

Ideas in Genomic Selection that Transformed Plant Molecular Breeding: A Review

Matthew McGowan
Molecular Plant Sciences Program
Washington State University
Pullman, WA 99164, USA

Jiabo Wang
Institute of Qinghai-Tibetan Plateau
Southwest Minzu University
Chengdu, Sichuan, China

Haixiao Dong
College of Plant Sciences
Jilin University
Changchun, China

Xiaolei Liu
Key Laboratory of Agricultural Animal Genetics, Breeding, and Reproduction
Ministry of Education
College of Animal Science and Technology
Huazhong Agricultural University
Wuhan, China

Yi Jia
Research and Development
Corteva Agriscience
Indianapolis, IN 46268, USA

Xianfeng Wang
Department of Crop Genomics and Bioinformatics
College of Agronomy and Biotechnology
National Maize Improvement Center of China
China Agricultural University
Beijing, China

Hiroyoshi Iwata
Department of Agricultural and Environmental Biology
University of Tokyo
Tokyo, Japan

Yutao Li
CSIRO Agriculture and Food
St. Lucia, Queensland, Australia

Alexander E. Lipka
Department of Crop Sciences
University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA

Zhiwu Zhang
Department of Crop and Soil Sciences
Washington State University
Pullman, WA 99164, USA

ABSTRACT

Estimation of breeding values through Best Linear Unbiased Prediction (BLUP) using pedigree-based kinship and Marker-Assisted Selection (MAS) are the two fundamental breeding methods used before and after the introduction of genetic markers, respectively. The emergence of high-density genome-wide markers has led to the development of two parallel series of approaches inspired by BLUP and MAS, which are collectively referred to as Genomic Selection (GS). The first series of GS methods alters pedigree-based BLUP by replacing pedigree-based kinship with marker-based kinship in a variety of ways, including weighting markers by their effects in genome-wide association study (GWAS), joining both pedigree and marker-based kinship together in a single-step BLUP, and substituting individuals with groups in a compressed BLUP. The second series of GS methods estimates the effects for all genetic markers simultaneously. For the second series methods, the marker effects are summed together regardless of their individual significance. Instead of fitting individuals as random effects like in the BLUP series, the second series fits markers as random effects. Differing assumptions regarding the underlying distribution of these marker effects have resulted in the development of many Bayesian-based GS methods. This review highlights critical concept developments for both of these series and explores ongoing GS developments in machine learning, multiple trait selection, and adaptation for hybrid breeding. Furthermore, considering the increasing use and variety of GS methods in plant breeding programs, this review addresses important concerns for future GS development and application, such as the use of GWAS-assisted GS, the long-term effectiveness of GS methods, and the valid assessment of prediction accuracy.

KEYWORDS: Plant breeding, genomic selection, Bayes, BLUP, machine learning

OUTLINE:

- I. INTRODUCTION
- II. BLUP ALPHABET
- III. BAYESIAN ALPHABET
- IV. MACHINE LEARNING
- V. GWAS-ASSISTED GENOMIC SELECTION
- VI. HYBRID BREEDING
- VII. MULTIPLE TRAITS
- VIII. LONG-TERM SELECTION
- IX. ASSESSMENT OF PREDICTION ACCURACY
- X. GS TRANSFORMED PLANT BREEDING
- XI. PROSPECTS

ABBREVIATIONS

BLUP	Best linear unbiased prediction
cBLUP	Compressed BLUP
CNN	Convolutional neural network
EMMA	Efficient mixed-model association
gBLUP	Genomic BLUP
GBM	Gradient boosting machine
GCA	General combining ability
GOCS	Genomic optimal contribution selection
GS	Genomic selection
GWAS	Genome-wide association study
MAS	Marker-assisted selection
ML	Machine learning
MLM	Mixed linear model
MPH	Mid-parent heterosis
MT-GS	Multi-trait genomic selection
OCS	Optimal contribution selection
RFLP	Restriction fragment length polymorphism
rrBLUP	Ridge-regression BLUP
SCA	Specific combining ability
ssBLUP	Single-step BLUP
SVR	Support vector regression

I. INTRODUCTION

Globally, crops represent 55% of calories and 40% of protein for human consumption, and 36% of calories and 53% of protein for animal feed (Cassidy et al. 2013). Many crops are also sources for industrial material and energy supply, such as maize for ethanol. The inevitable growth of the world's population demands a 50% increase in food production over the next three decades. As evidenced by past achievements, the development of crop varieties through breeding will continue to play a critical role in boosting crop production.

Before applications using genetic markers, traditional plant breeding relied on phenotyping to identify superior genetic lines. The introduction of genetic markers allowed breeders to partially eliminate phenotyping and conduct selection at an early stage with high selection accuracy. One such application is Marker-Assisted Selection (MAS), which is effective for simple traits controlled by major genes. For complex traits controlled by multiple genes, assessment of breeding values of genetic lines has to incorporate genetic markers spanning the entire genome, termed Genomic Selection (GS). In 1994, Rex Bernardo achieved an early milestone for the development of GS by replacing pedigree-based kinship (Henderson 1975b, 1975a, 1976a) with a marker-based kinship estimated using RFLP (Restriction Fragment Length Polymorphism) markers spanning the entire maize genome (Bernardo 1994). Instead of using only the markers associated with traits of interest, as in MAS, GS uses all the genomic markers to define the kinship among individuals. This marker-based kinship has the advantage of capturing Mendelian sampling errors and segregation distortion that were not able to be identified using pedigree-based kinship. As Best Linear Unbiased Prediction (BLUP) is used to evaluate the genomic value of recombinant inbred lines, the method has come to be known as genomic BLUP, or gBLUP (VanRaden 2008; Zhang et al. 2007).

Compared with MAS, which identifies genetic markers that co-segregate with causal genes and have statistically significant effects on traits of interest, GS uses all the available markers across a genome, regardless of whether they have significant effects or not. Inspired by MAS, the second milestone of GS was introduced by Meuwissen et al. in 2001, which first estimates the effects of all markers in a parental generation, then sums them together to predict performance of future progeny. These markers are treated as random effects. Although it is commonly understood that the random effects follow normal distributions, a variety of assumptions on the prior variance

distribution have prompted the development of different methods by which the Bayesian theorem has been used, resulting in a series of Bayesian alphabet-named methods. These methods include Bayes A, where markers are assumed to have different variances, and Bayes B, where a certain proportion of markers are restricted to have no effect. Similar to the Bayesian alphabet series, kinship-based BLUP methods have also evolved into their own BLUP alphabet series, including combining marker-based kinship and pedigree-based kinship in a single-step BLUP (ssBLUP) and predicting individuals' genetic effects based on their corresponding groups using kinship compression (compressed BLUP, or cBLUP).

The development of GS methods and their application remains one of the most active areas in genomic research. These developments have been reviewed multiple times over the last 25 years. The foci of these reviews include: 1) implementation procedure (Calus 2010); 2) training-testing population relationship, or marker properties (Goddard and Hayes 2007); 3) computational challenges (Mukhopadhyay and Kumar 2008; Nakaya and Isobe 2017; Tempelman 2015; Tier 2010); and 4) species or production specific, including plant (Crossa et al. 2017; Desta and Ortiz 2014; Heffner et al. 2009; Heslot et al. 2015; Jannink et al. 2010; Lorenz et al. 2011), animal (Hayes and Goddard 2010; Mandal et al. 2013; Meuwissen et al. 2013; Robertsen et al. 2019; Voss-Fels et al. 2019), forage (Maria et al. 2014), dairy (Hayes et al. 2009), and wheat (Larkin et al. 2019).

This review highlights the critical ideas that have transformed plant molecular breeding. First, we review the method developments to improve selection accuracy, including the first implementation of gBLUP in maize, the expansion from MAS to GS Bayesian methods, the development of machine learning methods, and the usage of Genome-Wide Association Studies (GWAS) for GS. Secondly, we review three factors that influence selection responses, including hybrid breeding, enhancement using multiple traits, and preservation of variation for long-term selection. Finally, we review the pitfalls of the assessment of prediction accuracy and discuss the transformation of plant breeding as a result of GS.

II. BLUP ALPHABET

The mixed linear model (MLM) equation invented by Charles Henderson in 1976 (Henderson 1976) combines random and fixed effects in a single equation to simultaneously calculate best linear unbiased predictions (BLUPs) for random effects and best linear unbiased estimates for fixed effects. Compared to the general linear model, which contains only fixed effects except for the residual, the MLM contains both the random residual and random animal effects. In many cases, the levels of random effects are equal to or larger than the number of observations. To avoid the overfitting problem, the MLM equation was built on the optimization of likelihood, instead of the least squared error used in general linear model. The maximum likelihood approach also allows the random effects to have a variance structure. In the case where random variables represent individual additive genetic effects, the kinship among individuals can be derived from pedigrees using the definition developed by Sewall Wright (Wright 1922). In 1976, Henderson further developed an algorithm to build the inverse of the numerator relationship matrix (A^{-1}) directly from a pedigree (Henderson 1976) in computing time that is linear to the number of individuals (Figure 1).

	$y = u + e$	
$Var(u) = K_p \sigma_u^2$ and K_p is pedigree-based kinship	Henderson	1960s
Directly derive inverse of K_p from pedigree	Henderson	1970s
$Var(u) = K_M \sigma_u^2$ and K_M is marker-based kinship	Bernado	1994
$u = \sum s_i$, $s_i \sim N(0, \sigma_i^2)$	Meuwissen et al.	2001
K_M replaces K_p in MTDFREML	Zhang et al.	2006
Efficient K_M algorithm for genomic BLUP (gBLUP)	VanRaden	2008
GWAS weighted K_M algorithm	Zhang et al.	2011
<i>Ridge Regression & gBLUP implemented in R (rrBLUP)</i>	Endelman	2011
K_M and K_p integration (single-step BLUP)	Misztal	2011
Build K_M with GWAS associated markers (SUPER BLUP)	Wang et al.	2019
Compress individuals to groups in K_M (Compressed BLUP)	Wang et al.	2019

Figure 1. Best Linear Unbiased Prediction and its deviations for genomic selection. Individual phenotypes (y) are decomposed into their genetic effects (u) and residuals (e). Under decomposition of phenotypes (y) of individuals into genetic effects (u) and residuals (e), Dr. Henderson developed the mixed linear model equation in the 1960s to directly solve for u as the Best Linear Unbiased Prediction of breeding values by using the pedigree-based kinship to define the variance structure of the genetic effects. The pedigree-based kinship was substituted with the marker-based kinship by Dr. Bernardo in 1994. Individuals' genetic effects were predicted indirectly by summing all the marker effects (s_i) following normal distribution with mean of zero and variance of σ_i^2 for $i=1$ to n covering the genome (Meuwissen et. al., 2001). Since then, the direct BLUP methods have been evolved into multiple formats, including the prediction of individuals' genetic effects based on their group effects (Wang et.al., 2019).

The MLM can also contain random effects for individuals that do not have observations, which makes it possible to evaluate the genetic merit of bulls from their daughters in animal breeding. Before genomic selection, the BLUP method, in conjunction with pedigree, served as the primary method in animal breeding, including dairy and beef cattle, sheep, chicken, and pig. Many free software packages have been developed, e.g., VCE (Groeneveld 1994), MTDFREML, DMU, ASReml, and BLUPF90 (Miszta et al. 2002). Since crops rarely perform differently due to sex, and more importantly, because plant breeding relies on the selection among siblings which cannot be differentiated by pedigree, BLUP was not used as much in crops as in livestock until genetic markers covering the entire genome became available. In 1994, Bernardo used genome-wide RFLP markers to define kinship among maize lines to evaluate their genetic merits using BLUP (Bernardo 1994). Marker-based kinship has advantages over pedigree-based kinship, because it captures the realized kinship, which includes factors that are not part of pedigree-based kinship, including Mendelian sampling error and segregation distortion. However, similarities between the two approaches have allowed many software packages originally developed for pedigree-based kinship to be extended to include marker-based kinship with minor modifications (**Figure 2**). In 2007, marker-based kinship was added into MTDFREML as the second way to define kinship among individuals (Zhang et al. 2007). A more efficient algorithm was developed (VanRaden 2008) by multiplying the genotype matrix and its transpose and adjusting for allele frequency:

$$G = \frac{WW'}{2 \sum P_i (1-P_i)} \quad (1)$$

where G is a marker-based kinship (genomic relationship matrix), W is a matrix of centralized genotypes, and allele frequency is noted as P_i . In 2011, gBLUP with the efficient kinship algorithm was implemented in an R package, rrBLUP (Endelman 2011).

$$\begin{array}{c}
 \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \dots \\ y_n \end{pmatrix} = \begin{matrix} \text{[Black Box]} \end{matrix} + \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ \dots \\ e_n \end{pmatrix} \\
 \mathbf{y} \quad \quad \quad \mathbf{e}
 \end{array}$$

$$\begin{array}{c}
 \begin{pmatrix} 1 & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & \dots & & \\ & & & & 1 & \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ \dots \\ u_n \end{pmatrix} & \begin{pmatrix} 0 & 2 & 1 & \dots & \dots & 2 \\ 1 & 1 & 2 & \dots & \dots & 1 \\ 2 & 0 & 1 & \dots & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 2 & 0 & \dots & \dots & 1 \end{pmatrix} \begin{pmatrix} g_1 \\ g_2 \\ g_3 \\ \dots \\ \dots \\ g_p \end{pmatrix} \\
 \mathbf{Z} \quad \mathbf{u} & \mathbf{M} \quad \mathbf{g}
 \end{array}$$

$$\mathbf{y} = \mathbf{Z}\mathbf{u} + \mathbf{e} \qquad \mathbf{y} = \mathbf{M}\mathbf{g} + \mathbf{e}$$

Direct estimates: $\hat{\mathbf{u}}$ Indirect estimates: $\mathbf{M}\hat{\mathbf{g}}$

$$\mathbf{u} \sim N(0, K \sigma_u^2), \hat{\mathbf{u}} = \mathbf{M}\hat{\mathbf{g}} \text{ if } K = \mathbf{M}'\mathbf{M} \text{ and } \mathbf{g} \sim N(0, I \sigma_g^2)$$

Figure 2. Direct and indirect estimation of genomic breeding values. When n individuals' phenotypes (y) are decomposed into a black box and residuals (e), the black box has to be parameterized to make it most likely to have the observed phenotypes for the given parameters. The direct approaches specify a genetic variance (σ_u^2) to define the distribution of the individuals' genetic effects by incorporating kinship (K) derived from pedigree or genetic markers. The indirect approaches estimate the effects of genetic markers (g) first and then sum them ($M\hat{g}$) to estimate genomic breeding values indirectly (note that the symbols with and without hat represent the estimate and the true effect, respectively). When K is defined as $M'M$ and marker effects have

independent normal distribution with mean of zero and a variance (σ_g^2), the direct and indirect approaches give the same estimates of genomic breeding values.

Since it is often not possible to genotype every individual, especially ancestral individuals, a single-step BLUP (or ssBLUP) method was developed that uses both marker-based kinship and pedigree-based kinship simultaneously (Aguilar et al. 2010; Christensen and Lund 2010; Legarra et al. 2009). The inverse of the joint kinship (H) is estimated by using an arbitrary weight (ω) between 0 and 1:

$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & (G^{-1} - \omega A_{22}^{-1}) \end{bmatrix} \quad (2)$$

where A_{22} is the pedigree relationship matrix among genotyped animals. Studies demonstrated that this method had better prediction accuracy than BLUP using pedigree alone (Chen et al. 2011). Since then, ssBLUP has dominated the application of genomic selection in animal breeding (Campos et al. 2018; Zhang et al. 2016).

Before genomic BLUP gained popularity, a different approach developed by Meuwissen et al. in 2001 boosted the usage of all genetic markers, as opposed to MAS, which uses only the statistically significant markers. The 2001 study fit all markers as a random effect and summed them together to predict the genetic merit of individuals. These marker effects were assumed to follow a normal distribution with zero mean and variances. When the variances are assumed to be a constant, the model become a ridge regression. When the variances are assumed to be random variables that follow an inverse Chi-square distribution, the model has to be solved using Bayesian methods. It was demonstrated that ridge regression gives a prediction identical to the gBLUP when the efficient kinship algorithm is used (VanRaden 2008).

Since the random effects in MLM are specific for a trait of interest, their variance structure should also be trait specific. With pedigree kinship, the specification is limited; however, marker-based kinship can be specific for a trait of interest. Instead of having all the genetic markers contribute equally to the marker-based kinship, these markers can be weighted for a particular trait based on their effects estimated by GWAS. This modified marker-based kinship method was named TABLUP (Zhang et al. 2010, 2011). After the introduction of the SUPER GWAS method (Wang et al. 2014), TABLUP was further specialized for traits controlled by a small number of genes. In this specialization, kinship is built using only the associated markers determined by the maximum likelihood method. Consequently, the method for genomic selection was named SUPER

BLUP, or sBLUP. Both real data and simulations demonstrated that SUPER BLUP was superior to regular gBLUP for traits controlled by a small number of genes (Tang et al. 2016; Jiabo Wang et al. 2018).

