



Advances in multi-trait genomic prediction approaches: classification, comparative analysis, and perspectives

Alain J. Mbebi ^{1,2}, Facundo Mercado¹, David Hobby¹, Hao Tong^{1,2}, Zoran Nikoloski ^{1,2,*}

¹Bioinformatics Department, Institute of Biochemistry and Biology, University of Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam-Golm, Brandenburg, Germany

²Systems Biology and Mathematical Modeling Group, Max Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1, 14476 Potsdam-Golm, Brandenburg, Germany

*Corresponding author. E-mail: nikoloski@mpimp-golm.mpg.de

Abstract

Traits in any organism are not independent, but show considerable integration, observed in a form of couplings and trade-offs. Therefore, improvement in one trait may affect other traits, often in undesired direction. To account for this problem, crop breeding increasingly relies on multi-trait genomic prediction (MT-GP) approaches that leverage the availability of genetic markers from different populations along with advances in high-throughput precision phenotyping. While significant progress has been made to jointly model multiple traits using a variety of statistical and machine learning approaches, there is no systematic comparison of advantages and shortcomings of the existing classes of MT-GP models. Here, we fill this knowledge gap by first classifying the existing MT-GP models and briefly summarizing their general principles, modeling assumptions, and potential limitations. We then perform an extensive comparative analysis with 10 traits measured in an *Oryza sativa* diversity panel using cross-validation scenarios relevant in breeding practice. Finally, we discuss directions that can enable the building of next generation MT-GP models in addressing pressing challenges in crop breeding.

Keywords: genomic prediction; multi-trait; machine learning; deep learning; crop improvement; breeding

Introduction

The food system is facing a challenge due to interrelated effects of climate change, growing world's population, and increasing scarcity of resources [1]. Breeding of crops resilient to changes in environmental cues with little to no penalty to yield provides one way to address this challenge [2, 3]. However, complex traits, such as disease resistance and drought tolerance, are often governed by genetic architecture characterized by multiple loci of small effects, epistasis [4], and pleiotropy [5], with the latter leading to extensive trait integration and trade-offs. Thus, efficiently decipher how genetic variants influence multiple traits simultaneously, optimize selection strategies, and develop crop varieties that maintain high productivity under changing environmental conditions requires understanding of the genetic and molecular mechanisms underpinning trait integration [6].

Due to the requirement of multiple crossing, selection, and testing steps often required, traditional breeding approaches to develop varieties with improved traits of interest is costly, labor-intensive, and time-consuming [7, 8]. Despite their success in introgressing and pyramiding genes, the effectiveness of breeding techniques based on marker-assisted selection for improving traits is reduced due to the inability of this approach to account for multiple quantitative trait loci (QTL) with minor effects [9–11], QTL-environment interactions, and the usage of non-generalizable mathematical methods [12, 13]. To mitigate

these shortcomings, Meuwissen *et al.* proposed genomic selection (GS) [14] that leverages the advances in genotyping technologies with genomic prediction (GP) models that facilitate shortening of the breeding cycle and reduction in usage of resources [15, 16].

The basic principle of GS involves designing a training set of genotypes for which genotypic and phenotypic data are available (Fig. 1). Models are then trained to predict the observed traits based on genomic data (i.e. genetic markers) using diverse machine learning (ML) approaches. The resulting GP model is in turn used to predict genomic estimated breeding values (GEBVs) for a testing set or selection population, for which only genotypic data are available [17]. The performance of the GP model (i.e. its prediction accuracy) is then computed as a correlation between GEBVs and the measured phenotypes of a trait using different cross-validation (CV) schemes. Since different ML approaches can account for large and small QTL effects [18], GP has been successfully applied to improve trait selection across several major crop species, including: rice, wheat, sorghum, and corn [19–23].

With the ability to perform detailed and precise simultaneous measurements of multiple traits at large-scale and under multiple environments, high-throughput phenotyping (HTP) technologies have opened the possibility for a more comprehensive understanding of genotype-by-environment interactions [24, 25]. As a result, these technologies have propelled the development of GP models that can predict multiple traits simultaneously.

Received: January 05, 2025. Revised: March 24, 2025. Accepted: April 20, 2025

© The Author(s) 2025. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

For commercial re-use, please contact journals.permissions@oup.com

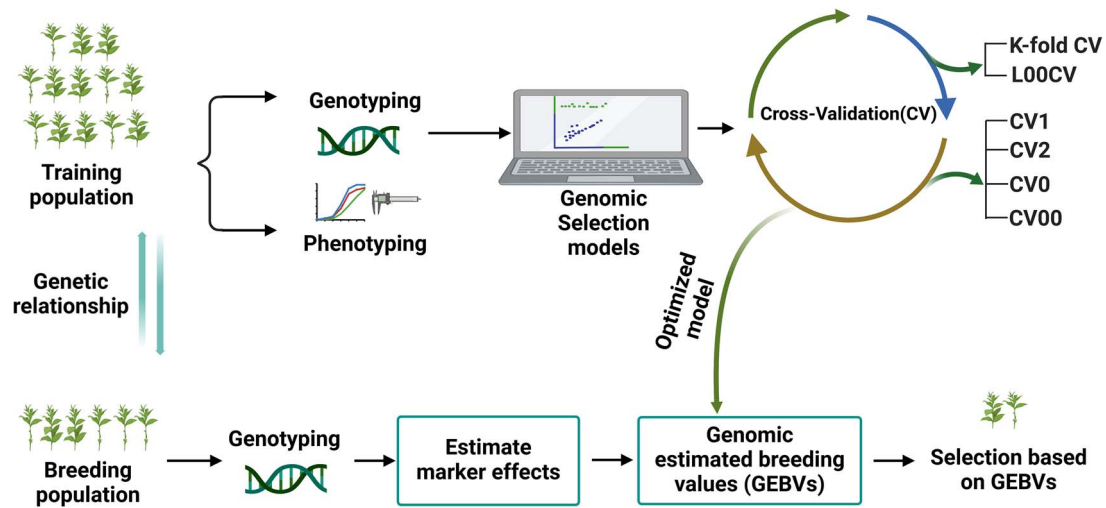


Figure 1. Schematic overview of GS. Showcased are the main steps involved in the GS, starting with the collection of phenotypic and genotypic data from a training population (e.g. inbreds or hybrids). Depending on the prediction objective and the sample size, different CV schemes along with collected data are used to train the predictive models; these models are subsequently used to determine GEBVs. The GEBVs are then applied to a testing population that is only phenotyped and from which individuals with desired performances are selected without the need for direct phenotyping. Briefly, in k -fold CV, the population under consideration is partitioned into k folds of approximately equal size; the model is trained on $(k - 1)$ folds while the remaining fold is used for validation until each fold has been used as a validation set. Leave one out CV is similar to the former except for the fact that a single individual is used for validation. On the other hand, CV0, CV00, CV1, and CV2 are employed under multiple environments settings and they correspond respectively to the prediction of seen genotypes in unseen environments, unseen genotypes in unseen environments, unseen genotypes in seen environments and genotypes seen in some environments to be predicted in other seen environments.

