (Holland et al. 2003). For genomic selection, the impact of GEI on gain can be understood through its potential to have a downward influence on the prediction accuracy (across the environments) component of the genomic breeder's equation $\Delta G_A = \frac{i * r_A * \sigma_A}{t}$, where ΔG_A is genetic gain in terms of breeding value, i is the standardised selection differential, r_A is the predictive accuracy, defined as the correlation between the estimated genotypic or breeding values in the training data sets and their corresponding true values in the TPE, σ_A is the additive genetic standard deviation of the target trait for the selection unit in the training population and t is the time to complete one cycle of the genomic selection cycle, which can be considerably faster than one cycle of the conventional breeding programme (Fig. 2). Thus, gain from genomic selection will be reduced whenever the incidence of GEI results in a decrease in the prediction accuracy (across environments). Likewise, if GEI is sufficiently large and breeding programmes have to be
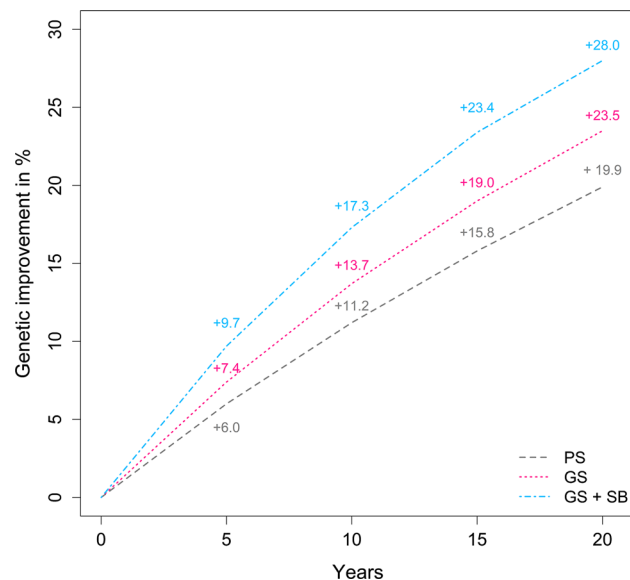
Fig. 2 Simulated genetic gain for spring wheat from three different breeding strategies. PS = phenotypic selection, GS = genomic selection, GS + SB = genomic selection + speed breeding. A fully additive genetic model with 1000 QTL was assumed, and effects were randomly sampled from a normal distribution. PS = phenotypic selection; GS = genomic selection; SB = 'speed breeding'. The cycle lengths for the three different programmes are PS = 5 years, GS = 4 years and PS + GS = 3 years. Breeding schemes were adapted from real CIMMYT breeding programmes. Simulation routines were initiated based on a spring wheat data set from CIMMYT. Data are extracted and adapted from simulations described in Voss-Fels et al. (2018a)

split to target specific production environments or regions, less genotypes can be tested in each of the sub-programmes which ultimately decreases the selection intensity, leading to lower rates of gain in each of the sub-programmes.

The ubiquitous nature of GEI for many crop species and their potential for negative impact on genetic gain motivates consideration of how to overcome these limitations. In principle, two components are required: (1) availability of appropriate descriptors of the environmental variables contributing to the GEI and (2) genetic diversity for the traits that contribute to differences in genotype performance in response to the causal environmental conditions. For example, if water availability is a key environmental variable contributing to genotypic variation for yield, firstly environments will need to be characterised to distinguish the different levels of water availability and their influences on yield performance (Chapman et al. 2000a; Löffler et al. 2005; Chenu et al. 2011). Secondly, trait variation contributing to differences in the yield performance of genotypes in response to the levels of water availability must exist and be phenotyped in the germplasm pools (and training population) of the breeding programme (Cooper et al. 2014a). To enable genomic selection that can account for the resulting GEI, methods are required to identify markers that can be used as descriptors of the trait variation associated with the GEI (e.g. Boer et al. 2007).

From the conventional plant breeding perspective, many methods have been advocated to deal with GEI. A widely used practice has been to conduct multi-environment trials (MET) to sample the diversity of environmental conditions within the TPE and select for genotypes demonstrating superior wide adaptation based on the results of these METs (Fig. 1A). When GEI is large, selecting for wide adaptation across highly diverse environment types may not be feasible. Practical extensions of this approach involve sub-dividing the TPE by grouping together environments with similar trait requirements and less GEI, and operating in separate breeding programmes for the defined environmental sub-groups. This strategy can work well when there is alignment between the biophysical causes of repeatable GEI and the identified sub-groups. In other examples, managed environments have been designed to stratify the sampling of some of the key environment types within the TPE (Fischer et al. 1989; Cooper et al. 1995; Campos et al. 2004; Kirigwi et al. 2004; Trethowan et al. 2005; Bänziger et al. 2006; Weber et al. 2012; Rebetzke et al. 2013; Cooper et al. 2014a). These conventional approaches to tackle GEI in plant breeding can be extended to enable GS by using the phenotypic data obtained from the METs and managed environments to construct training data sets to associate DNA sequence polymorphism with trait phenotypic variation (Cooper et al. 2014b; van Eeuwijk et al. 2018).

Environmental descriptors that explain components of GEI (Chapman et al. 2000a; Löffler et al. 2005; Chenu et al. 2011) have been used in combination with phenotypic data from METs to design GS methods to tackle GEI. The environmental descriptors have been used directly as covariates to index models for prediction to future environments (Heslot et al. 2014; van Eeuwijk et al. 2018; Malosetti et al. 2016). An alternative approach has been to use the environmental descriptors to group environments and design informative data breakouts from the total available MET data set. The data breakouts can then be used to construct separate training data sets for prediction to different environment types, e.g. drought, high temperature, low nitrogen, standability, specific biotic factors (e.g. nematodes, diseases, insects), favourable high input (low stress) or combinations of these environmental conditions (Cooper et al. 2014b). A logical progression from constructing such breakout data sets is targeted phenotyping of the training data sets in managed environments, e.g. for drought in maize (Cooper et al. 2014a). Designing METs based on combinations of both on-farm testing, to continually sample the diversity of the TPE, and managed environments, to ensure environment types that are important components of the TPE are consistently sampled, provides expanded experimental infrastructure to construct training data sets. Regardless of the approach, designing a MET that consistently provides samples of environments that are representative of the TPE is a key foundation for creating training data sets to enable GS. This also provides a suitable target for investing in extended phenotyping and advanced modelling methods to capture the nonlinear dynamics of trait adaptation that underpin components of GEI in genomic prediction models.

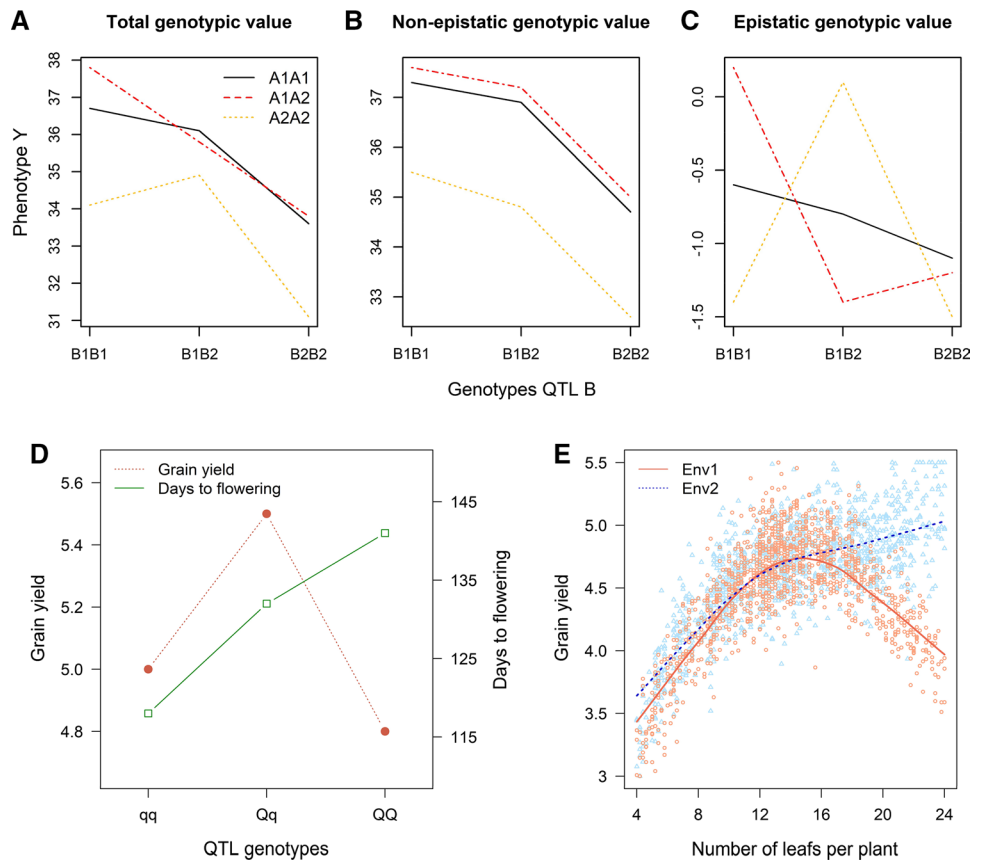## Non-additive genetic variance

The growing empirical molecular evidence that genes influencing complex traits can operate as members of networks has stimulated debates about the interpretation of gene effects and the relative contribution of additive and non-additive sources of genetic variance within reference populations and their potential influences on short- and long-term crop breeding trajectories (Hammer et al. 2006; Hill et al. 2008; Messina et al. 2011). Intra-locus (dominance) and inter-locus (epistasis) non-additivity for target traits present potential challenges for the implementation of GS in crop breeding. In particular, there is concern that in the presence of epistasis the substitution effects of QTL alleles can change among populations targeted for selection in crop breeding programmes. The most challenging case is where the changes in allele substitution effects are associated with changes in the ranking of the values of the alleles. Hill et al. (2008) reviewed theory for the relative importance of additive and non-additive genetic variance