Another modified marker-based kinship method considers hierarchical population structures that involve clustered groups of closely related individuals. In this case, individuals can cluster into a closely related group, which is at a greater distance from another group. For traits with low heritability, the phenotypic deviation of an individual from the population mean may be dominated by error rather than by the genetic difference. However, the phenotypic average across individuals in a group can be much closer to the genetic average of the group than the individual phenotype is to the individual genetic value. This group genetic value can be a better representation of individual genetic value than the prediction of individual genetic value based on individual phenotype. Simulations demonstrated that GWAS using groups as random effects for background control identified more causal mutations than when using individuals as random effects in an MLM (Zhang et al. 2010). Since individuals are compressed into groups for GWAS, the method was named compressed MLM. For GS, the method was similarly called compressed BLUP (cBLUP). Tests with both real and simulated data have demonstrated that cBLUP is superior to gBLUP and sBLUP for traits with low heritabilities (Wang et al. 2018) and populations with this hierarchical structure.

There is only one random effect in addition to the residual effect in all BLUP methods reviewed above, including the traditional BLUP based on pedigree relationships, regular gBLUP based on genomic relationships derived from all genetic markers, single-step BLUP combining pedigree-based and marker-based kinships together, TABLUP with markers weighted by GWAS, SUPER BLUP with kinship derived from associated markers only, and compressed BLUP with individuals predicted by their corresponding groups. In such cases, likelihood is a function of the ratio (λ) between residual variance and the genetic variance using the efficient mixed-model association (EMMA) algorithm (Kang et al. 2008). The ratio can be expressed as a function of heritability: $\lambda = \frac{1-h^2}{h^2}$. Interval search can be efficiently used to find the heritability that maximizes the likelihood by varying heritability from 0 to 1 with limited intervals. The EMMA algorithm has been implemented in rrBLUP for gBLUP (Endelman 2011), and in GAPIT for gBLUP, SUPER BLUP, and compressed BLUP (Tang et al. 2016).

III. BAYESIAN ALPHABET

When available markers cover an entire genome, the marker based-kinship can be transformed in a way that is not feasible for pedigree-based kinship. The total additive genetic effects of a trait can be decomposed into the sum of individual marker effects (**Figure 2**). The milestone research by Meuwissen et al. documented this decomposition (Meuwissen et al. 2001). To solve the “large p, small n” problem (i.e., more parameters than observations), genetic markers were treated as random effects. If these markers are fitted as fixed effects, it causes overfitting in the fixed-effect model. These marker effects were assumed to follow a normal distribution with zero mean and an unknown variance. In the 2001 study, one approach was to set the unknown variance as a constant, which turned the genomic prediction problem into a well-established statistical problem known as ridge regression (**Figure 3**).

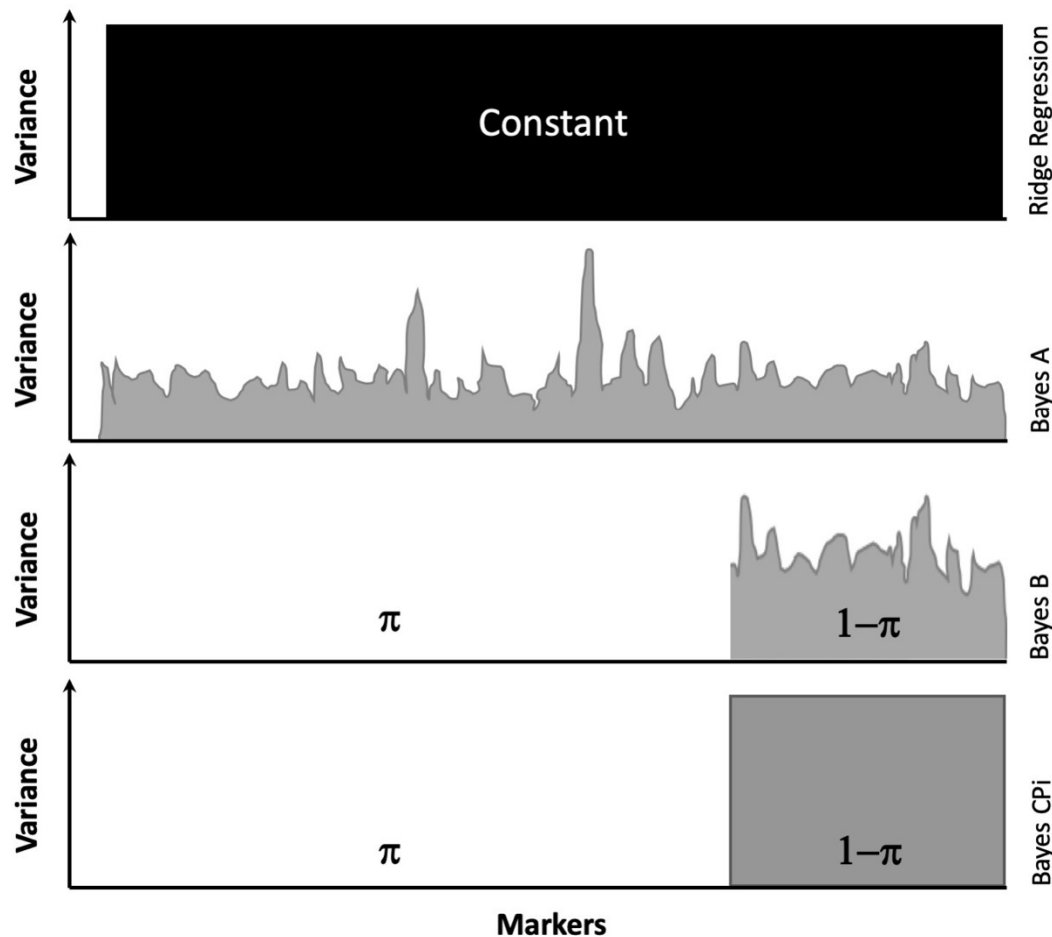


Figure 3. The graphic properties of the variance parameter for the distribution of marker effects. Markers are fitted simultaneously as random effects following normal distribution with mean of zero and variances that vary across methods. Ridge regression uses a constant (black) for all the variances. The variances are different across all markers and follow a prior distribution (gray) in Bayes A method. Bayes A turns to Bayes B when some of the variances are restricted to be zero for part of marker (π). Bayes B turns to Bayes Cpi when the non-zero variances are restricted to be the same.

In reality, it is rarely true that all genetic marker effects follow the same distribution, as assumed by the ridge regression method. On the other hand, treating these unknown variances as random variables also needs restriction to eliminate potential model overfitting. The Bayesian approach effectively solves this problem by setting the distribution hierarchy (Tipping and Faul 2003). For example, each marker effect can follow a different normal distribution with its variance

parameter. These variances are then required to follow a single inverse Chi-square distribution, governed by the hyperparameter prior distribution. This restriction for the prior distribution of variances was named Bayes A (**Figure 3**).

The Bayesian approach can also accommodate a variety of different prior distributions. Ridge regression has the highest restrictions on both marker effects and hyperparameters. All the marker effects follow the same normal distribution. The variance is a constant with a flat prior distribution domain from negative infinity to positive infinity. In contrast, Bayes A has the least restriction. Each marker has its own distribution with a distinct variance. Bayes A can be modified to restrict a proportion (π) of markers to have no effect. This modified method is termed Bayes B. Similarly, Bayes B can be further modified to create another method ($C\pi$) by restricting all non-zero marker effects to have the same normal distribution with variances following an inverse Chi-square distribution (**Figure 3**). Comparisons were conducted for the Bayesian methods developed between 2001 and 2011 (Kärkkäinen and Sillanpää 2012), including Bayesian Lasso, in which the inverse Chi-square distribution from $C\pi$ is changed to a Laplace distribution (double exponential).

For ridge regression, there is only one random effect besides the residual. The variance components can be explicitly estimated using algorithms such as EMMA (Kang et al. 2008). Bayesian methods have more parameters than ridge regression. In practice, the parameter estimation is conducted by sampling. Let y denote the data (e.g., phenotype and genotype), θ denotes the parameters (e.g., marker effects and variance parameters of marker effect distribution), $f(y)$ denotes the probability density of y (constant), $f(\theta)$ denotes the probability density for θ (prior), $f(y|\theta)$ denotes the likelihood of y given θ , $f(\theta|y)$ denotes the probability density of θ given y (posterior). With enough samples of the posterior, parameters for the posterior distribution can be inferred by using sample means. The problem is that the posterior distribution is hard to get, but Bayesian transformation can be used to transform a hard problem to an easier problem:

$$f(\theta|y) = \frac{f(y|\theta)f(\theta)}{f(y)} \propto f(y|\theta)f(\theta) \quad (3)$$

The equation transforms posterior $f(\theta|y)$ into the product of an easy-to-solve probability density of y given θ and a prior distribution which represents a belief. For joint distribution of multiple parameters, Gibbs sampling (García-Cortés and Sorensen 1996) is used for approximation using marginal distribution under Markov Chain Monte Carlo methods.

A challenge of using Bayesian approaches is the computing time due to the amount of sampling required to reach convergence. Several R packages have been developed, including BGLR (Pérez and De Los Campos 2004), BayesR (<https://github.com/syntheke/bayesR>), MCMCglmm (Hadfield 2010), LaplacesDemon (Hall 2018), and BMTME (Montesinos-López et al. 2019). Of these packages, BGLR incorporates multiple methods, including BayesA, BayesB, BayesC, Bayesian Lasso, and Bayesian Ridge Regression.

IV. MACHINE LEARNING

The genomic BLUP method introduced by Bernardo in 1994 can also be interpreted as a special case of machine learning. The kinship matrix is termed “kernel” in the field of engineering. A kinship matrix K can be multiplied by itself to derive a new kernel $K^2=K'K$. The process can be iterated until there are no further changes with multiplication. During the matrix multiplication process toward equilibrium, the optimum level of multiplication can be determined by the typical training process of machine learning, i.e., by dividing the entire population into training and testing populations. The recent application of machine learning in genomic prediction has expanded to include many domains of machine learning, including neural network or deep learning, which has been compared with other machine learning methods in crops like wheat. Only a few methods of machine learning for genomic prediction, including random forest, gradient descent, and deep learning, will be reviewed here.

Principles of Supervised and Deep Learning Methods

Machine learning (ML) is a domain of computer science that uses algorithms to gain the ability to automatically learn and improve from experience without being explicitly programmed (Samuel 1959). Some of the main advantages of ML methods over conventional statistical methods in large genomic data analysis include the following: (1) they have the ability to deal with “large p , small n ” problems; (2) they are black-box approaches that do not require any prior knowledge about distributions of responsible variables or underlying genetic models affecting traits of interest; (3) they can take multiple interactions or correlations among features into account; and (4) they can provide high prediction accuracy due to both training and validation procedures built into the processes of individual algorithms that allow users to predefine training and validation datasets, or allow ML to apply a randomly assigned cross-validation approach to a large population for

prediction of individual phenotypes (Chen and Ishwaran 2012; González-Recio et al. 2014; Lubke et al. 2013).

ML can be divided into supervised and unsupervised methods. Supervised learning involves using labelled input training data (e.g., SNPs or sequence variants and phenotypes) to learn a prediction function either for regression or classification of a response variable, then examining the prediction error of the derived function using validation data. A prediction error consists of two parts: a variance and a bias. The goal of a learning algorithm is to minimize both bias and variance. Examples of supervised learning methods include Random Forest (Breiman 2001) and Gradient Boosting Machine (GBM) (Friedman 2001). Both Random Forest and GBM are decision tree-based ensemble methods. The main difference between Random Forest and GBM is that decision trees in Random Forest are independently generated with replacement and prediction error of an individual feature (e.g., a SNP) indicated by a variable importance measurement value, which is calculated by averaging prediction errors of all the decision trees that contain the specific feature. GBM proceeds through a step-wise assembly of many “weak learners” (e.g., small subsets of SNPs) to build a predictive model, i.e., the output of the later trees relies heavily on the results from previously sampled trees (Li et al. 2018).

Unsupervised learning involves discovering existing patterns (e.g., groups or associations) among input data variables themselves without information from an output response variable (Hastie et al. 2009). Examples of unsupervised learning methods include principal component analysis and cluster analysis, two popular pattern identification methods that are used widely for high dimensionality omics data analysis. Unsupervised learning can provide features for GS but cannot directly be used for GS. For example, principal components derived from genetic markers are important features to predict genetic potential (Dong et al. 2018).

Deep learning is another family of ML methods that can be used for both supervised and unsupervised learning. It derives representative information from input data by using a hierarchy of multiple layers of units (neurons). Each neuron computes a weighted sum of its inputs, and the weighted sum is passed through a non-linear function (Montesinos-López et al. 2019). Each layer then transforms input data into increasingly abstract representations. Finding optimal weights for this neural network using input data is a goal of deep learning. Some of the popularly applied neural networks include densely connected networks, convolutional neural networks (CNNs), and recurrent neural networks (Montesinos-López et al. 2019). ML algorithms that perform supervised

learning are generally useful for genomic prediction (Khaki and Wang 2019; Miguel and Zingaretti 2019), while in rare cases, the classification and features from unsupervised learning can be directly used for GS.

Application of Supervised Machine Learning in Genomic Prediction

To date, studies using supervised ML methods and large SNP chips in genomic prediction have produced mixed results in livestock and plant species, depending on several factors. Some of these include presence or absence of population structure, additive and epistatic nature of phenotypes, ML methods used, data quality (e.g., the amount of missing data), and representation of training and validation populations (Azodi et al. 2019; González-camacho et al. 2018; Gonzalez-Recio et al. 2014; Grinberg et al. 2018; Howard et al. 2014; Montesinos-López et al. 2019; Waldmann 2016). For example, an extensive study systematically evaluated the genomic prediction accuracy of 6 linear and 6 non-linear algorithms with the data from 18 traits across 6 plant species with different marker densities and training population sizes. No overall winner was found when individual methods were compared (Azodi et al. 2019). However, the prediction using a combination of multiple ML methods did perform better than conventional linear methods. The six evaluated linear methods were BLUP (rrBLUP), Bayes A (BA), Bayes B (BB), Bayesian Lasso, Bayesian-RR, and support vector regression (SVR) with a linear kernel (SVRlin). The six non-linear ML methods included SVR with polynomial kernel (SVRpoly), SVR with radial basis function kernel (SVRrbf), gradient tree boosting, artificial neural networks, and convolutional neural networks (CNN). In another study with the datasets from yeast, wheat, and rice, six machine learning methods (elastic net, rrBLUP, LASSO regression, RF, GBM, and SVR) were compared with two conventional statistical methods (gBLUP and a two-step sequential method). The study concluded that “for almost all the phenotypes considered, standard machine learning methods outperformed the methods from classical statistical genetics” (Grinberg et al. 2019). Regarding individual methods, GBM performed very well for traits with greater complexity. In the presence of population structure, gBLUP was the best. However, in the presence of missing data, Random Forest was the most robust (Grinberg et al. 2019).

General consensus is that for moderate to high heritability traits that are driven by additive genetic models, there is no additional benefit to using machine learning methods for genomic prediction in comparison to conventional gBLUP or Bayesian methods (Azodi et al. 2019;

González-Recio et al. 2014; Grinberg et al. 2018). However, in the presence of additive and epistatic effects, machine learning methods outperform the gBLUP method (Grinberg et al. 2019; Montesinos-López, Martín-Vallejo, et al. 2019; Waldmann 2016).

Application of Deep Learning in Genomic Prediction

Early studies using shallow neural networks (e.g., single-layer networks) for nonparametric prediction of complex traits in plant and animal breeding suggested that neural networks could give high prediction accuracy (Gianola et al. 2011; González-Camacho et al. 2012; González-Camacho et al. 2016; Pérez-Rodríguez et al. 2012). However, these studies were limited by small sample sizes and number of SNPs. Performances of genomic prediction were compared using two deep learning approaches (Multilayer Perceptrons and CNNs) and Bayesian linear regression using significantly larger datasets consisting of five phenotypic measurements ($h^2 \sim 0.2-0.7$) and the genotypes of 500K SNPs from 100K distantly related European white Caucasian individuals (Bellot et al. 2018; Raymond et al. 2018). They found that the performance of multilayer perceptrons varied greatly when using different densities of SNP data or phenotypes, but the performance of CNNs was similar to Bayesian linear methods. Another study tested nine publicly available datasets with three deep learning methods (connected networks, CNNs, and recurrent neural networks). It was found that gBLUP models gave the best predictions when G*E interactions were considered in the models (Montesinos-lópez et al. 2018). However, when the interactions were ignored in both deep learning and gBLUP, the deep learning methods performed better than gBLUP due to their ability to handle large sets of correlated variables. Using wheat datasets from The International Maize and Wheat Improvement Center (CIMMYT), the performance of multi-trait deep learning models was compared with single-trait deep learning models on different types of phenotypes, including binary, ordinal, or continuous traits. Marginal gain was confirmed for multiple-trait deep learning models for continuous traits, but no difference was observed for binary or ordinal traits (Montesinos-lópez et al. 2018).