The existing single trait GP (ST-GP) models, including: ridge regression best linear unbiased prediction (rrBLUP) [26] and the those belonging to the Bayesian alphabet [27, 28], are not appropriate to address this problem as they neglect the genetic correlations between multiple traits [29]. Their applications with multiple traits usually involve transformations of the phenotype matrix (e.g. weighted linear combination of several traits [30] and vectorization) or training of multiple ST-GP models, for each single trait separately.

Trait integration is commonly observed in a form of couplings or trade-offs [6]. In the former, changes in one trait mimic those in another and are result of shared genetic architecture, as is the case for the functional interdependence of shoot and root growth [31]. In the case of trade-offs, improvement of one trait happens at the expense of another and is attributable to resource allocation, genetic constraints, and antagonistic pleiotropy; for instance, the well-studied relationship between plant growth and defense is an example of traits in trade-off [32]. These evidence of trait integration motivated recent developments that aimed to design GP models for multiple traits, termed MT-GP. The existing MT-GP models differ in the number of trait they can efficiently model and how population structure and marker effects are handled.

While significant progress has been made to jointly model multiple traits with a variety of statistical and machine learning approaches [33], there is no systematic comparison of advantages and shortcomings of the existing classes of MT-GP models. Here, we aim to fill this knowledge gap by first briefly summarizing the general principles, modeling assumptions, and potential limitations of the existing MT-GP models, followed by an outline of existing strategies used to assess the performance of MT-GP models. We then perform an extensive comparative analysis with different realistic CV scenarios, using 10 traits of which five are yield-related and five metabolic traits measured in an *Oryza sativa* diversity panel composed of 506 accessions characterized with genomic data on ~900k single nucleotide polymorphisms (SNPs). Finally, we discuss possible directions that can enable the

building of next generation MT-GP models in addressing pressing challenges in breeding.

Materials and methods

Main principles of MT-GP models

MT-GP models have gained significant traction over the past decade as the breeding community has sought to exploit the genetic correlation between traits to improve the accuracy and selection when simultaneously analyzing multiple traits [34, 35]. Unlike ST-GP models [29], that only account for between-individuals relationship and model single traits, MT-GP models use several traits simultaneously either by considering their matrix representation or a link function that jointly quantify the traits and/or their covariance. As a result, MT-GP models allow traits to borrow information from each other that can boost the predictability. For instance, improved prediction accuracies have been observed for traits with low heritability when modeled in the presence of correlation with high heritability traits, from which they can borrow information [36–38]. Like ST-GP, the aim of MT-GP models is to establish a mathematical relationship between multiple traits measured in a population of genotypes and the corresponding genomic data. When designing a MT-GP model, often two main scenarios are considered: (i) different traits measured in different populations of genotypes and (ii) different traits measured in the same population of genotypes, termed multitask and multiple output learning, respectively [39]. In doing so, a single model accounting for all traits of interest is trained and the between-traits and/or between-individuals correlation are explicitly modeled.

Assumptions of MT-GP models

The existing MT-GP models differ based on the assumptions regarding the underlying distribution of the data and the trade-offs between the resulting accuracy and interpretability. As a result, MT-GP models can be categorized into the following three

main groups (see Table 1): (i) parametric models that rely on specific distributional assumptions with a predefined link function between traits and markers; (ii) non-parametric approaches for which no predefined link function between phenotypes and genetic markers is defined nor strict distributional assumptions; and (iii) semi-parametric models that combine the former two [78, 79]. Most MT-GP models represent a special case of a multivariate linear mixed effects model (MLMM). Therefore, we follow the notations in [40] and consider that MLMM for s traits measured in n individuals can be written in matrix form as:

$$\mathbf{Y} = \mathbf{XB} + \mathbf{ZU} + \mathbf{E}, \quad (1)$$

with $\mathbf{Y} \in \mathbb{R}^{n \times s}$, an observed matrix of phenotypic values for s traits on n individuals, that serve as responses, $\mathbf{X} \in \mathbb{R}^{n \times p}$ a design matrix of covariates (i.e. fixed effects) including a row of 1s as intercepts for each response and corresponding to the coefficient matrix of fixed effects $\mathbf{B} \in \mathbb{R}^{p \times s}$; clearly, each column of \mathbf{B} represents the fixed effect coefficients of all covariates for a particular trait. The design matrix $\mathbf{Z} \in \mathbb{R}^{n \times r}$ contains the random effects $\mathbf{U} \in \mathbb{R}^{r \times s}$ that quantify the deviation from fixed effects such that each row of \mathbf{U} represents the random effect for a single genotype for all s traits. Finally, $\mathbf{E} \in \mathbb{R}^{n \times s}$ is the matrix of residuals for the s traits representing the part of \mathbf{Y} not explained by the model. It is further assumed that $\mathbf{E} \sim \text{MVN}(\mathbf{0}, \mathbf{I}_n, \mathbf{R})$ and $\mathbf{U} \sim \text{MVN}(\mathbf{0}, \mathbf{G}, \Sigma_g)$, with \mathbf{G} , the genomic relationship matrix (GRM) that can be derived using the method suggested in [80]. Specifically, $\mathbf{G} = \frac{\mathbf{WW}'}{2 \sum p_j(1 - p_j)}$, where \mathbf{W} is the centered SNP matrix with entries $w_{ij} = m_{ij} - 2p_j$, p_j is the allele frequency for the j^{th} SNP, m_{ij} the genotype coded 0, 1, and 2, denoting reference homozygote, heterozygote, and alternate homozygote, for the i^{th} sample and j^{th} SNP. Further, \mathbf{R} denotes the residual covariance matrix and \mathbf{I}_n the n -dimensional identity matrix. The notation $\mathbf{Y} \sim \text{MVN}(\mathbf{M}, \mathbf{U}, \mathbf{V})$, is used to denote that the matrix $\mathbf{Y} \in \mathbb{R}^{n \times s}$ follows a matrix-variate normal distribution [81] with location parameter $\mathbf{M} \in \mathbb{R}^{n \times s}$ (i.e. the expected value of \mathbf{Y}) and covariance matrices $\mathbf{U} \in \mathbb{R}^{n \times n}$ and $\mathbf{V} \in \mathbb{R}^{s \times s}$ capturing the extent to which the variables of interest co-vary.

Deriving closed form and unique solutions analytically for the maximum likelihood estimates of covariance matrices under matrix-variate normal distribution is a challenging task, particularly for unstructured covariance and high-dimensional data. In practice, this obstacle can be tackled by assuming separability of Σ the covariance between variables and samples, which amounts to writing Σ as a Kronecker product of the between and within variable covariances (i.e. $\Sigma = \mathbf{U} \otimes \mathbf{V}$), and by imposing identifiability constraints such as $\text{tr}(\mathbf{V}) = s$ (i.e., the between samples covariance is assumed to be the identity matrix) [82].