at the outbred reference population level in the presence of different models of gene action, which included dominance and epistasis. They applied the theory to review the extensive body of empirical evidence from studies of laboratory animals, livestock and twin studies in humans, with a more limited treatment of the empirical evidence from laboratory and experimental studies of crop plants. They demonstrated the expected and observed predominance of additive genetic variance for outbred reference populations, emphasising the important influence of extreme allele frequencies on the detectability of non-additive gene action in the reference populations. Similar investigations would be helpful for pedigree-related inbred populations that represent the reference populations of major crop plants. Following the framework used by Hill et al. (2008), samples of reference breeding populations that are targeted for GS by the breeder can be examined to quantify the levels of consistency of allele substitution effects for QTL. Attention should be given to determining potential influences of changes in the ranking of allele effects associated with detected changes in substitution values. Where consistency is predominantly observed core training populations can be developed to support broad application of GS for use across multiple breeding populations (Cooper et al. 2014b). However, where lack of consistency is predominantly observed, multiple training populations or iterative updating of training populations will be required (Podlich et al. 2004), making GS more expensive to implement for crop breeding.

Cheverud and Routman (1995) provided a quantitative genetic framework for interpreting the influence of epistasis at the genotypic level (physiological epistasis) on genetic variance components and applied this to demonstrate how physiological epistatic effects can contribute to increased levels of additive genetic variance as populations evolve through bottlenecks (Cheverud and Routman 1996). They provided empirical results for epistatic effects of two interacting QTL on adult body weight of mice, which they quantified as the deviation of the total from the non-epistatic genotypic values (Fig. 3A–C). They demonstrated how physiological epistasis detected between QTL can contribute to increased levels of additive variance for adult body weight at the population level. The theory provided by Cheverud and Routman (1995, 1996) could be relevant to the additive genetic variance created within pedigree-related inbred populations of crop breeding where inbreds with contrasting levels of trait expressions are crossed to create new recombinants. Messina et al. (2011) demonstrated how physiological epistasis at the genotypic level (Cheverud and Routman 1995, 1996) can be understood as physiological epistasis measured at the trait-by-trait interaction level to determine genetic variance and short- and long-term response to selection for quantitative traits. These extensions of quantitative genetic theory to explicitly incorporate the

Fig. 3 Examples of non-additive relationships between genotypes/intermediate traits and phenotypes. Total (**a**), non-epistatic (**b**) and epistatic (**c**) genotypic values at QTL A and B for phenotype Y. The deviation of the non-epistatic genotypic values (**b**) from the total genotypic values (**a**) is given in **c** and is a consequence of the epistatic interaction between QTL A and B. Data are simulated and follow the example given by Cheverud and Routman (1995). **d** An increasing number of QTL alleles increase days to flowering; however, yield is maximised when only one QTL allele is present. This relationship is highly context dependent and changes between distinct environments (e.g. terminal drought vs. well watered) **e** Relationship between total leaf number per plant and grain yield in two environments. Data are simulated and follow data presented in the study from Technow et al. (2015)

effects of epistasis could enable the exploration of novel additive genetic variance for complex traits, such as yield, in terms of the underlying biophysical principles including trait-by-trait and trait-by-environment interactions, as studied by crop physiology, to sustain and enhance long-term genetic improvement in crop breeding. They also motivate the extension of GS methodology that explicitly accounts for the effects of non-additivity.

Tackling the combined effects of epistasis and GEI

A range of gene-to-phenotype (G2P) models have been proposed as suitable approaches for capturing both additive and non-additive genetic effects and GEI to enhance prediction for complex traits in breeding (Cooper et al. 2002; Hammer et al. 2006; Marjoram et al. 2014; van Eeuwijk et al. 2018). A growing body of work has considered the potential for crop growth models (CGMs) to complement conventional GS strategies (Messina et al. 2011; Chenu et al. 2009; Chapman et al. 2003; Technow et al. 2015). CGMs are a coordinated set of biophysical functions (equations) that translate the time-indexed influences of key environmental variables (e.g. temperature, photoperiod, radiation, water, nitrogen) into crop growth and development dynamics. Ultimately, these CGMs that are correctly parameterised for the required

set of genotype coefficients can be used to predict biomass and grain yield of genotypes, given the required set of environmental inputs. Within this framework, the elements of the biophysical functions included in the CGM can be used to model components of the genetic variation for adaptive traits as they influence end-point traits such as biomass and grain yield. For example, root architecture directly influences the plants ability to capture water (Voss-Fels et al. 2018b; van Oosterom et al. 2016), while canopy architecture affects the amount of captured radiation (Hammer et al. 2009). A simple example for a non-additive relationship between genotype and the end-point trait "grain yield" that highly depends on the environment is given in Fig. 3D. While an increasing number of QTL alleles increase days to flowering, grain yield in this environment is maximised when only one QTL allele is present. Such a form of non-additive variation is difficult to capture using classical additive genetic models but would be more appropriately accounted for using CGMs. Representing the conditional effects of adaptive traits on yield for the important environment types of a TPE opens a number of opportunities to model sources of non-additive genetic variation for yield to the extent that effects can be represented by the trait functions of the CGM (Cooper et al. 2009; Chapman et al. 2003; Chenu et al. 2009; Messina et al. 2011).

Technow et al. (2015) used approximate Bayesian computation (ABC) within genomic prediction to demonstrate how a CGM can be used to capture additive and non-additive yield effects for growth and development traits that are responsive to environmental variation. In their study, non-additivity in terms of genetic effects on grain yield was a result of nonlinear relationships between physiological traits and grain yield. An example for this is given in Fig. 3E in which the different relationships between the total number of leaves per plant and grain yield in two different environments are shown. Based on a CGM, they successfully predicted grain yield for maize under GEI in new environments, referring to this method as CGM–WGP (where WGP is whole-genome prediction). Cooper et al. (2016) applied the CGM–WGP method to predict yield variation for maize hybrids within drought environments and Messina et al. (2018) demonstrated a wider range of applications of the CGM–WGP for the prediction of GEI for yield of maize hybrids. Modelling the genetic architecture of yield variation for maize hybrids in these examples provided an enhanced interpretation of non-additive effects contributing to the observed yield variation in terms of underlying physiological traits. For example, results from the three CGM–WGP maize studies (Technow et al. 2015; Cooper et al. 2016; Messina et al. 2018) demonstrate that the non-additive gene actions underlying variation for yield can be interpreted as interactions among the intermediate traits included in the CGM, e.g. radiation use efficiency, rates of water use determined by rates of transpiration and canopy size, and reproductive resiliency determined by silking dynamics, which ultimately determine yield in drought and favourable environments. Similarly, GEI for yield can be interpreted as differing optimal combinations of intermediate trait expressions, depending on the environmental conditions (e.g. drought vs. optimum conditions). Messina et al. (2018) demonstrated in maize that accounting for these non-additive effects through incorporating intermediate traits that were individually modelled applying additive QTL models in the genetic model via a CGM resulted in improved genomic prediction accuracies compared to traditional methods that focused on fitting only additive effects associated with the end-point trait (grain yield). Further validation work will be insightful to examine the performance of CGM–WGP models in situations where predictions are made in subsequent generations of a breeding programme that have undergone several rounds of meiosis.