Challenges and Opportunities for Future ML Applications in Breeding Programs

Genomic selection is revolutionizing genetic improvement programs. However, it still has difficulty predicting hard-to-measure and low-heritability traits (e.g., disease resistance traits). The digital revolution and high-throughput genomic technologies provide a new era for future smart

breeding programs that require drawing real-time large data information from biological (e.g., genomics), environmental (including climate records), and management sources to accurately predict future individual performance. A great challenge here is how to integrate multiple sources of data, especially complex genetics-by-environment-by-management (GxExM) interactions, into an efficient prediction model. Systematic approaches that combine multiple types of omics information with environment and other sources are needed to predict complex phenotypes (Misra et al. 2018; Sperschneider 2019). Machine learning methods (especially deep learning) have the ability to deal with complex phenotypes. Common pitfalls of applying ML methods in genomic prediction include crucial but non-trivial hyperparameter selection; determination of training, validation, and testing datasets to be used and ML methods to be applied; feature selection and dimension reduction; and dealing with overfitting and hard-to-interpret results (Crossa et al. 2019; de los Campos et al. 2018; Sperschneider 2019). However, a few studies have demonstrated that combining multiple ML methods (González-camacho et al. 2018; Grinberg et al. 2019), applying domain knowledge (Li et al. 2018; Li et al. 2019), and taking population structure into account (Grinberg et al. 2019) can improve genomic prediction accuracy of complex phenotypes.

V. GWAS-ASSISTED GENOMIC SELECTION

Molecular breeding started from single-marker applications and MAS still plays a key role in practice. Using combined validated and newly discovered markers has the potential to boost the prediction accuracy of genomic selection. Both theoretical and practical studies have been conducted to investigate the impact of incorporating these markers.

One of the key factors contributing to the success of GS is its use of genome-wide markers to capture the signals of both large- and small-effect causal mutations underlying traits (Lipka et al. 2015). Since GS methods can also be used for GWAS, the top-ranked markers from GWAS have been subsequently used for GS in two steps of ridge regression (Resende et al. 2012). As sequencing technology continues to improve, the likelihood of these causal mutations being included in the GS model increases. The direct benefit is to retain prediction accuracy over generations, because linkage disequilibrium breakdown between closely linked markers and causal mutations should theoretically become less frequent (Weller and Ron 2010). The indirect benefit is the increased number of validated mutations that provide options for conventional GS models.

In addition to integrating GWAS and GS in one process, such as in SUPER BLUP, there are several ways to conduct GWAS first and then merge the GWAS results into the GS process.

Various Approaches to Incorporate Validated Mutations

Considering the improvements in sequencing technology, various approaches for the specific incorporation of causal mutations into GS models have been explored (**Figure 4**). First, associated markers were incorporated as fixed effects (**Figure 4d**). A simulation study demonstrated that it improved prediction accuracy in the RR-BLUP model to include quantitative trait nucleotides, each explaining at least 10% of the genetic variance (Bernardo 2014). Real traits also demonstrated the same trend (Odilbekov et al. 2019).

Second, the standard GS model was extended to include an additional random effect for the causal mutations (Brøndum et al. 2015). The additional random effect has its own kernel derived from the markers at the peaks in a genome-wide association study (**Figure 4e**).

Third, the kernel, the genomic relationship matrix in the conventional GS model, was modified to put more weight on causal markers (Fragomeni et al. 2017). Such a genomic relationship matrix was derived from a weighting approach, where causal mutations underlying a trait received a stronger weight than the remaining genome-wide markers. In their simulation study, a 1.67-fold increase in prediction accuracy was observed when the genomic relationship matrix was weighted on the causal mutations (**Figure 4c**).

Fourth, the MultiBLUP method divides the chromosome into segments (**Figure 4f**, Speed and Balding 2014). Some of these segments are merged based on the likelihood ratio test. The remaining non-associated SNPs are merged into one segment as the background. Genomic relationship matrices are constructed for every segment and assigned to multiple random effects in gBLUP. Initial testing with human disease traits demonstrated improved accuracy over other GS methods (Speed and Balding 2014). However, testing in other species found MultiBLUP to be inconsistent depending on the trait and species. (Zeng and Zhou 2017).

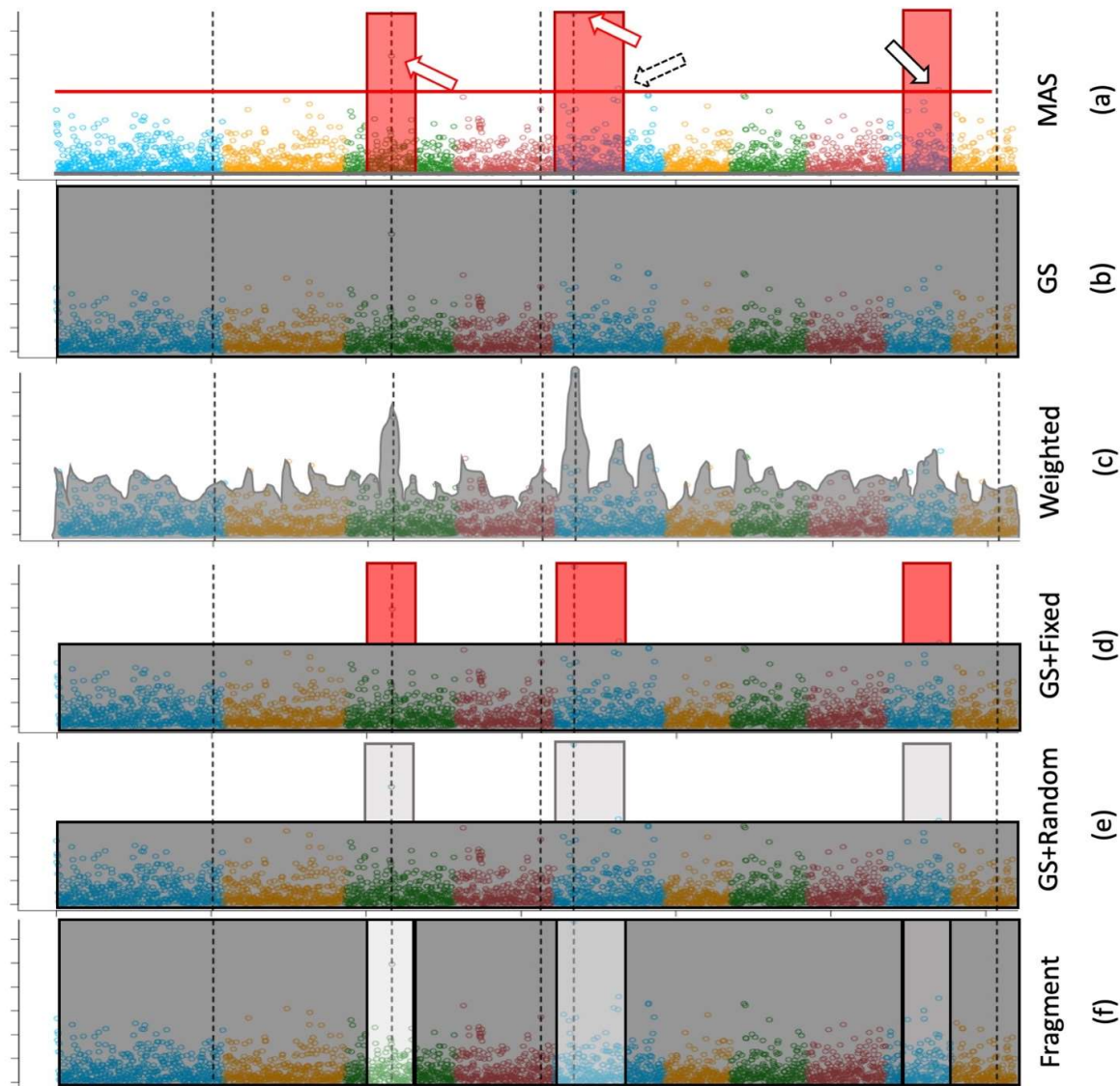


Figure 4. The connections between marker-assisted selection and genomic selection. Marker-Assisted Selection (MAS) uses the genetic markers that are significantly associated with a trait of interest by using a genome-wide threshold, the red line in (a) from Genome-Wide Association Study (GWAS). The associated markers may (red arrows) or may not (black arrows) be genetically linked to the causal genes indicated by the dashed vertical lines. Some of the associated markers may not be used for MAS, as they are close to markers with stronger association signals (black dashed arrow). Genomic Selection (GS) uses all markers to define the kinship among individuals to predict their breeding values either by weighting all the markers equally (b) or by their strength of association (c). Other variations include incorporating the associated SNPs as fixed effects (d), or random effects (e) in the regular GS (b). Another variation treats the fragments with associated SNPs and the rest of the fragments as different random effects indicated by different levels of gray (f).

Mixed Results for the Incorporation of Associated Markers

When applied in practice, mixed results were obtained when the GS model was modified to account for peak marker-trait associations (Lopes et al. 2017; Moore et al. 2017; Tan and Ingvarsson 2019). Prediction accuracy increased more than 10% when peak-associated markers from GWAS conducted in training sets were included as fixed-effect covariates for two out of four traits evaluated in rice (Spindel et al. 2016). Similarly, a 3-14% increase in prediction accuracy was reported for six Fusarium head blight resistance traits in wheat when a major-effect QTL was included in the GS model as a fixed-effect covariate (Arruda et al. 2016).

Other studies reported more modest increases, or even decreases, in prediction accuracy when using the same approach. When peak GWAS associations were included as fixed-effect covariates in gBLUP and Bayes B models, increases of 0.1-1% in prediction accuracy were observed for nine of 11 rice traits and two of three cattle traits (Zhang et al. 2014). Similarly, increases in prediction accuracies and bias were observed when meta-GWAS signals were included as fixed effect covariates in GS models evaluating stature in Jersey and Holstein bulls (Raymond et al. 2018). In a simulation study conducted in maize and sorghum, only 60 of 216 simulated genetic traits demonstrated improvement when peak-associated markers from a GWAS conducted on training sets were included in an RR-BLUP model (Rice and Lipka 2019). Moreover, the study observed an increase in the variance of prediction accuracies across replicate traits, as well as an increase in bias in the prediction accuracies, with the inclusion of such markers as fixed effects.

Similar mixed results were obtained when incorporating peak GWAS signals as random effects in the GS model. While an increase of up to 5% in reliability was observed for dairy production traits known to be controlled by major-effect genes, smaller increases in prediction accuracy were observed for mastitis and fertility, two traits where major-effect genes have not been consistently reported across dairy cattle breeds (Brøndum et al. 2015). Decrease in prediction accuracy was observed when the weighted genomic relationship matrix was based on GWAS results instead of the simulated effects of quantitative trait nucleotides (Fragomeni et al. 2017). These mixed results suggest that directly accounting for peak-associated genomic signals in GS models does not guarantee an increase in prediction accuracy and that the performance of such an augmented GS model depends upon the genetic architecture of the trait. Nevertheless, the adaptation of GS models to account for major-effect loci remains a relatively unexplored area of

research. In particular, a systematic study of the prediction accuracies of such augmented GS models as a function of sample size would shed light on their utility to the ever-increasing amounts of data that are becoming available.

VI. HYBRID BREEDING

Many major crops like maize and rice utilize heterosis for yield improvement and development of novel cultivars. Genomic selection in hybrids has been previously reviewed for circumstances of low marker density (Zhao et al. 2015), and here we focus on the application of high marker density, which creates both opportunities and challenges. For single-cross hybridization breeding in maize, GS is widely employed to select candidate combinations by predicting F_1 phenotypes from the corresponding genotypes of parental inbred lines (Cui et al. 2019; Riedelsheimer et al. 2012). Therefore, GS prediction of F_1 hybrids, which is used mainly to evaluate the hybrid performance of yield-related traits, must consider non-additive effects arising from heterotic loci in crop genomes (Zhou et al. 2012). In commercial maize breeding programs, application of GS usually partitions the training and prediction populations with the ratio of 1:4, to balance phenotyping costs and prediction accuracy (Erbe et al. 2010; Tan et al. 2017). The training population is composed of F_1 hybrids generated from crossing parental lines from two heterotic groups between which the progeny frequently expresses superior heterosis for grain yield. In a North Carolina II (NC-II) design (Garretsen and Keuls 1978) between the two heterotic groups of inbred lines, the F_1 hybrids with field-measured phenotypes in the training population accounted for 20% of all possible combinations, and the GS model predicted the phenotypes for the remaining 80% of combinations. The top 10-20% of F_1 combinations may be selected for field trials. Because heterotic effects may vary greatly among populations, traits, and environments, multiple factors need to be considered when designing a GS project for the best prediction of hybrid performance (Guo et al. 2019).

Factors Affecting Hybrid Prediction Accuracy

Many factors can affect hybrid prediction accuracy. First of all, the genetic backgrounds of training and prediction populations need to be uniform, which is achieved when the training samples are representative of the prediction population. This can be verified by examining the phylogenetic tree of parental lines built based on their genotypes. When this is not the case, the genetic

discrepancy between training and prediction populations will lead to problematic overfitting of the GS model, arising from the insufficiently expressed phenotypes, if phenotypes were collected in one place. For an extreme example in maize, if temperate and tropic maize lines are unevenly distributed in training and prediction populations, and the phenotypes are collected from temperate regions, the phenotypes for the tropic maize may not be fully expressed. Thus, model overfitting may occur as the correlation between genotype and phenotype is not correctly established for tropic maize.

Secondly, pedigree kinship and heterotic pattern of the two parental groups are essential for predictive effectiveness so that GS prediction with a much smaller training population than prediction population may achieve ideal accuracy to support decision making (Technow et al. 2014; Velazco et al. 2019). If the pedigree kinship and heterotic pattern were not explicit, heterotic performance of F_1 phenotypes may not be fully expressed, which may result in failure to establish the correct correlation between genotypes and phenotypes of F_1 hybrids.

Third, heterosis often involves interactions between genotype and environment ($G \times E$), including the influences arising from macro-environment and micro-environment (Li et al. 2018). While the macro-environment refers to the major ecological regions classified as different accumulative temperature belts, the micro-environment refers to the different planting conditions within the same ecological region, such as the local weather, nutritional conditions, microbial populations in the soil, and pathogenic conditions. Because macro-environment mainly influences flowering time attributed to photoperiodic genes with major effects, it is feasible to carry out a screening of parental lines based on prediction of the optimal ecological regions from genotypes to facilitate the design of training and prediction populations in a GS project.

Photo-thermal time was proposed to represent the environmental influence of one specific location on flowering time. Then, the correlation between genotype and photo-thermal time was modeled as a genotype-specific reaction-norm to quantitatively represent the contribution from $G \times E$ to phenotypic variations (Li et al. 2018). Thus, when the genotype of an inbred line and photo-thermal time of different locations are available, flowering time can be predicted to assist selection of the optimal environment for the inbred line to construct training populations. In contrast, the influence from micro-environment may be too complicated to be accurately modeled, as its effect is relatively minor and fluctuates. The ideal solution is to estimate its contribution to

phenotypic variations from multi-location field trials with the BLUP algorithm, so that the genetic contribution can be accurately modeled for GS prediction.

Heterosis Decomposition

Heterosis is generally considered the result of dominant and epistatic effects involving complex allelic, intra-genomic, and inter-genomic interactions. From the genetic perspective, heterosis can be dissected into two components based on the composition of an F_1 's phenotype. These components are the mid-parent value (the average of the corresponding maternal and paternal phenotypes), and the mid-parent heterosis (MPH) value (the deviation of F_1 phenotypes from the parental average) (Ali et al. 2012). The MP is considered the sum of all additive QTLs contributing to the trait, which may be inferred from the genomic relationship or kinship of the inbred lines in a pedigree (Bernardo 1994, Bernardo 1999; Massman et al. 2013; Crossa et al. 2017). The heterotic effect of a QTL is considered the combination of its dominant effect and all forms of di-genic epistatic effects, including additive-by-additive, additive-by-dominant, and dominant-by-dominant interactions, and the MPH of an F_1 hybrid is considered the summed effect of all the heterotic QTLs (Jiang et al. 2017). To accurately predict hybrid performance, all forms of additive and non-additive effects need to be considered in the GS model.

It is worth noting that composition of heterotic effects varies by trait. According to an analysis of hybrid rice, overdominant and pseudo-overdominant effects contribute most to the heterosis of grain yield, grain number per panicle, and grain weight; epistatic effects of dominant-by-dominant interactions contribute most to heterosis of tillers per plant, grain weight, and grain numbers; and single-locus dominant effects contribute small effects to almost all traits (Zhou et al. 2012). Therefore, GS prediction of different traits requires identification of heterotic QTLs, as well as quantitative inference of associated heterotic effects separately for each trait.