Parametric models

These type of MT-GP models rely on specific distributional assumptions with a predefined link function between traits and markers. One of the most common parametric methods is the multi-trait BLUP (MT-BLUP) [34, 58, 59], which is an extension of its ST version. To accommodate the presence of several traits and address selection bias, if correlated traits are analyzed individually [83], MT-BLUP incorporates genetic and residual correlations between traits [40]. MT-BLUP can be derived from (Eq. 1) by stacking columns of the respective matrices on top of each other (i.e. vectorization) and assuming a given probability distribution for all random effects (see, e.g. [44]). Under the assumption of same covariance for all loci and independent SNP effects, that is ultimately equivalent to assuming a common GRM

for all traits, its genomic variant (MT-GBLUP) can be derived in a similar way [42]. By extending the ST-rrBLUP constant variance assumption to constant covariance for all SNPs, its MT version (i.e. MT-rrBLUP) [45] was also proposed. In a Bayesian framework, covariance between SNPs across traits was explicitly accounted for through the introduction of latent variable in the expectation of random effects. Because of the Bayesian Markov chain Monte Carlo (MCMC) approach [84, 85] used for parameters estimation, this model is also referred to as MT-Bayesian rrBLUP.

This rather strong assumption of common \mathbf{G} for all traits was further relaxed in [60] to obtain a weighted multi-trait GBLUP (MT-wGBLUP) model in which a sparse latent variable model approach is used to estimate breeding values and SNP effects under heterogeneous SNP-covariance between genomic regions (e.g. chromosomes or group of SNPs) [86].

Nonetheless, the number of parameter to be estimated can quickly increase leading to a computational burden and unreliable estimates for unstructured covariance. To remedy this issue, MT-Bayesian models that utilize prior distributions to incorporate prior knowledge on marker effects and handle different types of genetic architectures more flexibly [87] have been proposed. We refer the interested reader to [40] for detailed explanations on how the Bayesian variants of the MT-BLUP and related models can be derived with proper priors assigned to the parameters of interest.

Additionally, to account for departure from normality, often exhibited by some traits, MT extensions of the Bayesian alphabet including: BayesA, BayesB, BayesC π , Bayesian LASSO, and Bayesian rr-BLUP have also been proposed (see [37, 41, 43]). For instance, MT-BayesC π estimates the marker effects by variable selection and assumes that each locus can have an effect on any combination of traits [42].

Another model called MT-BayesAS assumes similar genomic covariance (i.e. relationship between individuals based on shared genetic markers.) structure for SNPs within a given genomic regions, but different genomic covariance for SNPs in different genomic regions, and relies on multi-trait random regression model to explicitly model heterogeneous variance and covariance as a latent variables model [45, 86]. Additionally, MT-BayesA and MT-Bayesian LASSO that incorporate the assumption of heterogeneous covariances could also be derived from MT-BayesAS.

When information about multiple environments is available as it is often the case for genetic evaluation, the above MT models cannot account for genotype-environment ($G \times E$) and trait-genotype-environment ($T \times G \times E$) interactions. The Bayesian whole genome prediction [35] that makes use of an efficient MCMC procedure for parameter estimation with an exact Gibbs sampling for the posterior distribution was pioneered to fill this knowledge gap [29, 88]. For instance, several traits for maize and wheat have been predicted in diverse environments using Bayesian multiple-trait multiple-environment (BMTME) models [61]. This set of approaches incorporates $G \times T$ and $T \times G \times E$ interactions to improve predictability. A variant of BMTME that is equivalent to principal component analysis and termed SVD-BMTME, handles correlated traits by making use of singular value decomposition (SVD) on the trait matrix. This is performed to remove the information related to other covariates (i.e. de-correlation). Since the uncorrelated and decomposed vectors are subsequently used as traits, the model implementation becomes straightforward as computational tools designed for univariate models can be employed. The final MT predictions are derived by transforming back the decomposed vectors to their original scale (i.e. before decomposition) [61].

Table 1. Classification of multi-traits GP models. Shown is an non-exhaustive list of recently developed MT-GP models with their corresponding references.

Categories	Types	Models	References	Year
Parametric	Bayesian models	Bayesian MTME	[40]	2022
		MT-BayesA	[37]	2012
		MT-BayesB	[37, 41]	2018, 2012
		MT-BayesCrr	[42]	2020
		MT-Bayesian rr-BLUP	[43]	2021
		MT-BayesSSVS	[44]	2011
		MT-BayesAS	[45]	2014
		Bayesian MORS	[46, 47]	2019, 2020
		Bayesian MTRS	[48]	2016
	Penalized multivariate regression models	MT-Bayesian LASSO	[41]	2018
		MT-LASSO	[49, 50]	2019, 2021
		MT-RR	[50]	2021
		MRCE	[51]	2010
		MOR	[39]	2016
		L ₂₁ -Joint	[52]	2021
		L ₁ -MTL	[53]	2009
		Cluster-based MTL	[54]	2011
		L ₂₁ -MTL	[53]	2009
		Trace-norm regularized MTL	[55]	2009
		MTL group-EN	[56]	2014
		MTL group-RR	[56]	2014
		SPRING	[57]	2017
	Mixed linear models	MT-BLUP	[58, 59]	2017, 2018
		MT-GBLUP	[42]	2020
		MT-rrBLUP	[44, 45]	2011, 2014
		MT-wGBLUP	[60]	2018
		Multivariate pedigree-BLUP and -GBLUP	[37]	2012
	Others multivariate regression models	SVD-BMTME	[61]	2019
		MT-PLS	[62]	2023
		KMLASSO	[63]	2013
		MegaLMM	[64]	2021
		MTERC	[65]	2023
		CNNs	[66, 67]	2018, 2024
		RNNs	[68]	2018
		MLNN	[40, 69]	2024, 2022
		MT-DL, Mtcro	[68, 70–72]	2018, 2022, 2019, 2025
		MPDN	[73]	2020
		MT-RF	[74]	2024
Non-parametric		MT-RKHS	[75–77]	2022, 2023, 2025
		QMTSVR	[76]	2023
Semi-parametric				

BLUP, best linear unbiased prediction; CNNs, convolutional neural networks; DL, deep learning; EN, elastic net; KM, kernelized multivariate; LASSO, least absolute shrinkage and selection operator; ME, multiple-environment; MegaLMM, mega-scale linear mixed models; MLNN, multi-layer neural network; MOR, multiple output regression; MORS, multi-output regression stacking; MPDN, multi-trait Poisson deep neural network; MRCE, multivariate regression with covariance estimation; MT, multiple-trait; MTERC, multi-target ensemble regression chains; MTL, multi-task learning; MTRS, multi-target regressor stacking; PLS, partial least square; QMTSVR, quasi multitask SVR; RF, random forest; RKHS, reproducing kernel Hilbert space; RNNs, recurrent neural networks; RR, ridge regression; SPRING, structured regularization with underlying sparsity; SSVS, stochastic search variable selection; SVD, singular value decomposition; SVR, support vector regression; wGBLUP, weighted genomic BLUP.

In a different perspective, models accounting for complex covariance structures in simultaneous analysis of MT in several environments have also been proposed using a two-stage approach. The first stage uses the same ST-GP model to predict each trait. The predicted traits are then used in the second step as predictors in a MT model for final predictions. These are built upon a Bayesian extension of multi-target regressor stacking initially proposed in [48] and include variants of the Bayesian multi-output regression stacking (BMORS) [46, 47]. This framework present flexibility in term of modeling depending on the data at hands, whereby the user can account for both linear and non-linear MT relationships in the second stage by selecting the desired model in the first stage. In this regards, random forest

regression has been employ in the first step [33] to account for non-linearity.