While the example presented by Technow et al. (2015) focused on the use of a CGM to capture additive and non-additive effects for crop yield, their methodology represents a more general class of genomic prediction models in which additional biological models can be used to supervise the selection of the genetic model for the target trait within the training data set. The conceptual difference between classical phenotypic selection schemes and modern programmes which incorporate CGM-based genomic prediction is represented in Fig. 1. Additional biological models that could be considered include gene network models to predict important developmental transitions such as from vegetative to reproductive development (Dong et al. 2012), and biochemical and hormone pathway models to predict critical levels of metabolites or hormones that regulate development and adaptation (Guo et al. 2014; Marjoram et al. 2014). Therefore, this framework opens new opportunities for modelling the genetic architecture of quantitative traits by identifying connections between the non-additive genetic phenomena such as epistasis, pleiotropy and GEI observed for the end-point traits and the context-dependent effects of the elements of any suitable biological model that can be used to predict the target end-point trait.

The availability of G2P modelling methods, such as the CGM–WGP, motivates consideration of how to effectively use advanced phenotyping methods to further improve outcomes from genomic selection. van Eeuwijk et al. (2018) provide a comprehensive review of a range of alternative G2P modelling methods that can be applied to the data obtained from METs. The traditional approach is to seek additional phenotypes that can be collected from METs (e.g. direct measurement of time series traits such as biomass, remote sensing of biomass accumulation, canopy structure and function, metabolites from tissues) and include these data as additional variables in the prediction model. An alternative approach proposed by Messina et al. (2018) is to focus significant phenotyping effort towards improving the biophysical functions included in CGMs and defining prior trait distributions as data inputs for the CGM–WGP.

Given the recent advances in genomics and our current understanding of the genetic architecture of quantitative traits in the elite germplasm pools of breeding programmes, it is worth considering how to use information on non-additive genetic effects in genomic selection. Following the studies of long-term genetic gain for yield of hybrid maize in the US corn-belt (Duvick et al. 2004; Campos et al. 2006; Hammer et al. 2009; Messina et al. 2011) and the recent applications of genomic selection to develop drought-tolerant maize hybrids for the US corn-belt (Cooper et al. 2014a), we propose that a potential benefit from modelling non-additive genetic effects for quantitative traits is the possibility to design GS strategies to access trait genetic diversity for breeding that could otherwise be difficult to access if selection is focused only on the standing additive genetic effects. This could be achieved by recruiting novel trait diversity associated with the non-additive genetic variation into the additive core of the breeding programme to contribute to long-term genetic gain, building on the extensions of quantitative genetic theory provided by Cheverud and Routman (1995, 1996) for evolution and Messina et al. (2011) for

plant breeding. A proof of concept for this approach has been demonstrated for the development of drought-tolerant maize hybrids for the USA. Early hypotheses for improving drought tolerance focused predominantly on variation for root system architecture and function to capture additional water (Hammer et al. 2009; Reyes et al. 2015; van Oosterom et al. 2016). Field-based comparisons of maize hybrids with high yield in drought environments revealed that traits associated with changes in temporal patterns of water use were more consistently associated with high yield (Cooper et al. 2014a). In particular, characteristics that facilitated water conservation during vegetative development providing a greater reserve of water during the reproductive development and grain filling stages were of high significance. These findings created new phenotyping opportunities and indicated an important role of novel traits associated with reduced transpiration rates under high atmospheric vapour pressure deficit conditions (Messina et al. 2015). With this physiological understanding of yield variation under drought, managed environments were designed to target the expression of traits contributing to different levels of limited transpiration and assess their impact on yield in both drought and favourable environments. The managed-environment data were then used to design appropriate training data sets to enable GS for enhanced yield stability of maize hybrids (Cooper et al. 2014a), which were validated in the US corn-belt, representing the TPE of the breeding programme (Gaffney et al. 2015). Through this targeted approach, trait diversity contributing to components of non-additive genetic diversity and GEI for yield that was initially hidden from the view of the additive models within the training data sets for the TPE was revealed by appropriate phenotyping in designed drought managed environments and creation of targeted drought and favourable environment training data sets and breeding for novel trait combinations could be accelerated by GS using weighted combinations of the drought, favourable and TPE training data sets (Cooper et al. 2014a).

Integration of genomic selection and other modern technologies in breeding programmes

Implementing genomic selection in modern plant breeding programmes

The implementation of GS as a core operation of commercial crop and livestock breeding programmes for major species like maize or dairy cattle has led to significant increases in genetic gain (Gaffney et al. 2015; García-Ruiz et al. 2016), and increasingly GS is becoming an essential component of breeding programmes for other important crops. In addition to the outlined associated challenges the definition of

the most optimal way to incorporate GS in a breeding programme remains an active field of research and depends on several factors (Lin et al. 2014). For wheat for example, classical breeding programmes create novel diversity by crossing parental lines to initiate the breeding cycle. From then on, superior genotypes are selected in time- and cost-consuming multi-stage selection processes (Fig. 1A). Ultimately, a handful of fully homozygous elite breeding lines are moved forward as variety candidates, which potentially enter the market as commercial cultivars for farmers (Koebner and Summers 2003). While this strategy, which is built on classical phenotypic selection, has led to significant improvements of modern varieties, the implementation of GS in breeding schemes holds the potential to increase the rate of gain further (Fig. 1B). The most obvious scope of GS for existing breeding schemes can be to support selection of improved genotypes in the different selection stages, but more sophisticated approaches have been proposed. Heffner et al. (2009) firstly proposed the idea to separate the germplasm improvement cycle from the prediction model improvement cycle in a GS-featured plant breeding programme, and Gaynor et al. (2017) extended this idea in a simulation study that investigated different strategies in a simulated inbred line development programme (e.g. for cultivar release or hybrid parent development). Using stochastic simulation, Gaynor et al. (2017) showed that the most optimal way for integrating GS is to divide the two main operations of a programme into two main operative parts. This includes (1) creation of new variation and recurrent population improvement and (2) the selection of superior inbred lines for variety development. Using computer simulations, they show that two-stage breeding programmes generate up to 2.5 times more genetic gain than a conventional phenotypic selection scheme and up to 1.5 times more genetic gain than the best performing standard GS strategy in which GS is used to improve selection in the breeding programme. Implementing such a strategy is, however, associated with a complete reorganisation of existing conventional programmes, and there is arguably need for further, if possible empirical, investigations of most optimal implementations.

Increasing gain through rapid generation advancement and genomic selection

Reducing the length of the breeding cycle is one of the main potential advances of GS-based breeding schemes compared to classical phenotypic selection. Even though GS enables breeders to select genotypes earlier in the cycle, selection candidates for inbred crops still have to go through line development via selfing or doubled-haploid (DH) technology until they can be tested in METs. While DH technology has been widely adopted in major crops and has led to a significant reduction in generation time, main disadvantages

associated with DHs are that the technology can be costly, at times inefficient in terms of natural or chemical-induced chromosome doubling and the breeder cannot select for basic traits during line development. Rapid generation advance techniques like "speed breeding" (Watson et al. 2018) are gaining popularity because they allow the breeder to turnover many generations under glasshouse conditions while enabling the breeder to select for traits with a high heritability. By modifying temperature and light regimes, plant growth can be greatly accelerated. For wheat, barley, or rapeseed for example, this enables the breeder to develop F6 lines, suitable for yield field trials, in only 1–1.5 years, thereby significantly reducing the line development phase of a breeding programme. While under speed breeding conditions breeders can discard variety candidates with unfavourable basic characteristics (e.g. disease susceptibility) early in the line development process, this method is not adapted to winter crops yet. Furthermore, costs associated with installation and running of suitable facilities currently constrain the widespread application of the tool. Using simulations based on real wheat data sets from CIMMYT, it was shown that combining GS and speed breeding could potentially increase genetic gain significantly, compared to both classical phenotypic selection and standard GS-based breeding schemes (Voss-Fels et al. 2018a). This study also demonstrated that introgression of novel diversity was facilitated using a combination of these technologies, resulting in sustained long-term genetic gain as opposed to the other more conventional approaches. La Fuente et al. (2013) conceptually extended the idea of rapid generation advance by introducing the idea of in vitro nurseries. These nurseries could be formed by in vitro production and subsequent fusion of gametes, thereby theoretically overcoming the entire process of plant growth. Marker genotyping could be done on gametes or new cell lines (resulting from gamete fusion), and most optimal combinations could be selected based on GEBVs. This would enable an extremely fast turn-around of generations of genetic material which could ultimately further accelerate the process of combining favourable alleles. Examples from animal breeding show that selection of embryos based on GEBVs is achievable at high accuracy, implying a great potential to reduce the breeding cycle length (Shojaei Saadi et al. 2014). Considering costs and status of plant genotyping and cell culture techniques, this approach is currently impracticable for plant breeding, but this may change in the future.