Modeling Heterosis

In practice, breeders traditionally use combining ability, including general combining ability (GCA) and specific combining ability (SCA), to evaluate hybrid performance and select candidate combinations. While the GCA is generally considered a reflection of the additive background effects of one parental genome, the SCA reflects non-additive interaction effects between the two parental genomes (Aliu et al. 2008). This is consistent with the assumption that the strength of

heterosis expression of an F_1 hybrid is the summed result of additive and non-additive QTLs. Thus, the outstanding GCA and SCA together result in excellent heterosis. Because GCA and SCA may arise from different loci with distinct types of genetic effects, the appropriate solution to GS prediction on F_1 hybrids is to separately consider the effects from GCA and SCA.

Evaluation of GCA can be implemented through a linear model using the gBLUP (genomic Best Linear Unbiased Prediction) method, which uses the genomic relationship matrix inferred from whole-genome markers to predict phenotypes. Environmental factors can also be integrated into the gBLUP model as covariates (fixed effects). Because SCA is due to heterotic loci with non-additive effects, a small panel of heterotic SNPs with quantitatively estimated effects is essential for an accurate evaluation of SCA in two possible ways. The first way to evaluate SCA is to integrate heterotic SNPs into the gBLUP model as fixed effects to enhance prediction accuracy. The second way is to use non-linear models, such as Bayesian and machine learning methods, to predict MPH, followed by combining GCA and SCA to assist selection of candidate F_1 combinations for field trials. The compilation of the heterotic SNPs may be implemented via the GWAS analysis of the MPH value of a trait to first identify heterotic QTLs. Then, one of the haplotypic tagging SNPs with significant p-value, closest to the peak SNP of the QTL, may be selected as the heterotic SNP. The number of heterotic SNPs used in the panel is usually dependent on the number of significant QTLs, typically 3 to 5 SNPs depending on the trait.

VII. MULTIPLE TRAITS

It is common that new varieties of crops are selected for multiple traits. For example, crop breeders observe phenotypic data to evaluate multiple traits in different categories, such as resistance to biotic stress (e.g., insect, plant disease, fungal) or abiotic stress (e.g., cold, heat, wind, flood), grain quality (e.g., color, taste, shape, nutrient level), and yield components (e.g., grain weight or biomass). The multiple traits being selected for very often share a certain level of common genetic architecture and thus are genetically correlated. This applies to non-agricultural species as well; an evaluation of genetic correlations among 24 traits in human genetic studies demonstrated shared genetic controls between many complex traits and diseases (Bulik-Sullivan et al. 2015).

For a long time, researchers have taken advantage of genetically correlated traits to map the genetic control of complex traits through various quantitative genetics approaches (Banerjee et al. 2008; Jiang and Zeng 1995; Xu et al. 2009). However, when genomic selection models were

initially developed, the genome-wide marker effects were modeled to calculate the genomic estimated breeding value of a single trait (Bernardo 1994; Meuwissen et al. 2001). It took almost two decades before multi-trait genomic selection (MT-GS) received full attention. In 2011, the advantage of MT-GS was clearly demonstrated over single-trait genomic selection for genetically correlated traits using simulated data for animal breeding scenarios (Calus and Veerkamp 2011). In 2012, Jia and Jannink presented more detailed model development of multivariate BayesA and BayesC π using plant examples (Jia and Jannink 2012). In this report, the authors demonstrated for the first time that traits with low heritability can benefit from correlated traits with high heritability in GS. This insight provides one approach to apply MT-GS to accelerate genetic gain, especially for traits with low heritability or that are expensive to measure. Furthermore, the ability to impute phenotypes for focal traits with missing data using secondary traits measured on all individuals in both training and selection populations can improve the prediction accuracy of the individuals for selection by over 50% compared with prediction based on the focal traits alone.

The MT-GS statistical framework has been widely applied to broad breeding scenarios and human genetic studies after its initial published application in animals (Calus and Veerkamp 2011) and plants (dos Santos et al. 2020; Jia and Jannink 2012; Watson et al. 2019). In dairy cattle, reproductive and yield traits were modeled together through MT-GS (Guo et al. 2014) and confirmed the benefit of a trait with low heritability from MT-GS methods using a high-heritability trait. In humans, multivariate gBLUP was applied for genetic risk prediction of multiple diseases (Maier et al. 2015). With the MT-GS model, risk prediction is significantly improved, equivalent to the effects of increased sample size, by 34% for schizophrenia, 68% for bipolar disorder, and 76% for major depressive disorders with single-trait models (Maier et al. 2015). In plants, MT-GS has been evaluated in breeding data for soybean (Volpato et al. 2019), rice (Wang et al. 2017), cranberry (Covarrubias-Pazaran et al. 2018), rye (Schulthess et al. 2016), and hybrid wheat disease prediction (Schulthess et al. 2018).

Multi-trait Genomic Selection Models

The MT-GS model using gBLUP with marker-based kinship can be implemented as a regular BLUP with the numerator relationship matrix calculated from pedigree information. The only difference is the substitution of kinship. The MT-GS model of Bayesian methods is typically a linear regression as below.

$$\mathbf{y} = \mathbf{u} + \sum_{j=1}^p X_j \mathbf{a}_j \delta_j + \mathbf{e}$$

where \mathbf{y} is the a matrix ($n \times m$) for m trait observations on n individuals, X is a design matrix ($n \times p$) representing the p marker genotypes on n individuals, \mathbf{a}_j is a vector of size m for the effects of molecular marker j on all m traits, δ_j is an unknown indicator variable with value 0 if marker j is not in the model and value 1 otherwise, \mathbf{e} is a matrix ($n \times m$) of residuals.

The effect vector \mathbf{a}_j for marker j in a Bayesian MT-GS model is estimated from hierarchical models using multivariable normal distribution $N(0, \Sigma_{\mathbf{a}_j})$. The variance of the distribution for marker j can be 0 if the marker j is not in the model under the BayesC π framework where all markers share a common variance value. In contrast, under the BayesA method, the variance and covariance of each marker for multiple traits are estimated separately (Jia and Jannink 2012). Under the Bayesian framework, the correlation between the adjacent SNP markers on the chromosome due to linkage disequilibrium can be incorporated through the Bayesian multivariate antedependence model (Jiang et al. 2015). In addition, a multi-trait GS model can be expanded to a multi-trait and multi-environment Bayesian model with an available R package, named BMTME, standing for Bayesian multi-trait and multi-environment (Montesinos-Lopez et al. 2016; Montesinos-Lopez et al. 2019).

Besides the linear regression model through gBLUP and the Bayesian framework, the MT-GS model can also be realized through non-linear frameworks such as machine learning or a deep-learning algorithm. To model multiple phenotypes with mixed data types (e.g., binary, ordinal, and continuous), Montesinos-Lopez et al. 2019 developed the multiple-trait deep learning with mixed phenotypes (MTDLMP) model. The authors found that the MTDLMP model is equivalent to the univariate models with $G \times E$ interaction terms but shows better performance without $G \times E$ interaction terms (Montesinos-Lopez et al. 2019).

Multi-trait GS Applications in Plant Breeding

The MT-GS framework has been developed using different statistical or machine learning algorithms to fit different data types or needs in plant breeding. First, the primary and original application of MT-GS was to improve the prediction accuracy of traits of interest by integrating multiple correlated traits together. The insight that traits of low heritability can benefit from correlated traits with high heritability through the MT-GS model narrows the selection of favorable

secondary traits. Second, when a secondary trait is available for individuals in both training and selection populations, the framework of phenotype imputation (Jia and Jannink 2012) or trait-assisted genomic selection can improve the prediction accuracy by over 50% compared with the single-trait selection strategy (Fernandes et al. 2018). In practice, such secondary traits can be measured on individuals in an early stage of selection to implement stronger selection intensity. High-throughput phenotyping using platforms such as unmanned aerial vehicles has significantly enriched data collection for secondary traits (Hassan et al. 2019; Moreira et al. 2020). As the rapid development of high-throughput phenotyping techniques continues, more relevant and useful phenotypes or traits will be available for MT-GS applications in plant breeding to meet different breeding needs. Additionally, the MT-GS model has the potential to shed light on the molecular characterization underlying correlated traits, especially when a specific DNA segment is detected for multiple traits of interest. The confirmed pleiotropy can be used to identify targets of genome editing or another molecular engineering approach to introduce a beneficial mutation to unlink the negative correlation between the multiple traits of interest (Börgel 2018).

VIII. LONG-TERM SELECTION

To increase the response of long-term selection, it is necessary to balance genetic gain and genetic diversity. Under long-term selection, the effective population size is reduced because of selection. Consequently, genetic diversity decreases, causing the loss of favorable alleles and an increase in the rate of inbreeding (Bijma 2012). Although there are numerous simulation and cross-validation studies that report GS is more accurate than conventional selection, there is no direct evidence that GS increases long-term genetic gain in practical breeding schemes (Henryon et al. 2014).

In 2018, Doekes et al. evaluated genome-wide and region-specific diversity in Holstein–Friesian bulls in the Dutch–Flemish breeding program from 1986 to 2015 (Doekes et al. 2018). After the introduction of GS, rates of inbreeding and kinship substantially increased. Some methods used to reduce inbreeding, such as the method of optimal contribution selection (OCS), which aims to maximize genetic gain with a restricted rate of inbreeding, caused a drop in the rates of inbreeding and kinship (Meuwissen 1997), suggesting that careful management of genetic diversity is indispensable to enhancing long-term genetic gain in GS breeding. To enhance long-term genetic gain, various methods have been proposed. The methods are roughly divided into the five categories reviewed here, including (1) upweighting rare favorable alleles, (2) genomic

optimal contribution selection (GOCS), (3) selection based on potential progeny, (4) marker densities and prediction models, and (5) design of breeding populations.

Upweighting Rare Favorable Alleles

The idea behind this approach is to weight favorable alleles with lower frequencies more heavily to avoid losing such alleles from a population. In 2009, Goddard argued that the optimal weight to fix all favorable alleles in long-term selection depended only on the frequencies of the alleles, whilst the optimal weight in single-cycle selection depended only on the effects of the alleles, and suggested that the practical optimal is an intermediate of these two extremes (Goddard 2009). In 2010, Jannink proposed to weight markers based on both the effects and frequencies of favorable alleles. This approach has been modified to further improve long-term genetic gain (Jannink 2010). In 2014, Sun and VanRaden proposed the inclusion of a parameter in the weight, to balance long-term and short-term gains (Sun and VanRaden 2014). To increase long-term genetic gain within a fixed time horizon, a dynamic weighting method was proposed (Liu et al. 2015), which decreases weights on the rate of favorable alleles over the time horizon.

Genomic Optimal Contribution Selection

To balance genetic gain and the rate of inbreeding, Optimal Contribution Selection (OCS) has been efficiently used in animal breeding (Doekes et al. 2018). In OCS, rates of inbreeding are restricted by constraining the average relationship of candidates using an additive relational matrix inferred from pedigree information. In 2012, Sonesson et al. proposed Genomic OCS (GOCS), in which the additive relational matrix is replaced by a realized genomic relationship matrix calculated from marker data (Sonesson et al. 2012). Based on a similar idea to OCS, De Beukelaer et al. proposed two methods in the following year to balance genetic gain and genetic diversity (expected heterozygosity or maintenance of rare alleles) via index weighting of the gain and diversity (De Beukelaer et al. 2017). In 2018, Gorjanc et al. extended GOCS to select optimal crosses rather than optimal contributions (Gorjanc et al. 2018). Another method was proposed to select optimal crosses by balancing genetic gain and inbreeding in a different formulation (Akdemir and Sánchez 2016). Recently, it was found via simulations that OCS based on pedigree relationships realized greater genetic gain than GOCS, even when gBLUP was used as the prediction method (Mark Henryon et al. 2019), suggesting that OCS is better than GOCS even in GS breeding.

Selection Based on Potential Progeny

The idea behind this approach is to select candidates based on the genetic ability of potential progeny that can be produced from the candidates. In 2015, Daetwyler et al. proposed optimal haploid value selection, which predicts the best doubled haploid that can be produced from a candidate (Daetwyler et al. 2015). The authors found that optimum haploid value selection preserved a greater degree of genetic diversity than GS and was important in achieving long-term genetic gain. In 2018, Müller et al. proposed an expected maximum haploid breeding value, which predicts the expected GEBV of the best double haploid lines produced from a candidate (Müller et al. 2018a). This expected maximum haploid breeding value has two advantages over optimum haploid value. First, it can take the finite size of a double haploid population into account. Second, it does not require the partition of the genome into haplotype segments. In 2017, Goiffon et al. proposed an optimal population value selection, which is selection based on the maximum possible haploid value in a subset of the population (Goiffon et al. 2017). Optimal population value selection aims to select a set of candidates that can produce the best progeny, rather than candidates with superior individual abilities. In 2019, Moeinizade et al. extended optimal population value selection to “look-ahead selection” to improve genetic gain at a time horizon by maximizing expected genomic breeding values of the best progeny in the terminal generation (Moeinizade et al. 2019).

Marker Densities and Prediction Models

Long-term genetic gain is potentially influenced by marker densities and prediction models because weights on favorable alleles can be changed with differences in marker densities and prediction models. In 2014, MacLeod et al. performed simulations to evaluate whether whole-genome sequence data increased the accuracy over recurrent GS (MacLeod et al. 2014). They found that whole-genome data increased the accuracy with large effective population sizes but provided little advantage with reduced effective population size. In 2015, Liu et al. performed simulations and found that, compared to gBLUP, Bayesian Lasso led to less genetic drift and reduced loss of favorable alleles, and more effectively controlled the rate of both pedigree and genomic inbreeding (Liu et al. 2015).

Design of Breeding Populations

Genetic diversity can be maintained by designing breeding populations. In 2016, Yabe et al. proposed the island-model design of breeding populations, in which a large population was split into multiple subpopulations and each subpopulation received migrants from the others. The method was inspired by evolutionary algorithms, as well as population genetics. A simulation study suggested that the island-model GS better maintained genetic variation and genetic gain in later generations than did the normal GS (Yabe et al. 2016).

No study has comprehensively compared the potential of the methods in long-term selection, although some comparative studies have been published (de Rezende Neves et al. 2018; Müller et al. 2018). Moreover, various approaches can be extended from current and novel approaches. A “benchmark test” system for selection methods would be useful for the relative evaluation of the potential of the methods (Wang et al. 2018).

IX. ASSESSMENT OF PREDICTION ACCURACY

The ultimate goal of GS is to accurately predict the genetic potential for individuals that do not have phenotypes yet. To evaluate the accuracy of a prediction method or a breeding program, individuals with phenotypes must be used appropriately masked to conduct a valid assessment of prediction accuracy. Model fit in a training population cannot be used for the assessment of prediction accuracy due to overfitting problems, because most of the prediction involves more parameters than the number of observations. Validation of prediction must be conducted in an independent population, where the observations are not used to derive those parameters for prediction. The use of a separated population eliminates any bias and violation of using observations for both training and validation. The disadvantage is not being able to evaluate prediction accuracy when adding a proportion of the independent population as the training population. Consequently, cross-validation is one of the major approaches to evaluate prediction accuracy.

During the cross-validation process, the entire population is divided into k folds (e.g., five or ten). One of them is used as a testing population, while the rest are used as a training population. Then, a model is built using only the training population and used to predict phenotypes for the testing population. The Pearson correlation coefficient between the observed and the predicted phenotypes can be instantly calculated from the individuals in the current testing population. This process of training a model and predicting phenotypes for a single testing fold is iterated until all

folds have been treated separately as the testing population. When correlation coefficients are calculated separately for each testing fold and then averaged, this final accuracy value is referred to as the “instant” prediction accuracy. Alternatively, a single correlation can be calculated using predicted phenotypes across all folds. In this case, the accuracy value is referred to as the “hold” prediction accuracy (Zhou et al. 2016).

Both the hold and instant methods are used to evaluate prediction accuracy through cross-validation. The difference between the two methods has not been fully recognized by the public, because details have rarely been provided in literature reports regarding the method used to calculate the correlation coefficients. In particular, GS models with substantial negative cross-validated prediction accuracies may be suffering from bias related to how the accuracy was calculated. For example, significant negative correlation coefficients ranging from -0.24 to -0.42 were published in a genomic prediction study of maize (Massman et al. 2013). One explanation could be the switch of prediction sign. The other possibility is the expectation bias under the null hypothesis that there is no correlation between observed and predicted phenotypes in cross-validation, for either hold or instant methods or both.