These MT models are implemented as their ST versions, whereby vectorizations of the matrix of traits is applied; in the Bayesian setting, these models can quickly become computationally expensive because of the MCMC steps required during parameter estimation. To address this issue, multivariate regression models that incorporate different correlation structures to better exploit the possible shared between-traits and between-genotypes relationships have been proposed [39, 51, 52]. These models are built upon (Equation 1) and do not explicitly model genetic random effects (i.e. $\mathbf{Y} = \mathbf{XB} + \mathbf{E}$). Instead, with suitable distributional assumption on the residuals, they use a

multivariate likelihood framework and impose different regularization on parameters of interest (i.e. regression coefficients and covariance matrix). Under the assumption that the covariance is the identity matrix, these models can be further simplified and design as simple regularized multiple output linear regression. These include the multivariate extension of least absolute shrinkage and selection operator (LASSO), ridge regression and elastic net [50]. To account for possible non-linear dependence between the predictors and responses, Kernelized multivariate LASSO [63] that solves $\hat{\mathbf{B}}(\lambda) = \underset{\mathbf{B}}{\operatorname{argmin}} \|\mathbf{Y} - \Phi(\mathbf{X})\mathbf{B}\|_2 + \sqrt{\lambda}\|\mathbf{B}\|_1$, with kernel function Φ , has also been proposed and applied.

Another challenge often encountered in data used for GP is related to the derivation of parameter estimates unaltered by variables collinearity. In a multivariate regression framework, this has been explored using the MT equivalent of partial least square regression [89] (MT-PLS) and applied to enhance prediction accuracy of elite wheat yield, groundnut and rice [90], as well as potato cultivars [62] in new environments. MT-PLS has the ability to simultaneously accommodate different factors in small- n large- p setting and with multiple traits. Without the random effect part, MT-PLS can be derived from (Equation 1), the only difference with the previously mentioned multivariate regression being that \mathbf{Y} is regressed on latent variable or scores \mathbf{T} instead of the original predictors \mathbf{X} . The estimate is derived in an iterative procedure that maximizes the covariance between the original traits and the latent variables, and the final coefficients rotated or converted back to the original space of \mathbf{X} .

Non-parametric models

Unlike their parametric counterparts, non-parametric approaches do not rely on strict distributional assumptions, rendering them more robust to outliers or when the normality assumption is violated, as it is often the case with experimental data sets. Further, in non-parametric approaches, there is no predefined link function between phenotypes and genetic markers. Instead, this relationship is learned from the data. Driven by the need to accurately capture complex genetic architectures, the non-parametric class of models has recently gained popularity. These models relax the often assumed linear relationships between traits and genetic markers, providing more flexibility and the ability to reduce bias [79, 91].

Recently developed approaches include the multi-target ensemble regression chains that selects assistant trait automatically and predict the GEBVs of the target trait using genotypic information only [65]. MT deep learning (MT-DL) approaches, including: (i) convolutional neural networks (CNNs) [67], that can capture spatial dependencies between markers in modeling effects on phenotypic traits, provide hierarchical feature representations, and leverage large-scale genomic data; (ii) recurrent neural networks (RNNs), useful for modeling of time-series data; MtCro, an extension of deep learning to accommodate multi-task [72]; and (iv) multi-layer neural network (i.e. input, hidden, and output layers) [69] (MLNN) that bypasses the restrictive assumption of additive genetic effects of markers. More precisely, the input layer consists of SNPs, the neurons or mapping units in which the weighted sum of nodes from the SNPs is computed constitute the hidden layer and the output layer contains the model outcome. Nonetheless, for large number of SNPs, the MLNN can quickly become computational intractable requiring the usage of a subset of SNPs (e.g. significant SNPs

derived from MT- genome wide association studies). In particular, they have shown great promise thanks to their ability to model highly non-linear relationships and interactions among genetic markers [66].

Extension to this work aimed to account for MT interactions in multiple environments [68] and different type of responses (e.g. binary and ordinal) [92] have also been studied. Using three real data sets comprising elite maize and wheat lines, it was demonstrated that MT-DL model is less computational expensive and is a competitive alternative to the BMTME, with the best predictions achieved when the G x E interaction was ignored. In contrast, the BMTME model outperformed the MT-DL in terms of prediction accuracy when G x E interaction was considered.

The non-parametric nature of deep learning and related approaches allows the model to learn without predefined assumptions about the relationships between traits, making it ideal for handling complex trait architectures. Their flexibility can facilitate the incorporation of multi-environment and multi-trait data simultaneously, providing robust predictions across different conditions. Nonetheless, the application of MT-DL models in GP is still facing some challenges. They are more difficult to generalize, as they require larger sample size compared to the parametric counterparts [71], they are unable to estimate the between-traits covariance, and are often computationally expensive in terms of parameter tuning.

Semi-parametric models

Blending the parametric and non-parametric techniques thereby, allowing for flexibility while maintaining some level of interpretability, semi-parametric approaches have also emerged as a powerful tool for MT-GP. Previously used in ST-GP [93] and relying on kernel functions to build a covariance structure between traits, kernel-based regressions such as support vector (SVR) and reproducing kernel Hilbert space (RKHS) have also been adapted to accommodate MT-GP. These adaptations include: (i) the quasi multitask SVR (QMTSVR) with hyperparameter tuning derived from genetic algorithm, (ii) its weighted alternative, and (iii) the MT-RKHS, where the usual G matrix is replaced with a nonlinear kernel K from SVR [76]. RKHS method is particularly interesting as it leverages kernel functions to capture complex, non-linear relationships between genetic markers and traits while still maintaining some parametric structure [77]. This method has been shown to provide improved prediction accuracy on eight out of 13 lodgepole pine traits, and lower bias for all considered traits than a purely parametric model (i.e. MT-GBLUP) [75].

Models used in the comparative analysis and performance assessment

In what follows, the performance of the baseline ST genomic best linear unbiased prediction (ST-GBLUP) [80], the genomic variant of the rrBLUP is contrasted with that of five representative of previously discussed MT models including: MT Bayesian multi-output regressor stacking (MT-BMORS) [29], MT Multi output regression (MT-MOR) [39], MT Singular value decomposition (MT-SVD) [61], MT partial least square regression (MT-PLS) [62], and MT deep learning (MT-DL) [68]. It is worth noting that the selection of these representative MT models was based on the availability of implementation tools in the R programming language and to avoid as much as possible inclusion of models that have been assessed together in a previous comparative analysis with ST-GP. In addition to using a different data set, we note that our analysis

differs from those presented in other studies [43, 49, 94, 95] based on the number of considered MT models across different types and the number of traits investigated.

Performances of the considered five MT-GP approaches along with that of the baseline ST-GBLUP were quantified by the Pearson correlation coefficient between predicted and observed trait values in the testing set. Final prediction accuracy values were then computed as the average performances over 100 (i.e. 20 repetitions of 5-fold CV) and 200 (i.e. 20 repetitions of 10-fold CV) for three CV scenarios: (i) in CV-A, models were trained on Indica and Japonica to predict respectively Indica and Japonica. (ii) CV-B corresponds to the CV scenario where the contending models were trained on data from Indica and were used to assess the performance on Japonica and vice versa. (iii) Finally, CV-C is concerned with a random splitting with varying proportion of combination of Indica/Japonica samples to predict the remaining mixed samples of Indica and japonica accessions. Note, however, that, for MT-DL a controlled random CV was used, whereby instead of using a purely random split, we considered as testing set one fold from the previous 5- or 10-fold CV and the remaining folds as training set, and implemented following an adaption of the R script provided in [68].