Leveraging genomic selection to harness useful diversity from gene banks

A frequently proposed idea for crop improvement, especially with regard to enhancing adaptive capacity, is to introgress novel allelic diversity which is absent in modern elite germplasm pools using genetic resources (Huang and Han 2014). While it is commonly agreed that there is a lot of potentially useful variation locked within many millions of wild accessions and landraces stored in gene banks worldwide, the actual utilisation of genetically exotic material, which is mostly poorly adapted to modern production systems, remains a real challenge. A key step is of course that exotic accessions must be phenotyped, which can be technically and financially challenging in a breeding programme. Longin and Reif (2014) proposed a new strategy for the exploitation of wheat genetic resources in modern breeding programmes by using hybrid technology in combination with GS. Since genotyping is becoming increasingly feasible even for large populations, accurate phenotyping remains the actual bottleneck in evaluations of genetic resources for breeding. Potentially advantageous genes or alleles from exotic sources (e.g. conferring biotic or abiotic stress resistance) are almost always masked by major deleterious alleles and the absence of important agronomic genes. Longin and Reif (2014) propose to cross the wild "donor" accession to an elite tester to produce F1 hybrids. Due to the dominant action of many agronomically important genes, e.g. dwarfing genes which played a major role for increasing grain yields under high nitrogen inputs, the resulting hybrids can be tested in modern crop production systems. Using genome-wide markers and GS principles, breeding values for the exotic accessions can be determined and used to specifically reinstate diversity for target traits in a given germplasm pool (Longin and Reif 2014). Gorjanc et al. (2016) used stochastic simulations to provide decision support for defining the most optimal strategy of how to initiate a pre-breeding programme in maize based on introgressions from landraces. They found that GS could be particularly useful to increase frequencies of novel alleles with small favourable effects in bridging germplasm, which provides a means to efficiently channel new diversity into elite material. Yu et al. (2016) proposed a GS-based strategy to predict the performance of gene bank germplasm using a large empirical sorghum data set. Based on strategic sampling of the training population, they were able to extrapolate information to untested exotic genotypes and accurately predict agronomic traits. Cowling et al. (2017) simulated pre-breeding with exotic populations from gene banks deploying the concept of optimal contribution selection for long-term genetic gain. They see a potential of GS for the exploration of genetic diversity by breaking up large linkage blocks, which are prominent in gene bank germplasm, through recurrent selection. Tanaka and Iwata (2018) proposed a Bayesian optimisation algorithm for genomic prediction of superior genotypes based on a simulation study. They found this strategy to be most efficient and potentially useful for GS-based pre-breeding strategies, ultimately

reducing the number of phenotyped accessions needed for the identification of the best genotype in a large germplasm population.

Combining genomic selection with genome editing

As a result of rapid advances in molecular biology, genome editing (GE) technologies have become very popular in crop and livestock research. One major field of application is the reversal of deleterious mutations (Johnsson et al. 2018; Hirsch and Springer 2018), which are ubiquitous in crop genomes (Kono et al. 2016). Several studies identified and sized the effects of such mutations on agronomic performance in important crop species such as cassava (Ramu et al. 2017) or maize (Mezmouk and Ross-Ibarra 2014). Even though modern breeding has systematically selected against these variants, a hypothesis for the lack of their complete removal is that selection is constrained by LD with favourable alleles and limited population sizes (e.g. genetic drift), highlighting the potential of targeted mutagenesis using GE techniques (Gibson 2012; Yang et al. 2017). More recently, the potential of GE in combination with GS has been proposed (Hirsch and Springer 2018). For example, Yang et al. (2017) demonstrate that prediction accuracies for grain yield and plant height in maize could significantly be improved when information about deleterious alleles was used to inform GS models and Ramu et al. (2017) see a potential in combining GS with GE for improving cassava by facilitating the purging of deleterious mutations throughout the genome. Bernardo (2017) proposes to use CRISPR technology to induce targeted recombination breakpoints along chromosomes and estimates that such an approach could double the rate of gain for quantitative traits in maize. Within this framework, genomic prediction can be used to predict marker effects, which are used to target optimal recombination points on each chromosome.

Conclusions

Increasing the rate of crop genetic improvement is essential for future food security. To achieve this, new breeding strategies and technologies are required to boost genetic gain. The significant improvements in animal production that have been achieved through the implementation of GS showcase the potential of the technology, and GS has been successfully incorporated in modern breeding programmes for major crops as well, predominantly in the private sector. While the methodological frameworks have been well established over the past two decades and there is mounting evidence for the opportunity of GS to accelerate crop improvement, the most optimal strategy to incorporate GS in plant breeding programmes is very species and breeding programme dependant and requires continued research. Further simulation work and empirical studies will be essential for the development of efficient GS-featured breeding strategies which enable the delivery of significantly higher rates of genetic gain at an acceptable cost. Ongoing improvements to genotyping efficiency and phenotyping technologies will increase the adoption of GS in plant breeding. Experience and knowledge from the past decades of public and private research can be built on, for example, to develop GS-based solutions to tackle problems associated with genotype-by-environment interaction and other sources of non-additivity. As outlined in this review, extending classical GS frameworks through inclusion of crop growth models that incorporate biological functions to model the biophysical processes that jointly determine the targeted end-point traits, e.g. grain yield, provides opportunities for the redesign of current plant breeding programmes. This can potentially also enhance the utilisation of elite and exotic sources of genetic diversity. Finally, just implementing GS alone may not be sufficient to close the worrisome gap between current production trends and the projected future demand for plant-based products. Continued investments into research that develops strategies combining GS, gene editing, high-throughput phenotyping, rapid generation advancement and technologies from other disciplines are crucial.

Compliance with ethical statement

References

Abed A, Pérez-Rodríguez P, Crossa J, Belzile F (2018) When less can be better. How can we make genomic selection more cost-effective and accurate in barley? Theor Appl Genet 131(9):1873–1890. https://doi.org/10.1007/s00122-018-3120-8

Araus JL, Cairns JE (2014) Field high-throughput phenotyping. The new crop breeding frontier. Trends Plant Sci 19(1):52–61. https://doi.org/10.1016/j.tplants.2013.09.008

Araus JL, Kefauver SC, Zaman-Allah M, Olsen MS, Cairns JE (2018) Translating high-throughput phenotyping into genetic gain. Trends Plant Sci 23(5):451–466. https://doi.org/10.1016/j.tplants.2018.02.001

Auinger H-J, Schönleben M, Lehermeier C, Schmidt M, Korzun V, Geiger HH, Piepho H-P, Gordillo A, Wilde P, Bauer E, Schön C-C (2016) Model training across multiple breeding cycles significantly improves genomic prediction accuracy in rye (Secale cereale L.). Theor Appl Genet 129(11):2043–2053. https://doi.org/10.1007/s00122-016-2756-5

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS ONE 3(10):e3376. https://doi.org/10.1371/journal.pone.0003376

Baker RJ (1988) Differential response to environmental stress. In: Weir BS, Eisen EJ, Goodman MM, Namkoong G (eds) Proceedings of the international conference on quantitative genetics. Sinauer Associates Inc., Sunderland, MA, pp 492–504

Bänziger M, Setimela PS, Hodson D, Vivek B (2006) Breeding for improved abiotic stress tolerance in maize adapted to Southern Africa. Agric Water Manag 80(1–3):212–224. https://doi.org/10.1016/j.agwat.2005.07.014

Basford KE, Cooper M (1998) Genotype × environment interactions and some considerations of their implications for wheat breeding in Australia This review is one of a series commissioned by the Advisory Committee of the Journal. Aust J Agric Res 49(2):153. https://doi.org/10.1071/A97035

Bayer PE, Ruperao P, Mason AS, Stiller J, Chan C-KK, Hayashi S, Long Y, Meng J, Sutton T, Visendi P, Varshney RK, Batley J, Edwards D (2015) High-resolution skim genotyping by sequencing reveals the distribution of crossovers and gene conversions in Cicer arietinum and Brassica napus. Theor Appl Genet 128(6):1039–1047. https://doi.org/10.1007/s00122-015-2488-y

Bennewitz J, Solberg T, Meuwissen T (2009) Genomic breeding value estimation using nonparametric additive regression models. Genet Sel. Evol. GSE 41:20. https://doi.org/10.1186/1297-9686-41-20

Bernardo R (2017) Prospective targeted recombination and genetic gains for quantitative traits in maize. Plant Genome. https://doi.org/10.3835/plantgenome2016.11.0118

Bernardo R, Yu J (2007) Prospects for genomewide selection for quantitative traits in maize. Crop Sci 47(3):1082. https://doi.org/10.2135/cropsci2006.11.0690

Boer MP, Wright D, Feng L, Podlich DW, Luo L, Cooper M, van Eeuwijk FA (2007) A mixed-model quantitative trait loci (QTL) analysis for multiple-environment trial data using environmental covariables for QTL-by-environment interactions, with an example in maize. Genetics 177(3):1801–1813. https://doi.org/10.1534/genetics.107.071068