The difference between the instant and hold method accuracy was investigated. Using simulated and real data, it was demonstrated that both approaches exhibited downward bias under certain circumstances (Zhou et al. 2016). When the expected accuracy was zero, instant accuracy remained unbiased, whereas hold accuracy was negatively biased, especially with more folds. Instant accuracy exhibited bias only when expected accuracy was far from 0 and 1, and the inference population was small. This bias can be corrected to a nearly unbiased estimate by modifying the formula used to calculate the instant accuracy. Considering these potential biases, it is important that future reports involving cross-validated accuracies also include methodological details regarding how the accuracy was calculated.

In addition to systematic bias due to hold accuracy and Person correlation using small samples, there many other factors for biased estimation of prediction accuracy (Runcie and Cheng 2019), including bias in testing data selection, overlap between testing and training populations, and pre-selection of GWAS features before splitting data for cross-validation. Because specific details regarding cross-validation procedures are overlooked or incomplete, it is not possible to determine if cross-validation errors have contributed to the mixed results. Proper cross-validation

procedure involves dividing the entire population first, then GWAS should be conducted using the training population only (**Figure 5**).

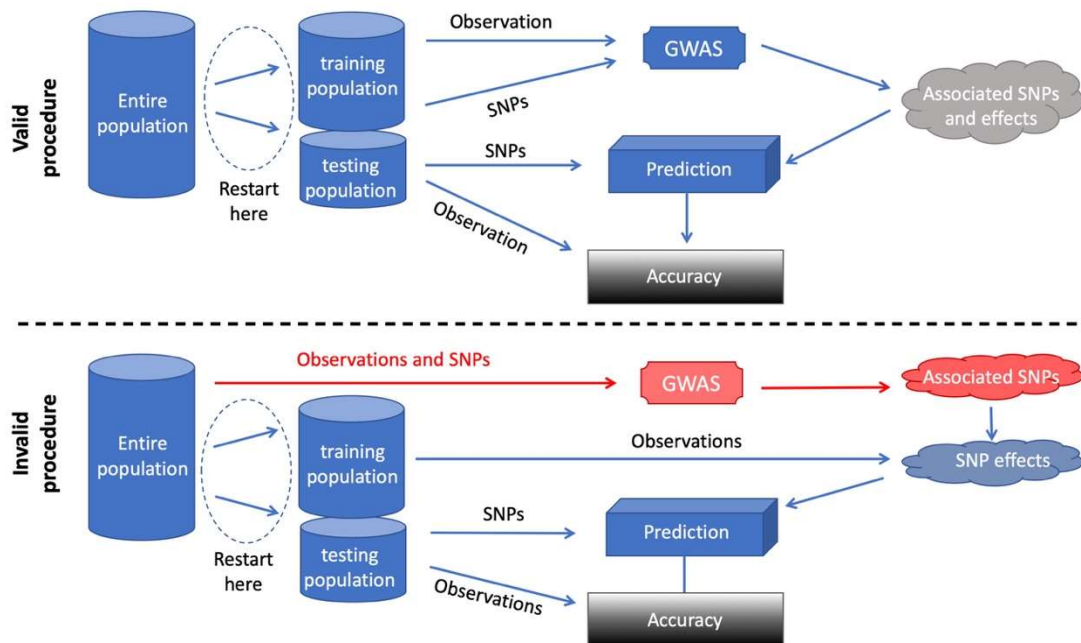


Figure 5. Valid and invalid procedures to evaluate enhancement of incorporating GWAS results in genomic selection. Prediction accuracy is evaluated in a testing population whose phenotypes should not be used in any format to develop the predictions. Any contamination by using the phenotypes of the testing population may result in errors in prediction. The valid procedure (above the dashed line) should divide the whole population into training and testing populations first. Then the phenotypes of the testing population should not be used in any cases except for comparison between the predicted and observed phenotypes. An invalid procedure (below the dashed line) conducts activities on the whole population before splitting it into training and testing populations. In this example, Genome-Wide Association Study (GWAS) is conducted on the whole population. The phenotypes in the testing population are used for marker classification. Even if effects of associated markers identified by GWAS are re-estimated in the training population only, the selection of the associated markers is contaminated by the usage of the phenotypes in the testing population.

However, if GWAS is conducted using the entire population before dividing it into training and testing populations, it violates this requirement even if the effects of the associated markers are reevaluated using only the testing population for prediction. To demonstrate the bias of overestimated prediction accuracy, we simulated a random number as a phenotype for each of the 926 individuals genotyped with 4853 SNPs from the standard dataset of Loblolly Pine (Resende et al. 2012). There were no connections between the phenotypes and genotypes. The expected prediction accuracy was zero. The model fitting in the training population exhibited overestimates

in all methods examined. The prediction accuracies in the testing population were zeros for using Ridge Regression, or GWAS assisted Ridge Regression using the valid procedure. However, when the GWAS assisted Ridge Regression was conducted using an invalid procedure, an average artifact of positive prediction accuracy (0.14) was observed (Figure 6).

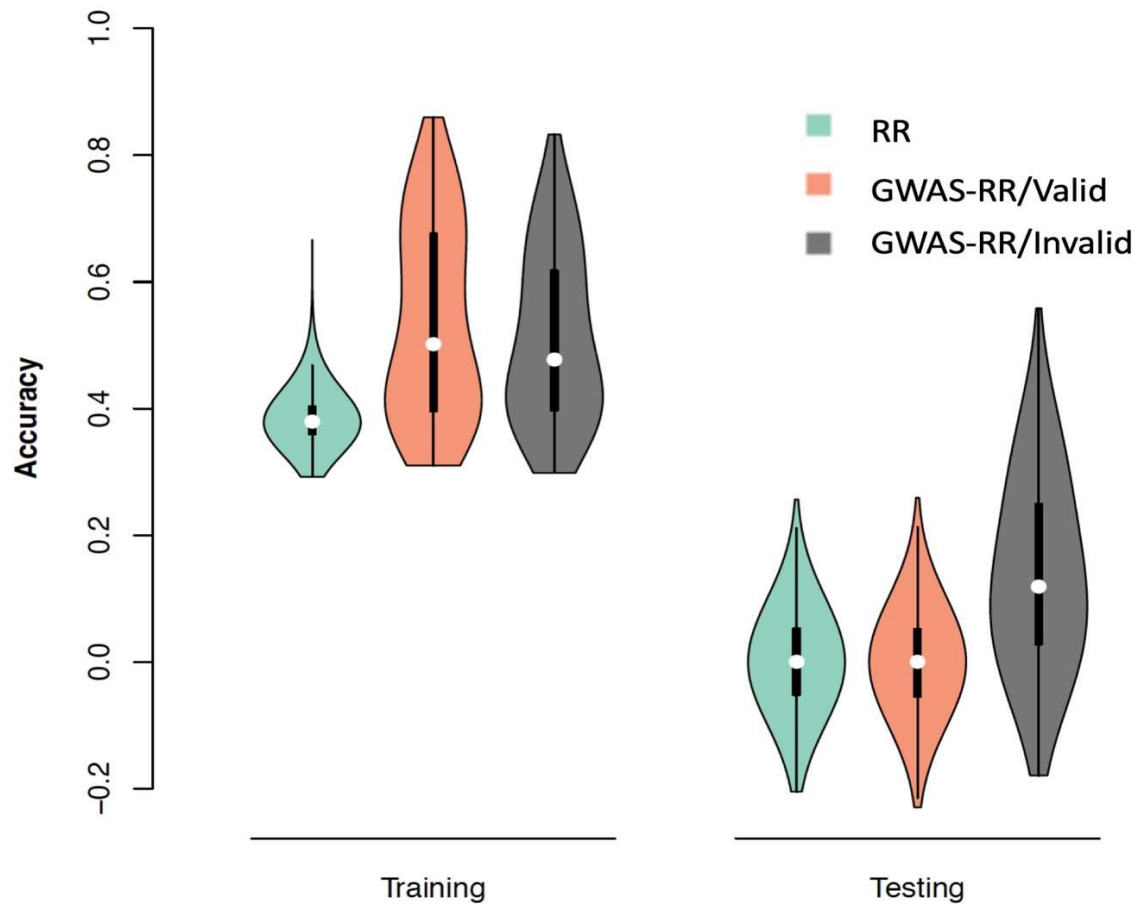


Figure 6. Accuracy of predictions on randomly generated noise. Random noises were generated as phenotypes for 926 individuals genotyped with 4853 SNPs. These SNPs were used to predict the random noises in training and testing populations by using two methods: 1) ridge regression (RR); and 2) genome-wide association study (GWAS)-assisted RR. GWAS-assisted RR required RR to be conducted twice; see Resende et al. (Genetics, 2012, Vol 190, 1503-1510) for details. The first RR ranked markers based on their absolute effects. The second RR re-estimated the effects for the markers that maximized the correlation between predicted and observed phenotypes. The 926 individuals were randomly separated into the training population (80%) and testing population (20%). The marker effects were estimated for all the SNPs using RR from the training population. The estimated marker effects were used to predict the random noises in the training and the testing populations, where Pearson correlation coefficients were calculated separately as prediction accuracies. The GWAS-assisted RR was conducted using two procedures: valid and invalid. The valid procedure (GWAS-RR/Valid) split training and testing populations first. GWAS and estimation of marker effects were conducted in the training population. The invalid procedure (GWAS-RR/Invalid) was contaminated by the usage of observations from the

testing population. GWAS was conducted first on the entire population to select the associated SNPs. These SNPs were then used to estimate marker effects in RR conducted in the training population. The process was replicated 1000 times. The results suggested an artifact of positive prediction accuracy in the training population, and also an artifact of positive prediction accuracy (0.14) in the testing population, when the invalid procedure was used.

X. GS TRANSFORMED PLANT BREEDING

Although GS methods were introduced to plant breeding earlier than in animals (Bernardo 1994; Meuwissen et al. 2001), animal breeders adapted to GS applications much faster due to savings of time and phenotyping (Hickey et al. 2017). Because marker-based kinship is easy to implement in existing animal genetic evaluation systems (VanRaden 2008; Zhiwu Zhang et al. 2007), gBLUP and its derivations such as single-step BLUP have been widely used in practical animal breeding (Hayes et al. 2009; Hickey et al. 2017; Jonas and De Koning 2013). Using cattle as an example, to reach reliable prediction of a bull's milk potential, hundreds of daughters are required. This can be done relatively soon after the birth of the bull at a cost of several hundred dollars, or even much less. In contrast, the incorporation of GS approaches into plant breeding programs initially lagged behind animal breeding programs. A 2011 in-depth review of genomic selection methodology and application in plants (Lorenz et al. 2011) identified only a single peer-reviewed study that focused on empirical GS accuracy in crops before 2011 (Lorenzana and Bernardo 2009).

A breakthrough for the incorporation of GS in plants was the reduced cost of acquiring high-throughput, genome-wide genotypes for large genomes with high repeat content, which is typical for many agronomically important plant species. In the last decade, two methodologies for genotyping plant species have emerged that can economically scale to the sample sizes required for GS and are adaptable to a wide range of species. These are high-density, single-nucleotide polymorphism arrays (SNP chips) and genotyping-by-sequencing (Ganal et al. 2019). A major review of GS in 2017 (Varshney et al. 2017) indicates that the availability of these high-throughput genotyping methods has helped GS to become more feasible in plant studies. Numerous empirical studies were listed that reported predictive accuracies of different GS models in several plant species. By far, the most represented plant group were the cereal grains with over 40 papers listed. Many non-grain species were also reported, with 7 in legumes, 5 in clonally propagated crops, and 13 for forest trees. Across all of these papers there was consensus that while the best specific GS modeling approach varied depending on the trait architecture, GS models in general were capable

of predicting many agronomically important traits with a high enough accuracy compared to phenotypic selection to warrant incorporation into plant breeding programs. Since this review in 2017, the number of studies demonstrating the predictive capabilities of GS has increased and expanded from previously studied crops such as wheat (Arruda et al. 2016; Belamkar et al. 2018; Haile et al. 2018; Lozada et al. 2019), pea (Annicchiarico et al. 2019), and alfalfa (Biazzi et al. 2017), to a wider variety of plant species including sunflower (Dimitrijevic and Horn 2018), coffee (Sousa et al. 2019), cassava (Wolfe et al. 2017), strawberry (Gezan et al. 2017), winter squash (Hernandez et al. 2020), and cranberry (Covarrubias-Pazaran et al. 2018). Cost-benefit analyses of incorporating GS into breeding programs were also discussed and indicated that GS can be substantially cheaper than phenotypic selection and brings the advantage that genotypes can be used for predicting multiple traits. While most discussion regarding the cost advantage of GS analysis has primarily focused on grains (Bernardo 2009; Endelman et al. 2014; Heffner et al. 2010; Voss-Fels et al. 2019), a large portion of the identified cost benefit is due to a decreased breeding cycle (Rutkoski 2019). Therefore, the cost advantage of GS is likely to be similar or even amplified in other species with longer breeding cycles, such as perennial crops and trees. However, considering the high structural diversity in plant breeding programs and biological differences between plant species, more specific studies are needed that explore the logistics of cost-effectively incorporating GS into different types of breeding programs. The ongoing dialogue surrounding GS in plant breeding is no longer primarily a question of capability but instead has shifted towards implementation.

XI. FUTURE PROSPECTS

The studies reviewed here reach both inside and outside the breeder's equation. With rapid advances in genotyping technologies during the first two decades in the 21st century, agronomics and genomics are fully connected as a straight line through genomic selection in plant breeding. This line is expanding into a surface through phenomics, formatting a Genomics-Phenomics-Agronomics (GPA) paradigm. Phenomics not only measures some existing agronomic traits in a high-throughput fashion, but also adds new heritable measurements that are correlated with conventional agronomic traits, or completely new traits with new agronomic content. Examples of this rapidly expanding research area include metabolomics (Zampieri et al. 2017), automated hyperspectral imaging of plant foliage (Thomas et al. 2018), and nondestructive large-scale

imaging of root systems with minirhizotrons (Svane et al. 2019). For decades to come, the GPA paradigm is expected to continually transform the breeder's equation, inside and outside. These transformations include at least five areas: 1) transformation from the prediction of parental performance to the prediction of the potential of progeny; 2) expansion of prediction beyond additive effects to include dominant, epistatic, and genetic-by-environment interaction effects; 3) utilization of pleiotropy among conventional agronomics traits and high-throughput phenomics traits; 4) a deeper understanding of the genetic architecture of agronomic traits through GWAS and integration of these architectures into GS models; and 5) the utilization of emerging big data and machine learning approaches to take advantage of the high dimensionality of high-throughput data, especially deep learning using artificial neural networks.

FUNDING:

This project was partially funded by the USDA NIFA (Hatch project 1014919, Award #s 2016-68004-24770, 2018-70005-28792, 2019-67013-29171, and 2020-67021-32460), JST CREST (Grant Number JPMJCR16O2), the Washington Grain Commission (Endowment and Award #s 126593 and 134574), and National Science Foundation of China (Grant Number 31871706).

ACKNOWLEDGMENTS:

The authors thank Cari Park for helpful comments and editing the manuscript.

CONFLICTS OF INTEREST:

The authors declare no conflict of interest.

LITERATURE CITED

- Aguilar, I., Misztal, I., Johnson, D. L., Legarra, A., Tsuruta, S., and Lawlor, T. J. 2010. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *Journal of Dairy Science*. 93(2): 743–752. DOI: 10.3168/jds.2009-2730
- Akdemir, D., and Sánchez, J. I. 2016. Efficient breeding by genomic mating. *Frontiers in Genetics*. 7(NOV). DOI: 10.3389/fgene.2016.00210
- Ali, F., Shah, I. A., Khyber, M. N., and Shahwar, D. 2012. Heterosis for yield and agronomic attributes in diverse maize germplasm From morphological to mechanistic insights on modern maize breeding View project Molecular behaviour of GmTGAs in soybean nodulation. *Australian Journal of Crop Science*. 6(3): 455–462.