Phenotypic and genotypic data

To contrast the predictability of selected approaches, we used freely available phenotypic and genotypic data from previous studies [96, 97]. The number of traits that can be accommodated by a MT-GP depends on the type of model under consideration. While some [64] can effectively model thousands of traits, the computational complexity of others becomes a challenge (e.g. estimation of large covariance matrix) as the number of traits increases [68]. These challenges call for adequate planning when choosing the number of traits for a given breeding objective, to ensure optimal performance. Therefore, a fair comparative analysis dictates using the number of traits (i.e., here ten) that can be handled by all models.

Genomic data

A diversity panel of 533 *O. sativa* accessions, including landraces and elite varieties, was obtained from various sources [98–100]. We obtained the bed format, together with the corresponding bim and fam files associated with the rice accessions in RiceVarMap2 database [101]. We then built a preprocessing pipeline in PLINK [102, 103] to select genotypes and variants of interest. Loci were filtered by utilizing a moving window of 100 gbp with a step of 10 gp while considering a quality threshold of 0.95. Selecting accessions for which traits information were available and removing SNPs with minor allele frequency (MAF) < 5%, yielded a final dataset of 506 accessions and 973 275 markers (see Supplementary File 1).

Yield-related traits

Field trials for yield-related traits were conducted in three environments: Huazhong Agricultural University, Wuhan, China (i.e. 2011 and 2012), and Lingshui County, Hainan Island (i.e. 2011). Rice seedlings were transplanted to fields in a randomized complete block design with two replications, and yield was measured from five plants per plot. This trait dataset contains rice genotypes from different populations (and subpopulations), including: Aromatic, Aus, Indica, and Japonica. Detailed information on the population structure at the genotype level is provided in (Supplementary Table S1). The selected yield-related traits [104] include: yield, plant (PH) height, grain weight (GW), heading date (HD), and panicle seed setting rate

(PSSR). Accession-specific values for each trait are provided in (Supplementary Table S2), whereas the min-max scaled traits are available in (Supplementary Table S3).

Metabolic traits

The metabolomic data set [96] includes 840 metabolites and replicates measured across 506 accessions as shown in (Supplementary Table S4). Metabolic traits in the assembled dataset come from a variety of classes including flavonoids, terpenes, fatty acids, amino acids, nucleic acid derivatives, polyphenols, and phenylamines. For the metabolite selection, marker-based heritability [105] were computed for each trait. Metabolites were then selected to represent high ($h^2 > 0.7$) and low ($h^2 < 0.5$) heritability (see Supplementary Table S5) and to ensure that they are representative of subpopulations present in the selected focal traits. To this end, a tricin derivative (i.e. spectra peak labeled mr1246) and C-pentosyl-apigenin O-p-coumaroylhexoside (i.e. mr1234), a flavonoid compound weighing around 711 Daltons with molecular formula $C_{35}H_{34}O_{16}$ [106] were retained as high-heritability traits (i.e. $h^2 = 0.89$). Subsequently, a tricin derivative with a spectral peak labeled as mr1198 [106] ($h^2 = 0.45$), a polyphenol named N-Feruloyltyramine ($h^2 = 0.31$), a compound weighing around 314 daltons with molecular formula $C_{18}H_{19}N_4O_4$ (i.e. mr1268) [107] and LPC(1-acyl 18:2) with a spectral peak labeled mr1418 ($h^2 = 0.24$) a fatty acid (i.e. $C_{26}H_{50}N_7P$) [107] as low-heritability traits.

Trait summary, heritability, and correlation

BLUP for metabolic traits were computed using a LME with genotypes and replicates as random effects. The above SNP data were used as input in TASSEL [108], to derive the GRM. The obtained GRM along with traits values were then used to derive variance components in a restricted maximum likelihood (REML) framework [109] using the heritability package [110] in R statistical software (RCoreTeam2021) [111].

Results

The comparative analysis of prediction accuracy across representative five multi-trait (i.e. MT-BMORS, MT-MOR, MT-SVD, MT-PLS, MT-DL) and a baseline single-trait (ST-GBLUP) models highlights distinct trends across different traits and CV scenario.

Our findings in (Fig. 2) clearly show the influence of CV schemes on prediction accuracy. However, no consistent trend could be observed across traits and models for CV-A, where the models are trained and tested within the same subspecies (Indica or Japonica) (Fig. 2a and b), as well as CV-B, where models trained on one subspecies (Indica or Japonica) (Fig. 2d and e) are validated on the other. However, for most traits the highest prediction accuracies are observed in CV-A. This can be attributed to the genetic homogeneity within each subspecies, which simplifies the prediction task. Conversely, in CV-B, accuracies drop substantially likely due to the genetic divergence between subspecies (i.e. training and testing populations), reflecting the expected challenges of transfer learning across genetically divergent groups in GP. CV-C, which involves random splits across combined Indica and Japonica (Fig. 2c) sub-populations, seems to reconcile this discrepancy, highlighting the benefit of using heterogeneous training set for predictions on mixed testing populations. These observations are in line with findings in a previous study [112], emphasizing the importance of using heterogeneous training data for robust GPs.