Boyer JS, Byrne P, Cassman KG, Cooper M, Delmer D, Greene T, Gruis F, Habben J, Hausmann N, Kenny N, Lafitte R, Paszkiewicz S, Porter D, Schlegel A, Schussler J, Setter T, Shanahan J, Sharp RE, Vyn TJ, Warner D, Gaffney J (2013) The U.S. drought of 2012 in perspective. A call to action. Glob Food Secur 2(3):139–143. https://doi.org/10.1016/j.gfs.2013.08.002

Brøndum RF, Su G, Janss L, Sahana G, Guldbrandtsen B, Boichard D, Lund MS (2015) Quantitative trait loci markers derived from whole genome sequence data increases the reliability of genomic prediction. J Dairy Sci 98(6):4107–4116. https://doi.org/10.3168/jds.2014-9005

Buckler ES, Ilut DC, Wang X, Kretzschmar T, Gore MA, Mitchell SE (2016) rAmpSeq. Using repetitive sequences for robust genotyping. BioRxiv. https://doi.org/10.1101/096628

Busemeyer L, Ruckelshausen A, Möller K, Melchinger AE, Alheit KV, Maurer HP, Hahn V, Weissmann EA, Reif JC, Würschum T (2013) Precision phenotyping of biomass accumulation in triticale reveals temporal genetic patterns of regulation. Sci Rep 3:2442. https://doi.org/10.1038/srep02442

Bustos-Korts D, Malosetti M, Chapman S, van Eeuwijk F (2016) Modelling of genotype by environment interaction and prediction of complex traits across multiple environments as a synthesis of crop growth modelling, genetics and statistics. In: Yin X, Struik PC (eds) Crop systems biology. Springer International Publishing, Cham, pp 55–82

Campos H, Cooper M, Habben JE, Edmeades GO, Schussler JR (2004) Improving drought tolerance in maize. A view from industry. Field Crops Res 90(1):19–34. https://doi.org/10.1016/j.fcr.2004.07.003

Campos H, Cooper M, Edmeades GO, Löffler C, Schussler JR, Ibanez M (2006) Changes in drought tolerance in maize associated with fifty years of breeding for yield in the US corn Belt. Maydica 51:369–381

Chapman SC, Cooper M, Butler DG, Henzell RG (2000a) Genotype by environment interactions affecting grain sorghum. Aust J Agric Res 51(2):197. https://doi.org/10.1071/AR99020

Chapman SC, Cooper M, Hammer GL, Butler DG (2000b) Genotype by environment interactions affecting grain sorghum. II. Frequencies of different seasonal patterns of drought stress are related to location effects on hybrid yields. Aust J Agric Res 51(2):209. https://doi.org/10.1071/ar99021

Chapman SC, Hammer GL, Butler DG, Cooper M (2000c) Genotype by environment interactions affecting grain sorghum. III. Temporal sequences and spatial patterns in the target population of environments. Aust J Agric Res 51(2):223. https://doi.org/10.1071/ar99022

Chapman SC, Cooper M, Podlich D, Hammer G (2003) Evaluating plant breeding strategies by simulating gene action and dryland environment effects. Agron J 95:99–113

Chenu K, Chapman SC, Tardieu F, McLean G, Welcker C, Hammer GL (2009) Simulating the yield impacts of organ-level quantitative trait loci associated with drought response in maize. A "gene-to-phenotype" modeling approach. Genetics 183(4):1507–1523. https://doi.org/10.1534/genetics.109.105429

Chenu K, Cooper M, Hammer GL, Mathews KL, Dreccer MF, Chapman SC (2011) Environment characterization as an aid to wheat improvement. Interpreting genotype-environment interactions by modelling water-deficit patterns in North-Eastern Australia. J Exp Bot 62(6):1743–1755. https://doi.org/10.1093/jxb/erq459

Cheverud JM, Routman EJ (1995) Epistasis and its contribution to genetic variance components. Genetics 139(3):1455–1461

Cheverud JM, Routman EJ (1996) Epistasis as a source of increased additive genetic variance at population bottlenecks. Evol Int J Org Evol 50(3):1042–1051. https://doi.org/10.1111/j.1558-5646.1996.tb02345.x

Comstock RE (1977) Quantitative genetics and the design of breeding programs. In: Pollack E, Kempthorne O, Bailey TB (eds) Proceedings of the international conference on quantitative genetics. Iowa State University Press, Ames, IA, USA, pp 705–718

Cooper M (1999) Concepts and strategies for plant adaptation research in rainfed lowland rice. Field Crops Res 64(1–2):13–34. https://doi.org/10.1016/S0378-4290(99)00048-9

Cooper M, Woodruff DR, Eisemann RL, Brennan PS, Delacy IH (1995) A selection strategy to accommodate genotype-by-environment interaction for grain yield of wheat. Managed-environments for selection among genotypes. Theor Appl Genet 90(3–4):492–502. https://doi.org/10.1007/bf00221995

Cooper M, Chapman SC, Podlich DW, Hammer GL (2002) The GP problem. Quantifying gene-to-phenotype relationships. In Silico Biol 2(2):151–164

Cooper M, van Eeuwijk FA, Hammer GL, Podlich DW, Messina C (2009) Modeling QTL for complex traits. Detection and context for plant breeding. Curr Opin Plant Biol 12(2):231–240. https://doi.org/10.1016/j.pbi.2009.01.006

Cooper M, Gho C, Leafgren R, Tang T, Messina C (2014a) Breeding drought-tolerant maize hybrids for the US corn-belt. Discovery to product. J Exp Bot 65(21):6191–6204. https://doi.org/10.1093/jxb/eru064

Cooper M, Messina CD, Podlich D, Totir LR, Baumgarten A, Hausmann NJ, Wright D, Graham G (2014b) Predicting the future of plant breeding. Complementing empirical evaluation with genetic prediction. Crop Pasture Sci 65(4):311. https://doi.org/10.1071/cp14007

Cooper M, Technow F, Messina C, Gho C, Totir LR (2016) Use of crop growth models with whole-genome prediction. Application to a

maize multienvironment trial. Crop Sci 56(5):2141. https://doi.org/10.2135/cropsci2015.08.0512

Cowling WA, Li L, Siddique KHM, Henryon M, Berg P, Banks RG, Kinghorn BP (2017) Evolving gene banks. Improving diverse populations of crop and exotic germplasm with optimal contribution selection. J Exp Bot 68(8):1927–1939. https://doi.org/10.1093/jxb/erw406

Daetwyler HD, Villanueva B, Woolliams JA (2008) Accuracy of predicting the genetic risk of disease using a genome-wide approach. PLoS ONE 3(10):e3395. https://doi.org/10.1371/journal.pone.0003395

Daetwyler HD, Swan AA, van der Werf JHJ, Hayes BJ (2012) Accuracy of pedigree and genomic predictions of carcass and novel meat quality traits in multi-breed sheep data assessed by cross-validation. Genet. Sel Evol GSE 44:33. https://doi.org/10.1186/1297-9686-44-33

Davies RW, Flint J, Myers S, Mott R (2016) Rapid genotype imputation from sequence without reference panels. Nat Genet 48(8):965–969. https://doi.org/10.1038/ng.3594

de Oliveira EJ, de Resende MDV, da Silva Santos V, Ferreira CF, Oliveira GAF, da Silva MS, de Oliveira LA, Aguilar-Vildoso CI (2012) Genome-wide selection in cassava. Euphytica 187(2):263–276. https://doi.org/10.1007/s10681-012-0722-0

DeLacy IH, Basford KE, Cooper M, Bull JK, McLaren CG (1996) Analysis of multi environmental trials: an historical perspective. In: Cooper M, Hammer GL (eds) Plant adaptation and crop improvement. Wallingford, UK, pp 39–124

Desta ZA, Ortiz R (2014) Genomic selection. Genome-wide prediction in plant improvement. Trends Plant Sci 19(9):592–601. https://doi.org/10.1016/j.tplants.2014.05.006

Dong Z, Danilevskaya O, Abadie T, Messina C, Coles N, Cooper M (2012) A gene regulatory network model for floral transition of the shoot apex in maize and its dynamic modeling. PLoS ONE 7(8):e43450. https://doi.org/10.1371/journal.pone.0043450

Duvick DN, Smith JSC, Cooper M (2004) Long-term selection in a commercial hybrid maize breeding program. Plant Breed Rev 24:109–151

Elias AA, Robbins KR, Doerge RW, Tuinstra MR (2016) Half a century of studying genotype × environment interactions in plant breeding experiments. Crop Sci 56(5):2090. https://doi.org/10.2135/cropsci2015.01.0061