- Aliu, S., Fetahu, S. H., and Salillari, A. 2008. Estimation of heterosis and combining ability in maize (*Zea mays* L.) for ear weight (EW) using the diallel crossing method. *Agronomijas Vestis*. 11: 7–12.
- Annicchiarico, P., Nazzicari, N., Pecetti, L., Romani, M., and Russi, L. 2019. Pea genomic selection for Italian environments. *BMC Genomics*. 20(1): 603. DOI: 10.1186/s12864-019-5920-x
- Arruda, M. P., Lipka, A. E., Brown, P. J., Krill, A. M., Thurber, C., Dong, Y., Foresman, B. J., and Kolb, F. L. 2016. Comparing genomic selection and marker-assisted selection for Fusarium head blight resistance in wheat (*Triticum aestivum* L.). *Molecular Breeding*. 36(7): 1–11. DOI: 10.1007/s11032-016-0508-5
- Azodi, C. B., Bolger, E., McCarren, A., Roantree, M., de Los Campos, G., and Shiu, S. H. 2019. Benchmarking Parametric and Machine Learning Models for Genomic Prediction of Complex Traits. *G3: Genes, Genomes, Genetics*. 9(11): 3691–3702. DOI: 10.1534/g3.119.400498
- Banerjee, S., Yandell, B. S., and Yi, N. 2008. Bayesian quantitative trait loci mapping for multiple traits. *Genetics*. 179(4): 2275–2289. DOI: 10.1534/genetics.108.088427
- Belamkar, V., Guttieri, M. J., Hussain, W., Jarquín, D., El-basyoni, I., Poland, J., Lorenz, A. J., and Baenziger, P. S. 2018. Genomic selection in preliminary yield trials in a winter wheat breeding program. *G3: Genes, Genomes, Genetics*. 8(8): 2735–2747. DOI: 10.1534/g3.118.200415
- Bellot, P., de los Campos, G., and Pérez-Enciso, M. 2018. Can deep learning improve genomic prediction of complex human traits? *Genetics*. 210(3): 809–819. DOI: 10.1534/genetics.118.301298
- Bernardo, R. 1994. Prediction of maize single-cross performance using RFLPs and information from related hybrids. *Crop Science*. 34(1): 20–25. DOI: 10.2135/cropsci1994.0011183X003400010003x
- Bernardo, R. 1999. Marker-assisted best linear unbiased prediction of single-cross performance. *Crop Science*. 39(5): 1277–1282. DOI: 10.2135/cropsci1999.3951277x
- Bernardo, R. 2009. Genomewide Selection for Rapid Introgression of Exotic Germplasm in Maize. *Crop Science*. 49(2): 419–425. DOI: 10.2135/cropsci2008.08.0452
- Bernardo, R. 2014. Genomewide selection when major genes are known. *Crop Science*. 54(1): 68–75. DOI: 10.2135/cropsci2013.05.0315
- Biazzi, E., Nazzicari, N., Pecetti, L., Brummer, E. C., Palmonari, A., Tava, A., and Annicchiarico, P. 2017. Genome-Wide Association Mapping and Genomic Selection for Alfalfa (*Medicago sativa*) Forage Quality Traits. *PLOS ONE*. 12(1): e0169234. DOI: 10.1371/journal.pone.0169234
- Bijma, P. 2012. Accuracies of estimated breeding values from ordinary genetic evaluations do not reflect the correlation between true and estimated breeding values in selected populations. *Journal of Animal Breeding and Genetics*. 129(5): 345–358. DOI: 10.1111/j.1439-0388.2012.00991.x
- Börgel, A. W. S. 2018. Evaluating the benefits and limitations of multiple-trait breeding assisted by genomics in cereal crops. Dissertation. Martin Luther University Halle-Wittenberg.
- Breiman, L. 2001. Random forests. *Machine Learning*. 45(1): 5–32. DOI: 10.1023/A:1010933404324
- Brøndum, R. F., Su, G., Janss, L., Sahana, G., Guldbrandtsen, B., Boichard, D., and Lund, M. S. 2015. Quantitative trait loci markers derived from whole genome sequence data increases the reliability of genomic prediction. *Journal of Dairy Science*. 98(6): 4107–4116. DOI: 10.3168/jds.2014-9005
- Bulik-Sullivan, B., Finucane, H. K., Anttila, V., Gusev, A., Day, F. R., Loh, P. R., ReproGen, C., Psychiatric Genomics, C., for Anorexia Nervosa of the Wellcome Trust Case Control, C., Duncan, L., Perry, J. R., Patterson, N., Robinson, E. B., Daly, M. J., Price, A. L., and Neale, B. M. 2015. An atlas of genetic correlations across human diseases and traits. *Nat Genet*. 47(11): 1236–1241. DOI: 10.1038/ng.3406
- Calus, M P L. 2010. Genomic breeding value prediction : methods and procedures. *Animal*. 4(2): 157–164. DOI: 10.1017/S1751731109991352
- Calus, Mario P.L., and Veerkamp, R. F. 2011. Accuracy of multi-trait genomic selection using different methods. *Genetics Selection Evolution*. 43(1): 26. DOI: 10.1186/1297-9686-43-26
- Campos, G. S., Reimann, F. A., Cardoso, L. L., Ferreira, C. E. R., Junqueira, V. S., Schmidt, P. I., Neto, J. B., Yokoo, M. J. I., Sollero, B. P., Boligon, A. A., and Cardoso, F. F. 2018. Genomic prediction using

- different estimation methodology, blending and cross-validation techniques for growth traits and visual scores in Hereford and Braford cattle. *Journal of Animal Science*. 96(7): 2579–2595. DOI: 10.1093/jas/sky175
- Cassidy, E. S., West, P. C., Gerber, J. S., and Foley, J. A. 2013. Redefining agricultural yields : from tonnes to people nourished per hectare. *Environmental Research Letters*. 8(3). DOI: 10.1088/1748-9326/8/3/034015
- Chen, C. Y., Misztal, I., Aguilar, I., Tsuruta, S., Meuwissen, T. H. E., Aggrey, S. E., Wing, T., and Muir, W. M. 2011. Genome-wide marker-assisted selection combining all pedigree phenotypic information with genotypic data in one step: An example using broiler chickens. *Journal of Animal Science*. 89(1): 23–28. DOI: 10.2527/jas.2010-3071
- Chen, X., and Ishwaran, H. 2012. Random forests for genomic data analysis. In *Genomics*. 99(6):323-329. DOI: 10.1016/j.ygeno.2012.04.003
- Christensen, O. F., and Lund, M. S. 2010. Genomic prediction when some animals are not genotyped. *Genetics Selection Evolution*. 42(1). DOI: 10.1186/1297-9686-42-2
- Covarrubias-Pazaran, G., Schlautman, B., Diaz-Garcia, L., Grygleski, E., Polashock, J., Johnson-Cicalese, J., Vorsa, N., Iorizzo, M., and Zalapa, J. 2018. Multivariate GBLUP Improves Accuracy of Genomic Selection for Yield and Fruit Weight in Biparental Populations of *Vaccinium macrocarpon* Ait. *Frontiers in Plant Science*. 9: 1310. DOI: 10.3389/fpls.2018.01310
- Crossa, J., Martini, J. W. R., Gianola, D., Pérez-rodríguez, P., Jarquin, D., Juliana, P., Montesinos-lópez, O., and Cuevas, J. 2019. Deep Kernel and Deep Learning for Genome-Based Prediction of Single Traits in Multienvironment Breeding Trials. *Frontiers in Plant Science*. 10: 1168. DOI: 10.3389/fgene.2019.01168
- Crossa, J., Pérez-rodríguez, P., Cuevas, J., Montesinos-lópez, O., Jarquín, D., Campos, G. D. L., Burgueño, J., González-camacho, J. M., Pérez-elizalde, S., Beyene, Y., Dreisigacker, S., Singh, R., Zhang, X., Gowda, M., Roorkiwal, M., Rutkoski, J., and Varshney, R. K. 2017. Genomic Selection in Plant Breeding : Methods , Models , and Perspectives. *Trends in Plant Science*. 22(11): 961–975. DOI: 10.1016/j.tplants.2017.08.011
- Cui, Y., Li, R., Li, G., Zhang, F., Zhu, T., Zhang, Q., Ali, J., Li, Z., and Xu, S. 2019. Hybrid breeding of rice via genomic selection. *Plant Biotechnology Journal*. DOI: 10.1111/pbi.13170
- Daetwyler, H. D., Hayden, M. J., Spangenberg, G. C., and Hayes, B. J. 2015. Selection on optimal haploid value increases genetic gain and preserves more genetic diversity relative to genomic selection. *Genetics*. 200(4): 1341–1348. DOI: 10.1534/genetics.115.178038
- De Beukelaer, H., Badke, Y., Fack, V., and De Meyer, G. 2017. Moving beyond managing realized genomic relationship in long-term genomic selection. *Genetics*. 206(2): 1127–1138. DOI: 10.1534/genetics.116.194449
- de los Campos, G., Vazquez, A. I., Hsu, S., and Lello, L. 2018. Complex-Trait Prediction in the Era of Big Data. *Trends in Genetics*. 34(10): 746–754. DOI: <https://doi.org/10.1016/j.tig.2018.07.004>
- de Rezende Neves, H. H., Carneiro, R., and de Queiroz, S. A. 2018. Trait-specific long-term consequences of genomic selection in beef cattle. *Genetica*. 146(1): 85–99. DOI: 10.1007/s10709-017-9999-1
- Desta, Z. A., and Ortiz, R. 2014. Genomic selection: genome-wide prediction in plant improvement. *Trends in Plant Science*. 19(9): 592–601. DOI: 10.1016/j.tplants.2014.05.006
- Dimitrijevic, A., and Horn, R. 2018. Sunflower hybrid breeding: From markers to genomic selection. In *Frontiers in Plant Science* (Vol. 8, p. 2238). Frontiers Media S.A. DOI: 10.3389/fpls.2017.02238
- Doekes, H. P., Veerkamp, R. F., Bijma, P., Hiemstra, S. J., and Windig, J. J. 2018. Trends in genome-wide and region-specific genetic diversity in the Dutch-Flemish Holstein-Friesian breeding program from 1986 to 2015. *Genetics Selection Evolution*. 50(1). DOI: 10.1186/s12711-018-0385-y
- Dong, H., Wang, R., Yuan, Y., Anderson, J., Pumphrey, M. O., Zhang, Z., and Chen, J. 2018. Evaluation of the Potential for Genomic Selection to Improve Spring Wheat Resistance to Fusarium Head Blight in

- the Pacific Northwest. *Frontiers in Plant Science*. 9(July): 911. DOI: 10.3389/FPLS.2018.00911
- dos Santos, J. P. R., Fernandes, S. B., McCoy, S., Lozano, R., Brown, P. J., Leakey, A. D. B., Buckler, E. S., Garcia, A. A. F., and Gore, M. A. 2020. Novel Bayesian Networks for Genomic Prediction of Developmental Traits in Biomass Sorghum. *G3: Genes/Genomes/Genetics*. 10(2): 769–781. DOI: 10.1534/g3.119.400759
- Endelman, J. B. 2011. Ridge regression and other kernels for genomic selection in the R package rrBLUP. *The Plant Genome*. 4(3): 250–255. DOI: 10.3835/plantgenome2011.08.0024
- Endelman, J. B., Atlin, G. N., Beyene, Y., Semagn, K., Zhang, X., Sorrells, M. E., and Jannink, J.-L. 2014. Optimal Design of Preliminary Yield Trials with Genome-Wide Markers. *Crop Science*. 54(1): 48–59. DOI: 10.2135/cropsci2013.03.0154
- Erbe, M., Pimentel, C. G., Sharifi, A. R., and Simianer, H. 2010. Assessment of cross-validation strategies for genomic prediction in cattle. Dissertation. Georg-August University
- Fernandes, S. B., Dias, K. O. G., Ferreira, D. F., and Brown, P. J. 2018. Efficiency of multi-trait, indirect, and trait-assisted genomic selection for improvement of biomass sorghum. *Theor Appl Genet*. 131(3): 747–755. DOI: 10.1007/s00122-017-3033-y
- Fragomeni, B. O., Lourenco, D. A. L., Masuda, Y., Legarra, A., and Misztal, I. 2017. Incorporation of causative quantitative trait nucleotides in single-step GBLUP. *Genetics Selection Evolution*. 49(1): 59. DOI: 10.1186/s12711-017-0335-0
- Friedman, J. H. 2001. Greedy function approximation: A gradient boosting machine. *Annals of Statistics*. 29(5): 1189–1232. DOI: 10.2307/2699986
- Ganal, M. W., Plieske, J., Hohmeyer, A., Polley, A., and Röder, M. S. 2019. High-Throughput Genotyping for Cereal Research and Breeding. *Applications of Genetic and Genomic Research in Cereals*. DOI: 10.1016/b978-0-08-102163-7.00001-6
- García-Cortés, L. A., and Sorensen, D. 1996. On a multivariate implementation of the Gibbs sampler. *Genetics Selection Evolution*. 28(1): 121–126. DOI: 10.1186/1297-9686-28-1-121
- Garretsen, F., and Keuls, M. 1978. A general method for the analysis of genetic variation in complete and incomplete diallels and North Carolina II (NC II) designs. Part II. Procedures and general formulas for the fixed model. *Euphytica*. 27(1): 49–68. DOI: 10.1007/BF00039119
- Gezan, S. A., Osorio, L. F., Verma, S., and Whitaker, V. M. 2017. An experimental validation of genomic selection in octoploid strawberry. *Horticulture Research*. 4(1): 1–9. DOI: 10.1038/hortres.2016.70
- Gianola, D., Okut, H., Weigel, K. A., and Rosa, G. J. 2011. Predicting complex quantitative traits with Bayesian neural networks: a case study with Jersey cows and wheat. *BMC Genetics*. 12(1): 87. DOI: 10.1186/1471-2156-12-87
- Goddard, M. 2009. Genomic selection: prediction of accuracy and maximisation of long term response. *Genetica*. 136(2): 245–257. DOI: 10.1007/s10709-008-9308-0
- Goddard, M. E., and Hayes, B. J. 2007. Genomic selection. *J. Anim. Breed. Genet*. 124: 323–330. DOI: 10.1111/j.1439-0388.2007.00702.x
- Goiffon, M., Kusmec, A., Wang, L., Hu, G., and Schnable, P. S. 2017. Improving response in genomic selection with a population-based selection strategy: Optimal population value selection. *Genetics*. 206(3): 1675–1682. DOI: 10.1534/genetics.116.197103
- González-Camacho, J. M., de los Campos, G., Pérez, P., Gianola, D., Cairns, J. E., Mahuku, G., Babu, R., and Crossa, J. 2012. Genome-enabled prediction of genetic values using radial basis function neural networks. *Theoretical and Applied Genetics*. 125(4): 759–771. DOI: 10.1007/s00122-012-1868-9
- González-camacho, J. M., Ornella, L., Pérez-rodríguez, P., Gianola, D., Dreisigacker, S., and Crossa, J. 2018. Applications of Machine Learning Methods to Genomic Selection in Breeding Wheat for Rust Resistance. *Plant Genome*. 11(2): 1–15. DOI: 10.3835/plantgenome2017.11.0104
- González-Camacho, Juan Manuel, Crossa, J., Pérez-Rodríguez, P., Ornella, L., and Gianola, D. 2016. Genome-enabled prediction using probabilistic neural network classifiers. *BMC Genomics*. 17(1): 208.

DOI: 10.1186/s12864-016-2553-1

- Gonzalez-Recio, O., Coffey, M. P., and Pryce, J. E. 2014. On the value of the phenotypes in the genomic era. *Journal of Dairy Science*. 97(12): 7905–7915. DOI: 10.3168/jds.2014-8125
- González-Recio, O., Rosa, G. J. M., and Gianola, D. 2014. Machine learning methods and predictive ability metrics for genome-wide prediction of complex traits. *Livestock Science*. 166: 217–231. DOI: 10.1016/j.livsci.2014.05.036
- Gorjanc, G., Gaynor, R. C., and Hickey, J. M. 2018. Optimal cross selection for long - term genetic gain in two - part programs with rapid recurrent genomic selection. *Theoretical and Applied Genetics*. 131(9): 1953–1966. DOI: 10.1007/s00122-018-3125-3
- Grinberg, N. F., Orhobor, O. I., and King, R. D. 2018. An Evaluation of Machine-learning for Predicting Phenotype: Studies in Yeast, Rice, and Wheat. *BioRxiv*: 105528. DOI: 10.1101/105528
- Grinberg, N. F., Orhobor, O. I., King, R. D., Grinberg, N. F., Orhobor, O. I., and King, R. D. 2019. An Evaluation of Machine-learning for Predicting Phenotype: Studies in Yeast, Rice, and Wheat. *Machine Learning*. 109: 251–277.
- Groeneveld, E. 1994. REML VCE - a multivariate multimodel restricted maximum likelihood (co)variance component estimation package. *Proceedings of an EC Symposium on Application of Mixed Linear Models in the Prediction of Genetic Merit in Pigs*.
- Guo, G., Zhao, F., Wang, Y., Zhang, Y., Du, L., and Su, G. 2014. Comparison of single-trait and multiple-trait genomic prediction models. *BMC Genetics*. 15(1): 1–7. DOI: 10.1186/1471-2156-15-30
- Guo, T., Yu, X., Li, X., Zhang, H., Zhu, C., Flint-garcia, S., Mcmullen, M. D., Holland, J. B., Szalma, S. J., Wisser, R. J., and Yu, J. 2019. Optimal Designs for Genomic Selection in Hybrid Crops. *Molecular Plant*. 12(3): 390–401. DOI: 10.1016/j.molp.2018.12.022
- Hadfield, J. 2010. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Jstatsoft.Org*.
- Haile, J. K., N'Diaye, A., Clarke, F., Clarke, J., Knox, R., Rutkoski, J., Bassi, F. M., and Pozniak, C. J. 2018. Genomic selection for grain yield and quality traits in durum wheat. *Molecular Breeding*. 38(6): 1–18. DOI: 10.1007/s11032-018-0818-x
- Hall, B. 2018. *LaplacesDemon: An R Package for Bayesian Inference*.
- Hassan, M. A., Yang, M., Fu, L., Rasheed, A., Zheng, B., Xia, X., Xiao, Y., and He, Z. 2019. Accuracy assessment of plant height using an unmanned aerial vehicle for quantitative genomic analysis in bread wheat. *Plant Methods*. 15(1): 37. DOI: 10.1186/s13007-019-0419-7
- Hastie, T., Tibshirani, R., and Friedman, J. 2009. *The elements of statistical learning: data mining, inference, and prediction*.
- Hayes, B., and Goddard, M. 2010. Genome-wide association and genomic selection in animal breeding. *Genome*. 53(11): 876–883. DOI: g10-076 [pii] 10.1139/G10-076
- Hayes, B. J., Bowman, P. J., Chamberlain, A. J., and Goddard, M. E. 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. *Journal of Dairy Science*. 92(2): 433–443. DOI: 10.3168/jds.2008-1646
- Heffner, E. L., Lorenz, A. J., Jannink, J. L., and Sorrells, M. E. 2010. Plant breeding with Genomic selection: Gain per unit time and cost. *Crop Science*. 50(5): 1681–1690. DOI: 10.2135/cropsci2009.11.0662
- Heffner, E. L., Sorrells, M. E., and Jannink, J. 2009. Genomic Selection for Crop Improvement. *Crop Science*. 49(Feb.): 1–12. DOI: 10.2135/cropsci2008.08.0512
- Henderson, C. R. 1975a. Best Linear Unbiased Estimation and Prediction under a Selection Model. *Biometrics*. 31(2): 423–447.
- Henderson, C. R. 1975b. Comparison of Alternative Sire Evaluation Methods. *J. Anim Sci.* 41(3): 760–770.
- Henderson, C. R. 1976. Simple Method for Computing Inverse of a Numerator Relationship Matrix Used in Prediction of Breeding Values. *Biometrics*. 32(1): 69–83.
- Henryon, M., Berg, P., and Sørensen, A. C. 2014. Animal-breeding schemes using genomic information