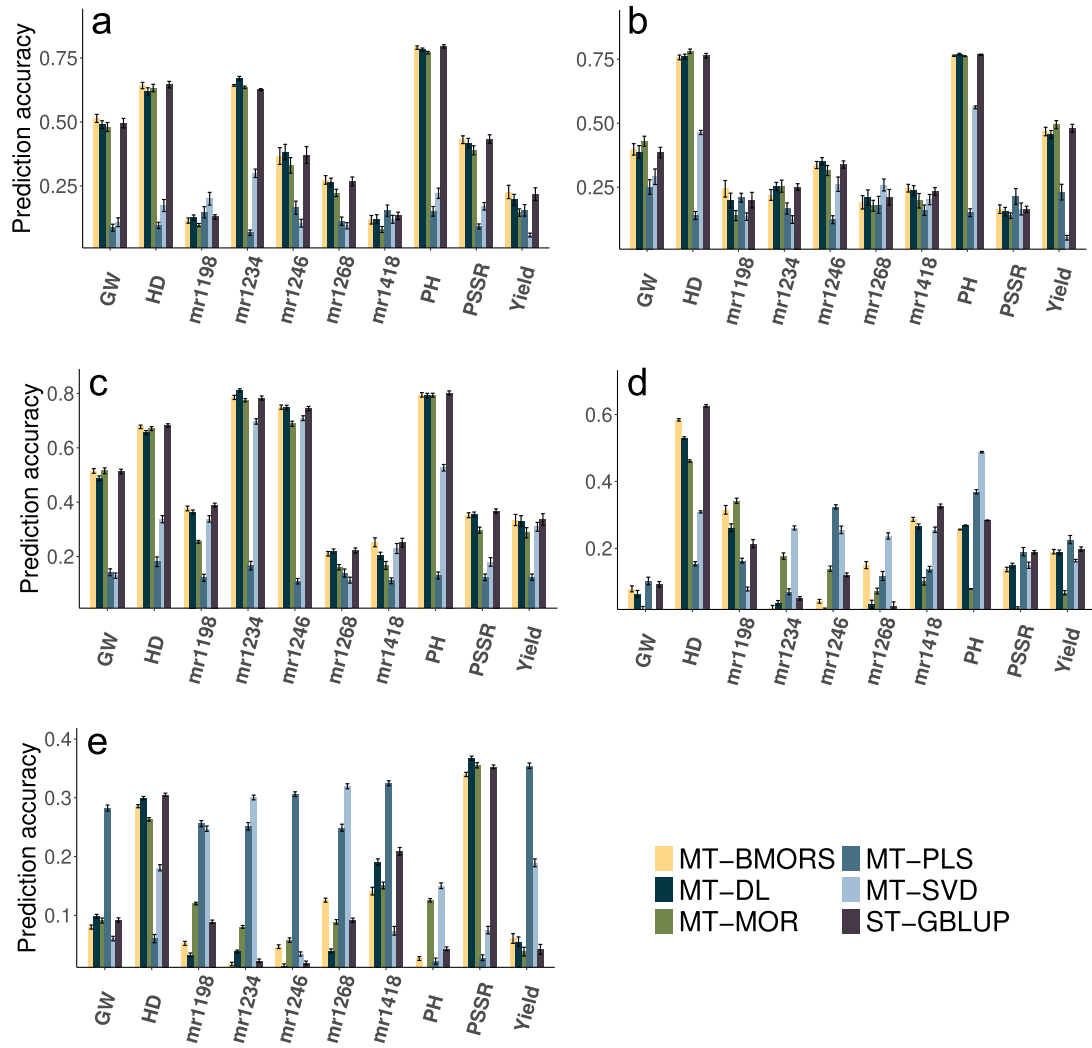


Figure 2. Comparison of predictabilities for MT and a baseline GP methods with a rice data set. We used five MT-GP models, namely: MT-BMORS, MT-MOR, MT-SVD, MT-PLS, and MT-DL, and ST-GBLUP to predict the levels of five metabolites (i.e. mr1198, mr1234, mr1246, mr1268, and mr1418; see [Metabolic traits](#) section for full description) as well as five yield-related traits (i.e. yield, GW, HD, PSSR, and PH). The predictability is computed as the average Pearson correlation coefficient between observed and predicted values for the ten traits in the validation set, based on 20 repetitions of 5- and 10-fold CV for respectively CV-A (a and b), CV-B (d and e), and CV-C (c). The average accuracy obtained from repeated CVs are reported as the height of the bars along with the standard errors. Panels a and b correspond to the CV schemes in which models were trained on Indica and Japonica to predict traits in Indica and Japonica accessions, respectively. In contrast, panels d and e correspond respectively to the CV scenario where the models were trained on data from Indica (Japonica) and used to predict the performance on Japonica (Indica). Finally, panel c is concerned with the random split with varying proportion of combined Indica/Japonica samples to predict the remaining mixed samples of Indica and japonica.

Except for MT-PLS, that exhibits the lowest accuracy, MT models slightly outperform the ST-GBLUP model in most scenarios, particularly in CV-A and CV-C, where trait correlations can be effectively leveraged. Traits such as grain weight (GW) and plant height (PH) show the most significant gains with MT models, reflecting their ability to exploit shared genetic architectures. However, in CV-B, where the training and validation sets are derived from divergent genetic backgrounds, the advantage of MT models diminishes slightly, likely because of the reduced utility of trait correlations under such conditions. Similar patterns have been reported in literature [95] highlighting the utility of MT approaches in scenarios with high genetic and phenotypic correlations. Notably, MT-DL and ST-GBLUP consistently exhibit high predictive accuracy, indicating their robustness.

Between the MT models, MT-DL and MT-BMORS consistently achieve the highest prediction accuracies across all CV designs and traits. The strength of MT-DL resides in its ability to handle

complex and non-linear relationships, while MT-BMORS benefits from its ability to account for trait correlations. Despite being less competitive than MT-DL and MT-BMORS for most traits, the performance of MT-MOR is not negligible in comparison to that of MT-PLS and MT-SVD. The poor performance and strong variability exhibited by MT-PLS and MT-SVD could suggest their sensitivity to specific trait architecture or a possible loss of information during the decompositions involved in the respective algorithm. Similar findings have previously been reported [66, 68] showcasing the potential of deep learning and Bayesian approaches for GP, particularly when simultaneously modeling multiple traits.

Overall, our findings underscore the value of multi-trait models in leveraging trait correlations, particularly for complex or highly polygenic traits. While single-trait models like ST-GBLUP remain effective and competitive in certain cases, MT approaches provide a significant advantage, especially when dealing with correlated traits or leveraging shared genetic architectures across

traits. The consistent superiority of MT-DL and MT-MOR across traits and CV schemes suggest their potential utility in breeding applications. However, the choice of model should also consider computational challenges and trait-specific requirements, since MT-DL may demand higher computational resources compared to models like MT-BMORS or ST-GBLUP. It could be interesting for future research to investigate how MT-GP approaches could further be optimized for the benefit of practical breeding, and to also explore the integration of different data types (e.g. environmental, high-throughput phenotyping, and transcriptomic) to enhance prediction accuracy and applicability.

Discussion

Our overview of newly developed MT-GP models and the comparative performances of parametric, semi-parametric, and non-parametric models revealed that no single approach is universally superior; rather, the choice of model often depends on the specific traits and their genetic architecture. Likewise, a similar conclusion could be reached with respect to the selected approaches used in the present analysis. Parametric models like MT-GBLUP and Bayesian methods are advantageous due to their simplicity and interpretability, particularly when the genetic architecture is well understood. Semi-parametric approaches such as MT-RKHS on the other hand, provide a balance between flexibility and interpretability, making them suitable for traits with moderate complexity. Non-parametric methods, including all variants of DL, excel in capturing intricate non-linear relationships but may require larger datasets and greater computational resources.

The future of MT-GP lies certainly in the integration of these diverse approaches, potentially through ensemble methods that combine the strengths of parametric, semi-parametric, and non-parametric models. Advances in computational power and algorithm development will further enhance the applicability and accuracy of these models. Moreover, as the availability of multi-omics data increases, integrating MT-GP with other layers of biological information, such as transcriptomics and metabolomics, promises to revolutionize our ability to predict complex traits with high precision [113, 114]. As highlighted in [68], the choice of appropriate hyperparameters for the implementation of non-parametric models such as MT-DL remains a challenge. This calls for the consideration of different network architectures and sets of hyperparameters that all together could enhance the reliability of the approach and full incorporation into practical breeding.

High-throughput phenotyping technologies have significantly advanced the capacity to capture large-scale, multi-dimensional phenotypic data, which is critical for MT-GP. By enabling the measurement of multiple traits across diverse environmental conditions with greater precision, high-throughput phenotyping could enhance the accuracy of trait heritability estimates and facilitate the identification of $G \times E$ interactions. This rich data collection allows MT-GP models to better account for the complex relationships between multiple traits and environmental factors, and could ultimately improve predictive power and selection efficiency in plant breeding programs [115]. A notable example is a study by [71] in which three traits were evaluated in 43 environments across several water regimes, showing that, despite the added complexity from the data, performance of multi-trait deep learning model matches that of GBLUP. In addition, the ability to gather data across different growth stages and environmental conditions also supports more robust modeling of the temporal dynamics of phenotypic expression.