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE 6(5):e19379. https://doi.org/10.1371/journal.pone.0019379

Fernandes SB, Dias KOG, Ferreira DF, Brown PJ (2018) Efficiency of multi-trait, indirect, and trait-assisted genomic selection for improvement of biomass sorghum. Theor Appl Genet 131(3):747–755. https://doi.org/10.1007/s00122-017-3033-y

Fischer KS, Edmeades GO, Johnson EC (1989) Selection for the improvment of maize yield under moisture deficits. Field Crops Res 22:227–243

Fischer T, Byerlee D, Edmeades G (2014) Crop yields and global food security. Will yield increase continue to feed the world? ACIAR Monograph, Canberra

Flint-Garcia SA, Thornsberry JM, Buckler ES (2003) Structure of linkage disequilibrium in plants. Annu Rev Plant Biol 54:357–374. https://doi.org/10.1146/annurev.arplant.54.031902.134907

Fukai S, Cooper M (1995) Development of drought-resistant cultivars using physiomorphological traits in rice. Field Crops Res 40(2):67–86. https://doi.org/10.1016/0378-4290(94)00096-U

Gaffney J, Schussler J, Löffler C, Cai W, Paszkiewicz S, Messina C, Groeteke J, Keaschall J, Cooper M (2015) Industry-scale evaluation of maize hybrids selected for increased yield in drought-stress conditions of the US corn belt. Crop Sci 55(4):1608. https://doi.org/10.2135/cropsci2014.09.0654

García-Ruiz A, Cole JB, VanRaden PM, Wiggans GR, Ruiz-López FJ, van Tassell CP (2016) Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. Proc Natl Acad Sci USA 113(28):E3995–E4004. https://doi.org/10.1073/pnas.1519061113

Garner JB, Douglas ML, Williams SRO, Wales WJ, Marett LC, Nguyen TTT, Reich CM, Hayes BJ (2016) Genomic selection improves heat tolerance in dairy cattle. Sci Rep 6:34114. https://doi.org/10.1038/srep34114

Gaynor RC, Gorjanc G, Bentley AR, Ober ES, Howell P, Jackson R, Mackay IJ, Hickey JM (2017) A two-part strategy for using genomic selection to develop inbred lines. Crop Sci 57(5):2372. https://doi.org/10.2135/cropsci2016.09.0742

Georges M, Charlier C, Hayes B (2018) Genomic selection of livestock and beyond. http://www.giga.uliege.be/upload/docs/application/pdf/2018-02/georges_nrg_2018.pdf

Gibson G (2012) Rare and common variants. Twenty arguments. Nature reviews. Genetics 13(2):135–145. https://doi.org/10.1038/nrg3118

Goddard ME, Hayes BJ (2009) Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nat Rev Genet 10(6):381–391. https://doi.org/10.1038/nrg2575

Gorjanc G, Jenko J, Hearne SJ, Hickey JM (2016) Initiating maize pre-breeding programs using genomic selection to harness polygenic variation from landrace populations. BMC Genom 17:30. https://doi.org/10.1186/s12864-015-2345-z

Gorjanc G, Dumasy J-F, Gonen S, Gaynor RC, Antolin R, Hickey JM (2017) Potential of low-coverage genotyping-by-sequencing and imputation for cost-effective genomic selection in biparental segregating populations. Crop Sci 57(3):1404. https://doi.org/10.2135/cropsci2016.08.0675

Gorjanc G, Gaynor RC, Hickey JM (2018) Optimal cross selection for long-term genetic gain in two-part programs with rapid recurrent genomic selection. Theor Appl Genet. https://doi.org/10.1007/s00122-018-3125-3

Guo M, Rupe MA, Wei J, Winkler C, Goncalves-Butruille M, Weers BP, Cerwick SF, Dieter JA, Duncan KE, Howard RJ, Hou Z, Löffler CM, Cooper M, Simmons CR (2014) Maize ARGOS1 (ZAR1) transgenic alleles increase hybrid maize yield. J Exp Bot 65(1):249–260. https://doi.org/10.1093/jxb/ert370

Haldane JBS (1946) The interaction of nature and nurture. Ann Eugen 13(1):197–205. https://doi.org/10.1111/j.1469-1809.1946.tb02358.x

Hall AJ, Richards RA (2013) Prognosis for genetic improvement of yield potential and water-limited yield of major grain crops. Field Crops Res 143:18–33. https://doi.org/10.1016/j.fcr.2012.05.014

Hammer G, Cooper M, Tardieu F, Welch S, Walsh B, van Eeuwijk F, Chapman S, Podlich D (2006) Models for navigating biological complexity in breeding improved crop plants. Trends Plant Sci 11(12):587–593. https://doi.org/10.1016/j.tplants.2006.10.006

Hammer GL, Dong Z, McLean G, Doherty A, Messina C, Schussler J, Zinselmeier C, Paszkiewicz S, Cooper M (2009) Can changes in canopy and/or root system architecture explain historical maize yield trends in the U.S. corn belt? Crop Sci 49(1):299. https://doi.org/10.2135/cropsci2008.03.0152

Hayes BJ, Bowman PJ, Chamberlain AC, Verbyla K, Goddard ME (2009a) Accuracy of genomic breeding values in multi-breed dairy cattle populations. Genet Sel Evol GSE 41:51. https://doi.org/10.1186/1297-9686-41-51

Hayes BJ, Visscher PM, Goddard ME (2009b) Increased accuracy of artificial selection by using the realized relationship matrix. Genet Res 91(1):47–60. https://doi.org/10.1017/S0016672308009981

Hayes BJ, Daetwyler HD, Goddard ME (2016) Models for genome × environment interaction. Examples in Livestock. Crop Sci 56(5):2251. https://doi.org/10.2135/cropsci2015.07.0451

Hayes BJ, Panozzo J, Walker CK, Choy AL, Kant S, Wong D, Tibbits J, Daetwyler HD, Rochfort S, Hayden MJ, Spangenberg GC (2017) Accelerating wheat breeding for end-use quality with multi-trait genomic predictions incorporating near infrared and nuclear magnetic resonance-derived phenotypes. TAG Theor Appl Genet 130(12):2505–2519. https://doi.org/10.1007/s00122-017-2972-7

Heffner EL, Sorrells ME, Jannink J-L (2009) Genomic selection for crop improvement. Crop Sci 49(1):1. https://doi.org/10.2135/cropsci2008.08.0512

Heffner EL, Jannink J-L, Sorrells ME (2011) Genomic selection accuracy using multifamily prediction models in a wheat breeding program. Plant Genome 4(1):65. https://doi.org/10.3835/plantgenome2010.12.0029

Heslot N, Yang H-P, Sorrells ME, Jannink J-L (2012) Genomic selection in plant breeding. A comparison of models. Crop Sci 52(1):146. https://doi.org/10.2135/cropsci2011.06.0297

Heslot N, Akdemir D, Sorrells ME, Jannink J-L (2014) Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. Theor Appl Genet 127(2):463–480. https://doi.org/10.1007/s00122-013-2231-5

Hickey JM, Chiurugwi T, Mackay I, Powell W (2017) Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery. Nat Genet 49(9):1297–1303. https://doi.org/10.1038/ng.3920

Hill WG, Goddard ME, Visscher PM (2008) Data and theory point to mainly additive genetic variance for complex traits. PLoS Genet 4(2):e1000008. https://doi.org/10.1371/journal.pgen.1000008

Hirsch CN, Springer NM (2018) Weeding out bad alleles. Nat Plants 4(4):193–194. https://doi.org/10.1038/s41477-018-0133-x

Holland JB, Nyquist W, Cervantes-Martínez CT (2003) Estimating and interpreting heritability for plant breeding: and update. Plant Breed Rev 22:9–112

Huang X, Han B (2014) Natural variations and genome-wide association studies in crop plants. Annu Rev Plant Biol 65:531–551. https://doi.org/10.1146/annurev-arplant-050213-035715

Hunt CH, van Eeuwijk FA, Mace ES, Hayes BJ, Jordan DR (2018) Development of genomic prediction in Sorghum. Crop Sci 58(2):690. https://doi.org/10.2135/cropsci2017.08.0469

Jannink J-L, Lorenz AJ, Iwata H (2010) Genomic selection in plant breeding. From theory to practice. Brief Funct Genom 9(2):166–177. https://doi.org/10.1093/bfgp/elq001

Johnsson M, Gaynor RC, Jenko J, Gorjanc G, de Koning D-J, Hickey JM (2018) Removal of alleles by genome editing—RAGE against the deleterious load. BioRxiv. https://doi.org/10.1101/335497