- need breeding plans designed to maximise long-term genetic gains. *Livestock Science*. 166: 38–47. DOI: 10.1016/j.livsci.2014.06.016
- Henryon, Mark, Liu, H., Berg, P., Su, G., Nielsen, H. M., Gebregiawergis, G. T., and Sørensen, A. C. 2019. Pedigree relationships to control inbreeding in optimum - contribution selection realise more genetic gain than genomic relationships. *Genetics Selection Evolution*: 39: 1–12. DOI: 10.1186/s12711-019-0475-5
- Hernandez, C. O., Wyatt, L. E., and Mazourek, M. R. 2020. Genomic Prediction and Selection for Fruit Traits in Winter Squash. *G3: Genes, Genomes, Genetics*. 10(10): g3.401215.2020. DOI: 10.1534/g3.120.401215
- Heslot, N., Jannink, J.-L., and Sorrells, M. E. 2015. Perspectives for Genomic Selection Applications and Research in Plants. *Crop Science*. 55(1): 1. DOI: 10.2135/cropsci2014.03.0249
- Hickey, J. M., Chiurugwi, T., Mackay, I., and Powell, W. 2017. Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery. *Nature Genetics* 49(9): 1297–1303. DOI: 10.1038/ng.3920
- Howard, R., Carriquiry, A. L., and Beavis, W. D. 2014. Parametric and Nonparametric Statistical Methods for Genomic Selection of Traits with Additive and Epistatic Genetic Architectures. In *G3: Genes, Genomes, Genetics*. 4(6): 1027–1046. DOI: 10.1534/g3.114.010298
- Jannink, J.-L., Lorenz, A. J., and Iwata, H. 2010. Genomic selection in plant breeding: from theory to practice. *Briefings in Functional Genomics*. 9(2): 166–177. DOI: 10.1093/bfpg/elq001
- Jannink, J. L. 2010. Dynamics of long-term genomic selection. *Genet Sel Evol*. 42: 35. DOI: 1297-9686-42-35
- Jia, Y., and Jannink, J. L. 2012. Multiple-trait genomic selection methods increase genetic value prediction accuracy. *Genetics*. 192(4): 1513–1522. DOI: genetics.112.144246 [pii]10.1534/genetics.112.144246
- Jiang, C., and Zeng, Z. B. 1995. Multiple trait analysis of genetic mapping for quantitative trait loci. *Genetics*. 140(3): 1111–1127.
- Jiang, J., Zhang, Q., Ma, L., Li, J., Wang, Z., and Liu, J. F. 2015. Joint prediction of multiple quantitative traits using a Bayesian multivariate antedependence model. *Heredity*. 115(1): 29–36. DOI: 10.1038/hdy.2015.9
- Jiang, Y., Schmidt, R. H., Zhao, Y., and Reif, J. C. 2017. Quantitative genetic framework highlights the role of epistatic effects for grain-yield heterosis in bread wheat. *Nature Genetics*. 49(12): 1741–1746. DOI: 10.1038/ng.3974
- Jonas, E., and De Koning, D. J. 2013. Does genomic selection have a future in plant breeding? *Trends in Biotechnology*. 31(9): 497–504. DOI: 10.1016/j.tibtech.2013.06.003
- Kang, H. M., Zaitlen, N. A., Wade, C. M., Kirby, A., Heckerman, D., Daly, M. J., and Eskin, E. 2008. Efficient control of population structure in model organism association mapping. *Genetics*. 178: 1709–1723. DOI: 10.1534/genetics.107.080101
- Kärkkäinen, H. P., and Sillanpää, M. J. 2012. Back to basics for Bayesian model building in genomic selection. *Genetics*. 191(3): 969–987. DOI: 10.1534/genetics.112.139014
- Khaki, S., and Wang, L. 2019. *Crop Yield Prediction Using Deep Neural Networks*. *Frontiers in Plant Science*. 10(May): 1–10. DOI: 10.3389/fpls.2019.00621
- Larkin, D. L., Lozada, D. N., and Mason, R. E. 2019. Genomic Selection — Considerations for Successful Implementation in Wheat Breeding Programs. *Agronomy*. 9(9): 1–18.
- Legarra, A., Aguilar, I., and Misztal, I. 2009. A relationship matrix including full pedigree and genomic information. *Journal of Dairy Science*. 92(9): 4656–4663. DOI: 10.3168/jds.2009-2061
- Li, B., Zhang, N., Wang, Y., George, A. W., and Reverter, A. 2018. Genomic Prediction of Breeding Values Using a Subset of SNPs Identified by Three Machine Learning Methods. *Frontiers in Genetics*. 9(July): 1–20. DOI: 10.3389/fgene.2018.00237
- Li, Q., Zhao, K., Bustamante, C. D., Ma, X., and Wong, W. H. 2019. Xrare: a machine learning method jointly

- modeling phenotypes and genetic evidence for rare disease diagnosis. *Genetics in Medicine*. 21(9): 2126–2134. DOI: 10.1038/s41436-019-0439-8
- Li, X., Guo, T., Mu, Q., Li, X., and Yu, J. 2018. Genomic and environmental determinants and their interplay underlying phenotypic plasticity. *Proceedings of the National Academy of Sciences of the United States of America*. 115(26): 6679–6684. DOI: 10.1073/pnas.1718326115
- Li, Z., Coffey, L., Garfin, J., Miller, N. D., White, M. R., Spalding, E. P., De Leon, N., Kaeppler, S. M., Schnable, P. S., Springer, N. M., and Hirsch, C. N. 2018. Genotype-by-environment interactions affecting heterosis in maize. *PLoS ONE*. 13(1). DOI: 10.1371/journal.pone.0191321
- Lipka, A. E., Kandianis, C. B., Hudson, M. E., Yu, J., Drnevich, J., Bradbury, P. J., and Gore, M. A. 2015. From association to prediction: statistical methods for the dissection and selection of complex traits in plants. *Current Opinion in Plant Biology*. 24: 110–118. DOI: 10.1016/j.pbi.2015.02.010
- Liu, H., Meuwissen, T. H. E., Sørensen, A. C., and Berg, P. 2015. Upweighting rare favourable alleles increases long-term genetic gain in genomic selection programs. *Genetics Selection Evolution*. 47(19): 1–14. DOI: 10.1186/s12711-015-0101-0
- Lopes, M. S., Bovenhuis, H., van Son, M., Nordbø, Ø., Grindflek, E. H., Knol, E. F., and Bastiaansen, J. W. M. 2017. Using markers with large effect in genetic and genomic predictions. *Journal of Animal Science*. 95(1): 59–71. DOI: 10.2527/jas.2016.0754
- Lorenz, A. J., Chao, S., Asoro, F. G., Heffner, E. L., Hayashi, T., Iwata, H., Smith, K. P., Sorrells, M. E., and Jannink, J.-L. 2011. Chapter 2: Genomic Selection in Plant Breeding: Knowledge and Prospects. *Advances in Agronomy*. 110: 77–123. DOI: 10.1016/B978-0-12-385531-2.00002-5
- Lorenzana, R. E., and Bernardo, R. 2009. Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. *TAG. Theoretical and Applied Genetics*. 120(1): 151–161. DOI: 10.1007/s00122-009-1166-3
- Lozada, D. N., Mason, R. E., Sarinelli, J. M., and Brown-Guedira, G. 2019. Accuracy of genomic selection for grain yield and agronomic traits in soft red winter wheat. *BMC Genetics*. 20(1): 1–12. DOI: 10.1186/s12863-019-0785-1
- Lubke, G., Laurin, C., Walters, R., Eriksson, N., Hysi, P., Spector, T., Montgomery, G., Martin, N., Medland, S., and Boomsma, D. I. 2013. Gradient Boosting as a SNP Filter: an Evaluation Using Simulated and Hair Morphology Data. *Journal of Data Mining in Genomics & Proteomics*. 4. DOI: 10.4172/2153-0602.1000143
- MacLeod, I. M., Hayes, B. J., and Goddard, M. E. 2014. The effects of demography and long-term selection on the accuracy of genomic prediction with sequence data. *Genetics*. 198(4): 1671–1684. DOI: 10.1534/genetics.114.168344
- Maier, R., Moser, G., Chen, G. B., Ripke, S., of the Psychiatric Genomics, C., Coryell, W., Potash, J. B., Scheftner, W. A., Shi, J., Weissman, M. M., Hultman, C. M., Landén, M., Levinson, D. F., Kendler, K. S., Smoller, J. W., Wray, N. R., and Lee, S. H. 2015. Joint analysis of psychiatric disorders increases accuracy of risk prediction for schizophrenia, bipolar disorder, and major depressive disorder. *Am J Hum Genet*. 96(2): 283–294. DOI: 10.1016/j.ajhg.2014.12.006
- Maier, R., Moser, G., Chen, G. B., Ripke, S., Coryell, W., Potash, J. B., Scheftner, W. A., Shi, J., Weissman, M. M., Hultman, C. M., Landén, M., Levinson, D. F., Kendler, K. S., Smoller, J. W., Wray, N. R., Lee, S. H., Absher, D., Agartz, I., Akil, H., ... Zöllner, S. 2015. Joint analysis of psychiatric disorders increases accuracy of risk prediction for schizophrenia, bipolar disorder, and major depressive disorder. *American Journal of Human Genetics*. 96(2): 283–294. DOI: 10.1016/j.ajhg.2014.12.006
- Mandal, A., Das, S. K., Karunakaran, M., Bhakat, C., and Dutta, T. K. 2013. Application of Genomic Selection in Livestock Improvement. *Research and Reviews: Journal of Dairy Science and Technology*. 2(3): 1–6.
- Maria, R., Resende, S., and Casler, M. D. 2014. Genomic Selection in Forage Breeding : Accuracy and Methods. *Crop Science*. 54(1): 1-4. DOI: 10.2135/cropsci2013.05.0353

- Massman, J. M., Gordillo, A., Lorenzana, R. E., and Bernardo, R. 2013. Genomewide predictions from maize single-cross data. *Theoretical and Applied Genetics*. 126(1): 13–22. DOI: 10.1007/s00122-012-1955-y
- Massman, J. M., Jung, H. G., and Bernardo, R. 2013. Genomewide Selection versus Marker-assisted Recurrent Selection to Improve Grain Yield and Stover-quality Traits for Cellulosic Ethanol in Maize. *Crop Science*. 53(1): 58–66. DOI: 10.2135/cropsci2012.02.0112
- Meuwissen, T. H. 1997. Maximizing the response of selection with a predefined rate of inbreeding. *J Anim Sci*. 75(4): 934–940.
- Meuwissen, T. H., Hayes, B. J., and Goddard, M. E. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*. 157(4): 1819–1829.
- Meuwissen, T., Hayes, B., and Goddard, M. 2013. Accelerating Improvement of Livestock with Genomic Selection. *Annu Rev Anim Biosci*: 221–239. DOI: 10.1146/annurev-animal-031412-103705
- Miguel, P., and Zingaretti, L. M. 2019. A Guide on Deep Learning for Complex Trait. *Genes*. 10(7): 553
- Misra, B. B., Langefeld, C. D., Olivier, M., and Cox, L. A. 2018. Integrated Omics: Tools, Advances, and Future Approaches. *Journal of Molecular Endocrinology*. DOI: 10.1530/JME-18-0055
- Misztal, I., Tsuruta, S., Strabel, T., Auvray, B., Druet, T., and Lee, D. H. 2002. BLUPF90 AND RELATED PROGRAMS (BGF90). *Proceedings from the 7th World Congress on Genetics Applied to Livestock Production*. 33: 743–744.
- Moeiniazade, S., Hu, G., Wang, L., and Schnable, P. S. 2019. Optimizing selection and mating in genomic selection with a look-ahead approach: An operations research framework. *G3: Genes, Genomes, Genetics*. 9(7): 2123–2133. DOI: 10.1534/g3.118.200842
- Montesinos-lópez, A., Montesinos-lópez, O. A., Gianola, D., Crossa, J., and Hernández-suárez, C. M. 2018. Multi-environment Genomic Prediction of Plant Traits Using Deep Learners With Dense Architecture. *G3: Genes, Genomes, Genetics*. 8(12): 3813–3828. DOI: 10.1534/g3.118.200740
- Montesinos-Lopez, O. A., Martin-Vallejo, J., Crossa, J., Gianola, D., Hernandez-Suarez, C. M., Montesinos-Lopez, A., Juliana, P., and Singh, R. 2019a. New Deep Learning Genomic-Based Prediction Model for Multiple Traits with Binary, Ordinal, and Continuous Phenotypes. *G3: Genes, Genomes, Genetics*. 9(5): 1545–1556. DOI: 10.1534/g3.119.300585
- Montesinos-López, O. A., Montesinos-López, A., Crossa, J., Cuevas, J., Montesinos-López, J. C., Gutiérrez, Z. S., Lillemo, M., Philomin, J., and Singh, R. 2019b. A Bayesian Genomic Multi-output Regressor Stacking Model for Predicting Multi-trait Multi-environment Plant Breeding Data. *G3: Genes, Genomes, Genetics*. 9(10): 3381–3393. DOI: 10.1534/g3.119.400336
- Montesinos-Lopez, O. A., Montesinos-Lopez, A., Crossa, J., Toledo, F. H., Perez-Hernandez, O., Eskridge, K. M., and Rutkoski, J. 2016. A Genomic Bayesian Multi-trait and Multi-environment Model. *G3: Genes, Genomes, Genetics*. 6(9): 2725–2744. DOI: 10.1534/g3.116.032359
- Montesinos-Lopez, O. A., Montesinos-Lopez, A., Luna-Vazquez, F. J., Toledo, F. H., Perez-Rodriguez, P., Lillemo, M., and Crossa, J. 2019c. An R Package for Bayesian Analysis of Multi-environment and Multi-trait Multi-environment Data for Genome-Based Prediction. *G3: Genes, Genomes, Genetics*. 9(5): 1355–1369. DOI: 10.1534/g3.119.400126
- Montesinos-López, O. A., Montesinos-López, A., Tuberosa, R., Maccaferri, M., Sciara, G., Ammar, K., and Crossa, J. 2019d. Multi-Trait, Multi-Environment Genomic Prediction of Durum Wheat With Genomic Best Linear Unbiased Predictor and Deep Learning Methods. *Frontiers in Plant Science*. 10: 1311. DOI: 10.3389/fpls.2019.01311
- Moore, J. K., Manmathan, H. K., Anderson, V. A., Poland, J. A., Morris, C. F., and Haley, S. D. 2017. Improving Genomic Prediction for Pre-Harvest Sprouting Tolerance in Wheat by Weighting Large-Effect Quantitative Trait Loci. *Crop Science*. 57(3): 1315–1324. DOI: 10.2135/cropsci2016.06.0453
- Moreira, F. F., Oliveira, H. R., Volenec, J. J., Rainey, K. M., and Brito, L. F. 2020. Integrating High-Throughput Phenotyping and Statistical Genomic Methods to Genetically Improve Longitudinal Traits in Crops.