Nonetheless, the improved accuracy exhibited by multi-trait models when using multi-dimensional phenotypic data from high-throughput phenotyping platforms should be taken with caution. While offering model flexibility the availability of these large scale datasets constitutes a challenge on its own, thereby one needs through some pre-processing, genetic correlation or heritability to select a representative subset of traits to be analyzed. Even for high genetic correlation, it has been shown that multivariate models may not be the best approach when predicting lines with small genetic relatedness [95]. Systematic assessment of multi-trait models for different combination of heritability and genetic correlation should be kept in mind during traits selection and model evaluation. For instance, simultaneous modeling for: (i) traits with low heritability and high genetic correlation with others and (ii) traits with high heritability and high negative genetic correlation with others [116].

Another critical limitation faced by traditional MT-GP models is the number of traits they can accommodate while preserving their natural multidimensional structure and their possible shared relationships. These multi-trait models have difficulties to cope with increasing trait numbers due to computational demands (e.g. MT-DL) and poor estimate of high-dimensional covariance matrix (e.g. regularized multivariate regression in the maximum likelihood framework). Additionally, many MT-GP models rely on complete phenotypic datasets, reducing their effectivity in partially observed traits setting. Mega-scale linear mixed models (MegaLMM) [64] that has been shown to handle thousands of traits simultaneously overcomes this constraint through the usage of sparsity-inducing priors while maintaining computational efficiency and improved prediction accuracy.

One of the key challenges in MT-GP, as for their ST counterpart, is the transferability of predictive models across environments, thereby models trained in one environment often struggle to predict trait performance accurately in a different environment due to differences in environmental conditions, $G \times E$ interactions, and trait plasticity. While multi-trait models can leverage shared genetic architecture among correlated traits, they still face difficulties when the environmental context changes, as the model might capture trait associations that are specific to the training environment. This limitation, akin to what is observed in ST-GP [117], presents a significant challenge for breeding programs targeting broad environmental adaptability. Multi-trait multi-environment [29] and transfer learning [118, 119] GP models have been proposed to address this issue, nonetheless their applications in GP remains underdeveloped.

The fast developing field of artificial intelligence (AI) could open incredible possibilities for multi-trait GP models. Some possible directions include: (i) leveraging explainable AI techniques such as Shapley Additive exPlanations [120] to identify key genomic markers driving multiple traits or trait interactions to make prediction from MT-GP models more biologically meaningful. (ii) Attention mechanisms [121] can be adapted to identify which genomic features contribute to the predictability of multiple traits.

While time series methods have shown potential for phenomic prediction by capturing temporal patterns in trait development [70], their application in MT-GP remains largely unexplored. Incorporating time-series data could provide a more comprehensive understanding of how multiple traits co-evolve over time and under different environmental conditions, thereby improving the predictive accuracy and offering new insights into complex trait

inter-dependencies. This approach could be particularly valuable for many agronomically relevant traits, including growth and yield potential, that are dynamic throughout the plant life cycle.

The integration of genome wide association studies (GWAS) into GP models provides another means to improve the biological interpretability of GP models [77, 122]. GWAS can identify genetic variants associated with multiple traits, allowing for the inclusion of trait-specific marker effects in prediction models. By leveraging GWAS results, GP models can be refined to account for pleiotropic effects (i.e. a single locus influences multiple traits), thereby improving the accuracy when predicting correlated traits. However, integrating GWAS data into MT-GP prediction models remains challenging due to the complexity of polygenic traits and the large number of genetic markers involved.

Despite successful applications across multiple model organisms and practical breeding programs, MT-GP models are prone to overfitting and a decrease of prediction accuracy has been observed with respect to their ST alternative in several comparative analysis. This makes model selection a challenging task especially when secondary traits measured on genotypes from the testing population are used to predict focal traits [123]. Additionally, results from multi-locus shrinkage modeling revealed that the majority of agronomic traits in crops exhibit considerable polygenicity due to small effects of multiple QTL [124–126]. Choosing the appropriate MT model to accurately predict such traits while simultaneously accounting for the small- n large- p problem remains a challenge in MT-GP. This calls for appropriate and unbiased methodology such as the commonly used CV [127, 128] for performance (i.e. predictability or prediction accuracy) evaluation. However, CV strategies in MT-GP must be adapted to emulate realistic plant-breeding mechanisms, account for the targeted prediction scenarios and insure independence between the training set and the testing one. These challenges are addressed by using CV1 and CV2 [117] corresponding respectively to the prediction of untested lines (newly developed genotypes) in tested environments and sparse testing where lines tested in some environments are to be predicted in other tested environments.

Finally, addressing the increasing global food demand in the context of a growing population is one of the most pressing challenges in plant breeding, thereby crop production must increase significantly to ensure food security. This challenge is compounded by the effects of climate change, which threaten to reduce agricultural productivity. One way to accelerate crop improvement that enables the selection of plants with optimal combinations of traits, such as higher yield and disease resistance is the usage of MT-GP models. However, the complexity of breeding for multiple traits simultaneously, especially under changing environmental conditions, presents significant obstacles. Advances in computational methods, along with the integration of multi-environment GP [17], will be crucial to addressing these challenges and ensuring a resilient global food supply.

Key Points

- We provided a classification of computational approaches for multi-trait genomic prediction (GP).
- We pointed the differences of the underlying principles, assumptions, and potential limitations of the three classes of computational approaches for multi-trait GP.

- We compared the performance of five representative approaches for GP of 10 traits related to yield and metabolism in a rice diversity panel.
- We observed a consistent superior performance of multi-trait deep learning as well as multi-output regression across the studied traits and cross-validation schemes.
- Nevertheless, the choice of approach to be used in practice depends on several factors, including the traits considered, the population structure, and genetic architecture of the traits.

Author contributions

Alain J. Mbebi (Method implementation, Method testing, Data analysis, Figure preparation, Manuscript writing), Facundo Mercado (Method implementation, Data analysis), David Hobby (Method implementation, Data analysis), Hao Tong (Method implementation, Data analysis, Method testing), Zoran Nikoloski (Conceptualization, Data analysis, Funding acquisition, Manuscript writing).

Supplementary material

[Supplementary material](#) is available at *Briefings in Bioinformatics* online.

Conflict of interest

All authors declare that they have no conflict of interest.

Funding

This project was funded by the Horizon Europe research and innovation program, project BOLERO (Breeding for coffee and cocoa root resilience in low-input farming systems based on improved rootstock, HORIZON-CL6-2021-BIODIV-01-13), under grant agreement ID: 101060393.

Data availability

We implemented all statistical models using R programming language and the codes can be freely accessed from https://github.com/alainmbebi/MT_Review. Phenotypic and genotypic data used in this study were from [96] and are available for query from reference within.