Kirigwi FM, van Ginkel M, Trethowan R, Sears RG, Rajaram S, Paulsen GM (2004) Evaluation of selection strategies for wheat adaptation across water regimes. Euphytica 135(3):361–371. https://doi.org/10.1023/B:EUPH.0000013375.66104.04

Koebner RMD, Summers RW (2003) 21st century wheat breeding. Plot selection or plate detection? Trends Biotechnol 21(2):59–63. https://doi.org/10.1016/s0167-7799(02)00036-7

Kono TJY, Fu F, Mohammadi M, Hoffman PJ, Liu C, Stupar RM, Smith KP, Tiffin P, Fay JC, Morrell PL (2016) The role of deleterious substitutions in crop genomes. Mol Biol Evol 33(9):2307–2317. https://doi.org/10.1093/molbev/msw102

Krause MR, González Pérez L, Crossa J, Pérez-Rodríguez P, Montesinos-López OA, Singh R, Dreisigacker S, Poland J, Rutkoski J, Sorrells ME, Gore MA, Mondal S (2018) Use of hyperspectral reflectance-derived relationship matrices for genomic prediction of grain yield in wheat. BioRxiv. https://doi.org/10.1101/389825

La Fuente GND, Frei UK, Lübberstedt T (2013) Accelerating plant breeding. Trends Plant Sci 18(12):667–672. https://doi.org/10.1016/j.tplants.2013.09.001

Lesk C, Rowhani P, Ramankutty N (2016) Influence of extreme weather disasters on global crop production. Nature 529(7584):84–87. https://doi.org/10.1038/nature16467

Lin Z, Hayes BJ, Daetwyler HD (2014) Genomic selection in crops, trees and forages. A review. Crop Pasture Sci 65(11):1177. https://doi.org/10.1071/CP13363

Lin Z, Cogan NOI, Pembleton LW, Spangenberg GC, Forster JW, Hayes BJ, Daetwyler HD (2016) Genetic gain and inbreeding from genomic selection in a simulated commercial breeding program for perennial ryegrass. Plant Genome. https://doi.org/10.3835/plantgenome2015.06.0046

Löffler CM, Wei J, Fast T, Gogerty J, Langton S, Bergman M, Merrill B, Cooper M (2005) Classification of maize environments using crop simulation and geographic information systems. Crop Sci 45(5):1708. https://doi.org/10.2135/cropsci2004.0370

Longin CFH, Reif JC (2014) Redesigning the exploitation of wheat genetic resources. Trends Plant Sci 19(10):631–636. https://doi.org/10.1016/j.tplants.2014.06.012

Lorenz AJ, Smith KP (2015) Adding genetically distant individuals to training populations reduces genomic prediction accuracy in barley. Crop Sci 55(6):2657. https://doi.org/10.2135/cropsci2014.12.0827

Lorenz AJ, Smith KP, Jannink J-L (2012) Potential and optimization of genomic selection for fusarium head blight resistance in six-row barley. Crop Sci 52(4):1609. https://doi.org/10.2135/cropsci2011.09.0503

Lu D, Akanno EC, Crowley JJ, Schenkel F, Li H, de Pauw M, Moore SS, Wang Z, Li C, Stothard P, Plastow G, Miller SP, Basarab JA (2016) Accuracy of genomic predictions for feed efficiency traits of beef cattle using 50 K and imputed HD genotypes. J Anim Sci 94(4):1342–1353. https://doi.org/10.2527/jas.2015-0126

Ly D, Hamblin M, Rabbi I, Melaku G, Bakare M, Gauch HG, Okechukwu R, Dixon AGO, Kulakow P, Jannink J-L (2013) Relatedness and genotype × environment interaction affect prediction accuracies in genomic selection. A study in Cassava. Crop Sci 53(4):1312. https://doi.org/10.2135/cropsci2012.11.0653

MacLeod IM, Bowman PJ, Vander Jagt CJ, Haile-Mariam M, Kemper KE, Chamberlain AJ, Schrooten C, Hayes BJ, Goddard ME (2016) Exploiting biological priors and sequence variants enhances QTL discovery and genomic prediction of complex traits. BMC Genom 17:144. https://doi.org/10.1186/s12864-016-2443-6

Malosetti M, Bustos-Korts D, Boer MP, van Eeuwijk FA (2016) Predicting responses in multiple environments. Issues in relation to genotype × environment interactions. Crop Sci 56(5):2210. https://doi.org/10.2135/cropsci2015.05.0311

Marjoram P, Zubair A, Nuzhdin SV (2014) Post-GWAS. Where next more samples, more SNPs or more biology? Heredity 112(1):79–88. https://doi.org/10.1038/hdy.2013.52

Messina CD, Podlich D, Dong Z, Samples M, Cooper M (2011) Yield-trait performance landscapes. From theory to application in breeding maize for drought tolerance. J Exp Bot 62(3):855–868. https://doi.org/10.1093/jxb/erq329

Messina CD, Sinclair TR, Hammer GL, Curan D, Thompson J, Oler Z, Gho C, Cooper M (2015) Limited-transpiration trait may increase maize drought tolerance in the US Corn belt. Agron J 107(6):1978. https://doi.org/10.2134/agronj15.0016

Messina CD, Technow F, Tang T, Totir R, Gho C, Cooper M (2018) Leveraging biological insight and environmental variation to improve phenotypic prediction. Integrating crop growth models (CGM) with whole genome prediction (WGP). Eur J Agron. https://doi.org/10.1016/j.eja.2018.01.007

Meuwissen T, Goddard M (2010) Accurate prediction of genetic values for complex traits by whole-genome resequencing. Genetics 185(2):623–631. https://doi.org/10.1534/genetics.110.116590

Meuwissen TH, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157(4):1819–1829

Meuwissen T, Hayes B, Goddard M (2016) Genomic selection. A paradigm shift in animal breeding. Anim Front 6(1):6–14. https://doi.org/10.2527/af.2016-0002

Mezmouk S, Ross-Ibarra J (2014) The pattern and distribution of deleterious mutations in maize. G3 (Bethesda, Md.) 4(1):163–171. https://doi.org/10.1534/g3.113.008870

Podlich DW, Winkler CR, Cooper M (2004) Mapping as you go: an effective approach for marker-assisted selection of complex traits. Crop Sci 44(5):1560–1571

Poland J, Endelman J, Dawson J, Rutkoski J, Wu S, Manes Y, Dreisigacker S, Crossa J, Sánchez-Villeda H, Sorrells M, Jannink J-L (2012) Genomic selection in wheat breeding using genotyping-by-sequencing. Plant Genome 5(3):103. https://doi.org/10.3835/plantgenome2012.06.0006

Qian L, Hickey LT, Stahl A, Werner CR, Hayes B, Snowdon RJ, Voss-Fels KP (2017) Exploring and harnessing haplotype diversity to improve yield stability in crops. Front Plant Sci 8:1534. https://doi.org/10.3389/fpls.2017.01534

Ramu P, Esuma W, Kawuki R, Rabbi IY, Egesi C, Bredeson JV, Bart RS, Verma J, Buckler ES, Lu F (2017) Cassava haplotype map highlights fixation of deleterious mutations during clonal propagation. Nat Genet 49(6):959–963. https://doi.org/10.1038/ng.3845

Rebetzke GJ, Chenu K, Biddulph B, Moeller C, Deery DM, Rattey AR, Bennett D, Barrett-Lennard EG, Mayer JE (2013) A multisite managed environment facility for targeted trait and germplasm phenotyping. Funct Plant Biol 40(1):1. https://doi.org/10.1071/FP12180

Reyes A, Messina CD, Hammer GL, Liu L, van Oosterom E, Lafitte R, Cooper M (2015) Soil water capture trends over 50 years of single-cross maize (Zea mays L.) breeding in the US corn-belt. J Exp Bot 66(22):7339–7346. https://doi.org/10.1093/jxb/erv430

Riedelsheimer C, Czedik-Eysenberg A, Grieder C, Lisec J, Technow F, Sulpice R, Altmann T, Stitt M, Willmitzer L, Melchinger AE (2012) Genomic and metabolic prediction of complex heterotic traits in hybrid maize. Nat Genet 44(2):217–220. https://doi.org/10.1038/ng.1033

Rutkoski J, Benson J, Jia Y, Brown-Guedira G, Jannink J-L, Sorrells M (2012) Evaluation of genomic prediction methods for fusarium head blight resistance in wheat. Plant Genome 5(2):51. https://doi.org/10.3835/plantgenome2012.02.0001

Rutkoski J, Poland J, Mondal S, Autrique E, Pérez LG, Crossa J, Reynolds M, Singh R (2016) Canopy temperature and vegetation indices from high-throughput phenotyping improve accuracy of pedigree and genomic selection for grain yield in wheat. G3 (Bethesda Md.) 6(9):2799–2808. https://doi.org/10.1534/g3.116.032888