- Frontiers in Plant Science*. 11: 681. DOI: 10.3389/fpls.2020.00681
- Mukhopadhyay, C. S., and Kumar, D. 2008. *Genomic Selection : Its Prospects and Challenges*. Project Directorate on Cattle. Military Farm School and Research Centre, Meerut Cantt, UP, India
- Müller, D., Schopp, P., and Melchinger, A. E. 2018a. Selection on expected maximum haploid breeding values can increase genetic gain in recurrent genomic selection. *G3: Genes, Genomes, Genetics*. 8(4): 1173–1181. DOI: 10.1534/g3.118.200091
- Müller, D., Schopp, P., and Melchinger, A. E. 2018b. Selection on expected maximum haploid breeding values can increase genetic gain in recurrent genomic selection. *G3: Genes, Genomes, Genetics*. 8(4): 1173–1181. DOI: 10.1534/g3.118.200091
- Nakaya, A., and Isobe, S. N. 2017. Will genomic selection be a practical method for plant breeding? *Ann Bot*. 110(6): 1303–1316. DOI: mcs109 [pii] 10.1093/aob/mcs109
- Odilbekov, F., Armoniené, R., Koc, A., Svensson, J., and Chawade, A. 2019. GWAS-Assisted Genomic Prediction to Predict Resistance to *Septoria Tritici* Blotch in Nordic Winter Wheat at Seedling Stage. *Frontiers in Genetics*. 10: 1224. DOI: 10.3389/fgene.2019.01224
- Pérez-Rodríguez, P., Gianola, D., González-Camacho, J. M., Crossa, J., Manès, Y., and Dreisigacker, S. 2012. Comparison between linear and non-parametric regression models for genome-enabled prediction in wheat. *G3: Genes, Genomes, Genetics*. 2(12): 1595–1605. DOI: 10.1534/g3.112.003665
- Pérez, P., and De Los Campos, G. 2004. BGLR: A Statistical Package for Whole Genome Regression and Prediction.
- Raymond, B., Bouwman, A. C., Schrooten, C., Houwing-Duistermaat, J., and Veerkamp, R. F. 2018. Utility of whole-genome sequence data for across-breed genomic prediction. *Genetics Selection Evolution*. 50(1). DOI: 10.1186/s12711-018-0396-8
- Resende, M. F. R., Muñoz, P., Resende, M. D. V., Garrick, D. J., Fernando, R. L., Davis, J. M., Jokela, E. J., Martin, T. A., Peter, G. F., and Kirst, M. 2012. Accuracy of Genomic Selection Methods in a Standard Data Set of Loblolly Pine (*Pinus taeda* L.). *Genetics*. 190(4): 1503–1510. DOI: 10.1534/genetics.111.137026
- Rice, B., and Lipka, A. E. 2019. Evaluation of RR-BLUP Genomic Selection Models that Incorporate Peak Genome-Wide Association Study Signals in Maize and Sorghum. *The Plant Genome*. 12(1): 1–14. DOI: 10.3835/plantgenome2018.07.0052
- Riedelsheimer, C., Czedik-Eysenberg, A., Grieder, C., Lisec, J., Technow, F., Sulpice, R., Altmann, T., Stitt, M., Willmitzer, L., and Melchinger, A. E. 2012. Genomic and metabolic prediction of complex heterotic traits in hybrid maize. *Nat Genet*. DOI: ng.1033 [pii] 10.1038/ng.1033
- Robertson, C. D., Hjortshøj, R. L., and Janss, L. L. 2019. Genomic Selection in Cereal Breeding. *Agronomy*. 9(2): 1–16. DOI: 10.3390/agronomy9020095
- Runcie, D., and Cheng, H. 2019. Pitfalls and remedies for cross validation with multi-trait genomic prediction methods. *G3: Genes, Genomes, Genetics*. 9(11): 3727–3741. DOI: 10.1534/g3.119.400598
- Rutkoski, J. E. 2019. A practical guide to genetic gain. In D. L. B. T.-A. in A. Sparks (Ed.), *Advances in Agronomy*. 157: 217–249. Academic Press. DOI: https://doi.org/10.1016/bs.agron.2019.05.001
- Samuel, A. L. 1959. Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*. 3(3): 210–229.
- Schulthess, A. W., Wang, Y., Miedaner, T., Wilde, P., Reif, J. C., and Zhao, Y. 2016. Multiple-trait- and selection indices-genomic predictions for grain yield and protein content in rye for feeding purposes. *Theor Appl Genet*. 129(2): 273–287. DOI: 10.1007/s00122-015-2626-6
- Schulthess, A. W., Zhao, Y., Longin, C. F. H., and Reif, J. C. 2018. Advantages and limitations of multiple-trait genomic prediction for Fusarium head blight severity in hybrid wheat (*Triticum aestivum* L.). *Theor Appl Genet*. 131(3): 685–701. DOI: 10.1007/s00122-017-3029-7
- Sonesson, A. K., Woolliams, J. A., and Meuwissen, T. H. 2012. Genomic selection requires genomic control of inbreeding. *Genet Sel Evol*. 44: 27. DOI: 1297-9686-44-27 [pii] 10.1186/1297-9686-44-27

- Sousa, T. V., Caixeta, E. T., Alkimim, E. R., Oliveira, A. C. B., Pereira, A. A., Sakiyama, N. S., Zambolim, L., and Resende, M. D. V. 2019. Early Selection Enabled by the Implementation of Genomic Selection in *Coffea arabica* Breeding. *Frontiers in Plant Science*. 9: 1934. DOI: 10.3389/fpls.2018.01934
- Speed, D., and Balding, D. J. 2014. MultiBLUP: Improved SNP-based prediction for complex traits. *Genome Research*. 24(9): 1550–1557. DOI: 10.1101/gr.169375.113
- Sperschneider, J. 2019. Machine learning in plant–pathogen interactions: empowering biological predictions from field scale to genome scale. *New Phytologist*: nph.15771. DOI: 10.1111/nph.15771
- Spindel, J. E., Begum, H., Akdemir, D., Collard, B., Redoña, E., Jannink, J., and Mccouch, S. 2016. Genome-wide prediction models that incorporate de novo GWAS are a powerful new tool for tropical rice improvement. *Heredity*. 116(4): 395–408. DOI: 10.1038/hdy.2015.113
- Sun, C., and VanRaden, P. M. 2014. Increasing Long-Term Response by Selecting for Favorable Minor Alleles. *PLoS ONE*. 9(2): e88510. DOI: 10.1371/journal.pone.0088510
- Svane, S. F., Jensen, C. S., and Thorup-Kristensen, K. 2019. Construction of a large-scale semi-field facility to study genotypic differences in deep root growth and resources acquisition. *Plant Methods*. 15(1). DOI: 10.1186/s13007-019-0409-9
- Tan, B., Grattapaglia, D., Martins, G. S., Ferreira, K. Z., and Sundberg, B. 2017. Evaluating the accuracy of genomic prediction of growth and wood traits in two Eucalyptus species and their F 1 hybrids. *BMC Plant biology*. 17(1): 110. DOI: 10.1186/s12870-017-1059-6
- Tan, B., and Ingvarsson, P. K. 2019. Integrating genome-wide association mapping of additive and dominance genetic effects to improve genomic prediction accuracy in Eucalyptus. *BioRxiv*: 841049. DOI: 10.1101/841049
- Tang, Y., Liu, X., Wang, J., Li, M., Wang, Q., Tian, F., Su, Z., Pan, Y., Liu, D., Lipka, A. E., Buckler, E. S. E. S., Zhang, Z., and Zhang, Z. 2016. GAPIT Version 2: An Enhanced Integrated Tool for Genomic Association and Prediction. *The Plant Genome*. 9(2): 1–9. DOI: 10.3835/plantgenome2015.11.0120
- Technow, F., Schrag, T. A., Schipprack, W., Bauer, E., Simianer, H., and Melchinger, A. E. 2014. Genome Properties and Prospects of Genomic Prediction of Hybrid Performance in a Breeding Program of Maize. *Genetics*. 197(4): 1343–1355. DOI: 10.1534/genetics.114.165860
- Tempelman, R. J. 2015. Statistical and Computational Challenges in Whole Genome Prediction and Genome-Wide Association Analyses for Plant and Animal Breeding. *Journal of Agricultural, Biological, and Environmental Statistics*. 20(4): 442–466. DOI: 10.1007/s13253-015-0225-2
- Thomas, S., Kuska, M. T., Bohnenkamp, D., Brugger, A., Alisaac, E., Wahabzada, M., Behmann, J., and Mahlein, A. K. 2018. Benefits of hyperspectral imaging for plant disease detection and plant protection: a technical perspective. *Journal of Plant Diseases and Protection*. 125(1): 5-20. DOI: 10.1007/s41348-017-0124-6
- Tier, B. 2010. Genomic selection: promises and propriety. *J Anim Breed Genet*. 127(3): 169–170. DOI: JBG876 [pii]10.1111/j.1439-0388.2010.00876.x
- Tipping, M. E., and Faul, A. C. 2003. Fast Marginal Likelihood Maximisation for Sparse Bayesian Models. *AISTATS*
- VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. *Journal of Dairy Science*. 91: 4414–4423. DOI: 10.3168/jds.2007-0980
- Varshney, R. K., Roorkiwal, M., and Sorrells, M. E. 2017. Genomic selection for crop improvement: New molecular breeding strategies for crop improvement. *Genomic Selection for Crop Improvement: New Molecular Breeding Strategies for Crop Improvement*. Springer International Publishing. DOI: 10.1007/978-3-319-63170-7
- Velazco, J. G., Malosetti, M., Hunt, C. H., Mace, E. S., Jordan, D. R., and van Eeuwijk, F. A. 2019. Combining pedigree and genomic information to improve prediction quality: an example in sorghum. *Theoretical and Applied Genetics*. 132(7): 2055–2067. DOI: 10.1007/s00122-019-03337-w
- Volpato, L., Alves, R. S., Teodoro, P. E., de Resende, M. D., Nascimento, M., Nascimento, A. C. C., Ludke,

- W. H., da Silva, F., and Borem, A. 2019. Multi-trait multi-environment models in the genetic selection of segregating soybean progeny. *PLoS One*. 14(4): e0215315. DOI: 10.1371/journal.pone.0215315
- Voss-Fels, K. P., Cooper, M., and Hayes, B. J. 2019. Accelerating crop genetic gains with genomic selection. *Theoretical and Applied Genetics*. 132(3): 669–686. DOI: 10.1007/s00122-018-3270-8
- Waldmann, P. 2016. Genome-wide prediction using Bayesian additive regression trees. *Genetics Selection Evolution*. 48(1): 42. DOI: 10.1186/s12711-016-0219-8
- Wang, J., Zhou, Z., Zhang, Z., Li, H., Liu, D., Zhang, Q., Bradbury, P. J., Buckler, E. S., and Zhang, Z. 2018. Expanding the BLUP alphabet for genomic prediction adaptable to the genetic architectures of complex traits. *Heredity*. 121(6): 648–662. DOI: 10.1038/s41437-018-0075-0
- Wang, Q., Tian, F., Pan, Y., Buckler, E. S. E. S., and Zhang, Z. 2014. A SUPER Powerful Method for Genome Wide Association Study. *PLoS ONE*. 9(9): e107684. DOI: 10.1371/journal.pone.0107684
- Wang, X., Li, L., Yang, Z., Zheng, X., Yu, S., Xu, C., and Hu, Z. 2017. Predicting rice hybrid performance using univariate and multivariate GBLUP models based on North Carolina mating design II. *Heredity*. 118(3): 302–310. DOI: 10.1038/hdy.2016.87
- Wang, Xin, Xu, Y., Hu, Z., and Xu, C. 2018. Genomic selection methods for crop improvement: Current status and prospects. *Crop Journal*. 6(4): 330–340. DOI: 10.1016/j.cj.2018.03.001
- Watson, A., Hickey, L. T., Christopher, J., Rutkoski, J., Poland, J., and Hayes, B. J. 2019. Multivariate Genomic Selection and Potential of Rapid Indirect Selection with Speed Breeding in Spring Wheat. *Crop Science*. 59(5): 1945–1959. DOI: 10.2135/cropsci2018.12.0757
- Weller, J. I., and Ron, M. 2010. Invited review: quantitative trait nucleotide determination in the era of genomic selection. *J Dairy Sci*. 94(3): 1082–1090. DOI: S0022-0302(11)00063-4 [pii] 10.3168/jds.2010-3793
- Wolfe, M. D., Del Carpio, D. P., Alabi, O., Ezenwaka, L. C., Ikeogu, U. N., Kayondo, I. S., Lozano, R., Okeke, U. G., Ozimati, A. A., Williams, E., Egesi, C., Kawuki, R. S., Kulakow, P., Rabbi, I. Y., and Jannink, J.-L. 2017. Prospects for Genomic Selection in Cassava Breeding. *The Plant Genome*. 10(3): 1–19. DOI: 10.3835/plantgenome2017.03.0015
- Wright, S. 1922. Coefficients of Inbreeding and Relationship. *The American Naturalist*. 56(645): 330–338. DOI: 10.1086/279872
- Xu, C. W., Wang, X. F., Li, Z. K., and Xu, S. Z. 2009. Mapping QTL for multiple traits using Bayesian statistics. *Genetics Research*. 91(1): 23–37. DOI: 10.1017/S0016672308009956
- Yabe, S., Yamasaki, M., Ebana, K., Hayashi, T., and Iwata, H. 2016. Island-Model Genomic Selection for Long-Term Genetic Improvement of Autogamous Crops. *PLoS one*. 11(4): 1–21. DOI: 10.1371/journal.pone.0153945
- Zampieri, M., Sekar, K., Zamboni, N., and Sauer, U. 2017. Frontiers of high-throughput metabolomics. *Current Opinion in Chemical Biology*. 36: 15–23. DOI: 10.1016/j.cbpa.2016.12.006
- Zeng, P., and Zhou, X. 2017. Non-parametric genetic prediction of complex traits with latent Dirichlet process regression models. *Nature Communications*. 8(1). DOI: 10.1038/s41467-017-00470-2
- Zhang, X., Lourenco, D., Aguilar, I., Legarra, A., and Misztal, I. 2016. Weighting strategies for single-step genomic BLUP: An iterative approach for accurate calculation of GEBV and GWAS. *Frontiers in Genetics*. 7: 151. DOI: 10.3389/fgene.2016.00151
- Zhang, Z., Todhunter, R. J., Buckler, E. S., and Van Vleck, L. D. 2007. Technical note: Use of marker-based relationships with multiple-trait derivative-free restricted maximal likelihood. *Journal of Animal Science*. 85(4): 881–885. DOI: 10.2527/jas.2006-656
- Zhang, Zhe, Ding, X., Liu, J., de Koning, D.-J., and Zhang, Q. 2011. Genomic selection for QTL-MAS data using a trait-specific relationship matrix. *BMC Proceedings*. 5(S3): S15. DOI: 10.1186/1753-6561-5-S3-S15
- Zhang, Zhe, Liu, J., Ding, X., Bijma, P., de Koning, D. J., and Zhang, Q. 2010. Best linear unbiased prediction of genomic breeding values using a trait-specific marker-derived relationship matrix. *PLoS ONE*. 5(9):

- 1–8. DOI: 10.1371/journal.pone.0012648
- Zhang, Zhe, Ober, U., Erbe, M., Zhang, H., Gao, N., He, J., Li, J., and Simianer, H. 2014. Improving the Accuracy of Whole Genome Prediction for Complex Traits Using the Results of Genome Wide Association Studies. *PloS one*. 9(3): 1–12. DOI: 10.1371/journal.pone.0093017
- Zhang, Zhiwu, Ersoz, E., Lai, C.-Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., Bradbury, P. J., Yu, J., Arnett, D. K., Ordovas, J. M., and Buckler, E. S. 2010. Mixed linear model approach adapted for genome-wide association studies. *Nature Genetics*. 42(4): 355–360. DOI: 10.1038/ng.546
- Zhao, Y., Mette, M. F., and Reif, J. C. 2015. Genomic selection in hybrid breeding. *Plant Breeding*. 134(1): 1–10. DOI: 10.1111/pbr.12231
- Zhou, G., Chen, Y., Yao, W., Zhang, C., Xie, W., Hua, J., Xing, Y., Xiao, J., and Zhang, Q. 2012. Genetic composition of yield heterosis in an elite rice hybrid. *Proceedings of the National Academy of Sciences of the United States of America*. 109(39): 15847–15852. DOI: 10.1073/pnas.1214141109
- Zhou, Y., Isabel Vales, M., Wang, A., and Zhang, Z. 2016. Systematic bias of correlation coefficient may explain negative accuracy of genomic prediction. *Briefings in Bioinformatics*. 18(5): 744-753