Scheben A, Batley J, Edwards D (2017) Genotyping-by-sequencing approaches to characterize crop genomes. Choosing the right tool for the right application. Plant Biotechnol J 15(2):149–161. https://doi.org/10.1111/pbi.12645

Schön CC, Simianer H (2015) Resemblance between two relatives—animal and plant breeding. J Anim Breed Genet Zeitschrift fur Tierzuchtung und Zuchtungsbiologie 132(1):1–2. https://doi.org/10.1111/jbg.12137

Shojaei Saadi HA, Vigneault C, Sargolzaei M, Gagné D, Fournier É, de Montera B, Chesnais J, Blondin P, Robert C (2014) Impact of whole-genome amplification on the reliability of pre-transfer cattle embryo breeding value estimates. BMC Genom 15:889. https://doi.org/10.1186/1471-2164-15-889

Singh AK, Ganapathysubramanian B, Sarkar S, Singh A (2018) Deep learning for plant stress phenotyping. Trends and future perspectives. Trends Plant Sci 23(10):883–898. https://doi.org/10.1016/j.tplants.2018.07.004

Spindel J, Begum H, Akdemir D, Virk P, Collard B, Redoña E, Atlin G, Jannink J-L, McCouch SR (2015) 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(2):e1004982. https://doi.org/10.1371/journal.pgen.1004982

Tanaka R, Iwata H (2018) Bayesian optimization for genomic selection. A method for discovering the best genotype among a large number of candidates. TAG Theoret Appl Genet Theoretische und angewandte Genetik 131(1):93–105. https://doi.org/10.1007/s00122-017-2988-z

Tardieu F, Cabrera-Bosquet L, Pridmore T, Bennett M (2017) Plant phenomics, from sensors to knowledge. Curr Biol CB 27(15):R770–R783. https://doi.org/10.1016/j.cub.2017.05.055

Technow F, Messina CD, Totir LR, Cooper M (2015) Integrating crop growth models with whole genome prediction through approximate bayesian computation. PLoS ONE 10(6):e0130855. https://doi.org/10.1371/journal.pone.0130855

Tester M, Langridge P (2010) Breeding technologies to increase crop production in a changing world. Science (New York NY) 327(5967):818–822. https://doi.org/10.1126/science.1183700

Trethowan RM, Reynolds M, Ortiz-Monasterio KS (2005) Adapting wheat cultivars to resource conserving farming practices and human nutritional needs. Ann Appl Biol 146(4):405–413. https://doi.org/10.1111/j.1744-7348.2005.040137.x

Trick M, Adamski NM, Mugford SG, Jiang C-C, Febrer M, Uauy C (2012) Combining SNP discovery from next-generation sequencing data with bulked segregant analysis (BSA) to fine-map genes in polyploid wheat. BMC Plant Biol 12:14. https://doi.org/10.1186/1471-2229-12-14

van Eeuwijk FA, Bustos-Korts D, Millet EJ, Boer MP, Kruijer W, Thompson A, Malosetti M, Iwata H, Quiroz R, Kuppe C, Muller O, Blazakis KN, Yu K, Tardieu F, Chapman SC (2018) Modelling strategies for assessing and increasing the effectiveness of new phenotyping techniques in plant breeding. Plant Sci 1:1. https://doi.org/10.1016/j.plantsci.2018.06.018

van Oosterom EJ, Yang Z, Zhang F, Deifel KS, Cooper M, Messina CD, Hammer GL (2016) Hybrid variation for root system efficiency in maize. Potential links to drought adaptation. Functional Plant Biol. 43(6):502. https://doi.org/10.1071/fp15308

Voss-Fels KP, Herzog E, Dreisigacker S, Sukurmaran S, Watson A, Frisch M, Hayes BJ, Hickey LT (2018a) SpeedGS to accelerate genetic gain in spring wheat. In: Miedaner T, Korzun V (eds) Applications of genetic and genomic research in cereals, 1st edn. Woodhead Publishing, Cambridge

Voss-Fels KP, Snowdon RJ, Hickey LT (2018b) Designer roots for future crops. Trends Plant Sci. https://doi.org/10.1016/j.tplants.2018.08.004

Wang X, Singh D, Marla S, Morris G, Poland J (2018) Field-based high-throughput phenotyping of plant height in sorghum using different sensing technologies. Plant Methods 14:53. https://doi.org/10.1186/s13007-018-0324-5

Watson A, Ghosh S, Williams MJ, Cuddy WS, Simmonds J, Rey M-D, Asyraf Hatta M, Hinchliffe A, Steed A, Reynolds D, Adamski NM, Breakspear A, Korolev A, Rayner T, Dixon LE, Riaz A, Martin W, Ryan M, Edwards D, Batley J, Raman H, Carter J, Rogers C, Domoney C, Moore G, Harwood W, Nicholson P, Dieters MJ, DeLacy IH, Zhou J, Uauy C, Boden SA, Park RF, Wulff BBH, Hickey LT (2018) Speed breeding is a powerful tool to accelerate crop research and breeding. Nat Plants 4(1):23–29. https://doi.org/10.1038/s41477-017-0083-8

Weber VS, Melchinger AE, Magorokosho C, Makumbi D, Bänziger M, Atlin GN (2012) Efficiency of managed-stress screening of elite maize hybrids under drought and low nitrogen for yield under rainfed conditions in Southern Africa. Crop Sci 52(3):1011. https://doi.org/10.2135/cropsci2011.09.0486

Werner CR, Qian L, Voss-Fels KP, Abbadi A, Leckband G, Frisch M, Snowdon RJ (2018a) Genome-wide regression models considering general and specific combining ability predict hybrid performance in oilseed rape with similar accuracy regardless of trait architecture. Theor Appl Genet 131(2):299–317. https://doi.org/10.1007/s00122-017-3002-5

Werner CR, Voss-Fels KP, Miller CN, Qian W, Hua W, Guan C-Y, Snowdon RJ, Qian L (2018b) Effective genomic selection in a narrow-genepool crop with low-density markers. Asian rapeseed as an example. Plant Genome. https://doi.org/10.3835/plantgenome2017.09.0084

Wolc A, Zhao HH, Arango J, Settar P, Fulton JE, O'Sullivan NP, Preisinger R, Stricker C, Habier D, Fernando RL, Garrick DJ, Lamont SJ, Dekkers JCM (2015) Response and inbreeding from a genomic selection experiment in layer chickens. Genet Sel Evol GSE 47:59. https://doi.org/10.1186/s12711-015-0133-5

Xu Y (2016) Envirotyping for deciphering environmental impacts on crop plants. TAG Theoret Appl Genet Theoretische und angewandte Genetik 129(4):653–673. https://doi.org/10.1007/s00122-016-2691-5

Yabe S, Iwata H, Jannink J-L (2017) A simple package to script and simulate breeding schemes. The breeding scheme language. Crop Sci 57(3):1347. https://doi.org/10.2135/cropsci2016.06.0538

Yang S, Fresnedo-Ramírez J, Wang M, Cote L, Schweitzer P, Barba P, Takacs EM, Clark M, Luby J, Manns DC, Sacks G, Mansfield AK, Londo J, Fennell A, Gadoury D, Reisch B, Cadle-Davidson L, Sun Q (2016) A next-generation marker genotyping platform (AmpSeq) in heterozygous crops A case study for marker-assisted selection in grapevine. Hortic Res 3:16002. https://doi.org/10.1038/hortres.2016.2

Yang J, Mezmouk S, Baumgarten A, Buckler ES, Guill KE, McMullen MD, Mumm RH, Ross-Ibarra J (2017) Incomplete dominance of deleterious alleles contributes substantially to trait variation and heterosis in maize. PLoS Genet 13(9):e1007019. https://doi.org/10.1371/journal.pgen.1007019

Yu X, Li X, Guo T, Zhu C, Wu Y, Mitchell SE, Roozeboom KL, Wang D, Wang ML, Pederson GA, Tesso TT, Schnable PS, Bernardo R, Yu J (2016) Genomic prediction contributing to a promising global strategy to turbocharge gene banks. Nat Plants 2:16150. https://doi.org/10.1038/nplants.2016.150

Zhao Y, Gowda M, Liu W, Würschum T, Maurer HP, Longin FH, Ranc N, Reif JC (2012) Accuracy of genomic selection in European maize elite breeding populations. Theor Appl Genet 124(4):769–776. https://doi.org/10.1007/s00122-011-1745-y

Zhong S, Dekkers JCM, Fernando RL, Jannink J-L (2009) Factors affecting accuracy from genomic selection in populations derived from multiple inbred lines. A Barley case study. Genetics 182(1):355–364. https://doi.org/10.1534/genetics.108.098277