References

1. Van Dijk, Morley T, Rau ML. et al. A meta-analysis of projected global food demand and population at risk of hunger for the period 2010–2050. *Nat Food* 2021;**2**:494–501. <https://doi.org/10.1038/s43016-021-00322-9>
2. Tester M, Langridge P. Breeding technologies to increase crop production in a changing world. *Science* 2010;**327**:818–22. <https://doi.org/10.1126/science.1183700>
3. McCouch S, Baute GJ, Bradeen J. et al. Feeding the future. *Nature* 2013;**499**:23–4. <https://doi.org/10.1038/499023a>

4. Dwivedi SL, Heslop-Harrison P, Amas J. et al. Epistasis and pleiotropy-induced variation for plant breeding. *Plant Biotechnol J* 2024;**22**:2788–807. <https://doi.org/10.1111/pbi.14405>
5. Mackay TFC, Anholt RRH. Pleiotropy, epistasis and the genetic architecture of quantitative traits. *Nat Rev Genet* 2024;**25**: 639–57. <https://doi.org/10.1038/s41576-024-00711-3>
6. Laitinen RAE, Nikoloski Z. Strategies to identify and dissect trade-offs in plants. *Mol Ecol* 2024;**33**:e16780. <https://doi.org/10.1111/mec.16780>
7. Tuberosa R. Phenotyping for drought tolerance of crops in the genomics era. *Front Physiol* 2012;**3**:347. <https://doi.org/10.3389/fphys.2012.00347>
8. Ahmar S, Gill RA, Jung K-H. et al. Conventional and molecular techniques from simple breeding to speed breeding in crop plants: recent advances and future outlook. *Int J Mol Sci* 2020;**21**:2590. <https://doi.org/10.3390/ijms21072590>
9. Bernardo R. Molecular markers and selection for complex traits in plants: learning from the last 20 years. *Crop Sci* 2008;**48**: 1649–64. <https://doi.org/10.2135/cropsci2008.03.0131>
10. Jannink J-L, Lorenz AJ, Iwata H. Genomic selection in plant breeding: from theory to practice. *Brief Funct Genomics* 2010;**9**: 166–77. <https://doi.org/10.1093/bfpg/elq001>
11. Riedelsheimer C, Lisec J, Czedik-Eysenberg A. et al. Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize. *Proc Natl Acad Sci* 2012;**109**:8872–7.
12. Ben-Ari G, Lavi U. Marker-assisted selection in plant breeding. In: Altman A, Hasegawa PM (eds.), *Plant Biotechnology and Agriculture*, 163–84. San Diego: Academic Press, 2012. <https://doi.org/10.1016/B978-0-12-381466-1.00011-0>.
13. Heffner EL, Sorrells ME, Jannink J-L. Genomic selection for crop improvement. *Crop Sci* 2009;**49**:1–12. <https://doi.org/10.2135/cropsci2008.08.0512>
14. Meuwissen THE, Hayes BJ, Goddard ME. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 2001;**157**:1819–29. <https://doi.org/10.1093/genetics/157.4.1819>
15. Fernandez O, Urrutia M, Bernillon S. et al. Fortune telling: metabolic markers of plant performance. *Metabolomics* 2016;**12**: 1–14.
16. Schrag TA, Westhues M, Schipprack W. et al. Beyond genomic prediction: combining different types of omics data can improve prediction of hybrid performance in maize. *Genetics* 2018;**208**:1373–85. <https://doi.org/10.1534/genetics.117.300374>
17. Crossa J, Pérez-Rodríguez P, Cuevas J. et al. Genomic selection in plant breeding: methods, models, and perspectives. *Trends Plant Sci* 2017;**22**:961–75. <https://doi.org/10.1016/j.tplants.2017.08.011>
18. Tong H, Nikoloski Z. Machine learning approaches for crop improvement: leveraging phenotypic and genotypic big data. *J Plant Physiol* 2021;**257**:153354. <https://doi.org/10.1016/j.jplph.2020.153354>
19. Heffner EL, Jannink J-L, Iwata H. et al. Genomic selection accuracy for grain quality traits in biparental wheat populations. *Crop Sci* 2011;**51**:2597–606. <https://doi.org/10.2135/cropsci2011.05.0253>
20. Marulanda JJ, Mi X, Melchinger AE. et al. Optimum breeding strategies using genomic selection for hybrid breeding in wheat, maize, rye, barley, rice and triticale. *Theor Appl Genet* 2016;**129**:1901–13. <https://doi.org/10.1007/s00122-016-2748-5>
21. Ishimori M, Hattori T, Yamazaki K. et al. Impacts of dominance effects on genomic prediction of sorghum hybrid performance. *Breed Sci* 2020;**70**:605–16. <https://doi.org/10.1270/jsbbs.20042>
22. Budhlakoti N, Kushwaha AK, Anil Rai KK. et al. Genomic selection: a tool for accelerating the efficiency of molecular breeding for development of climate-resilient crops. *Front Genet* 2022;**13**:832153. <https://doi.org/10.3389/fgene.2022.832153>
23. Meena MR, Chinnaswamy Appunu R, Arun Kumar R. et al. Recent advances in sugarcane genomics, physiology, and phenomics for superior agronomic traits. *Front Genet* 2022;**13**:854936. <https://doi.org/10.3389/fgene.2022.854936>
24. Li D, Quan C, Song Z. et al. High-throughput plant phenotyping platform (ht3p) as a novel tool for estimating agronomic traits from the lab to the field. *Front Bioeng Biotechnol* 2021;**8**:623705. <https://doi.org/10.3389/fbioe.2020.623705>
25. Song P, Wang J, Guo X. et al. High-throughput phenotyping: Breaking through the bottleneck in future crop breeding. *Crop J* 2021;**9**:633–45. <https://doi.org/10.1016/j.cj.2021.03.015>
26. Piepho H-P. Ridge regression and extensions for genomewide selection in maize. *Crop Sci* 2009;**49**:1165–76. <https://doi.org/10.2135/cropsci2008.10.0595>
27. Habier D, Fernando RL, Kizilkaya K. et al. Extension of the bayesian alphabet for genomic selection. *BMC Bioinform* 2011;**12**:1–12. <https://doi.org/10.1186/1471-2105-12-186>
28. Meher PK, Rustgi S, Kumar A. Performance of Bayesian and blup alphabets for genomic prediction: analysis, comparison and results. *Heredity* 2022;**128**:519–30. <https://doi.org/10.1038/s41437-022-00539-9>
29. Montesinos-López OA, Montesinos-López A, Crossa J. et al. A Bayesian genomic multi-output regressor stacking model for predicting multi-trait multi-environment plant breeding data. *G3: Genes Genomes Genet* 2019;**9**:3381–93. <https://doi.org/10.1534/g3.119.400336>
30. Habyarimana E, Lopez-Cruz M, Baloch FS. Genomic selection for optimum index with dry biomass yield, dry mass fraction of fresh material, and plant height in biomass sorghum. *Genes* 2020;**11**:61. <https://doi.org/10.3390/genes11010061>
31. Bouteillé M, Rolland G, Balsera C. et al. Disentangling the intertwined genetic bases of root and shoot growth in arabidopsis. *PloS One* 2012;**7**:e32319. <https://doi.org/10.1371/journal.pone.0032319>
32. He Z, Webster S, He SY. Growth–defense trade-offs in plants. *Curr Biol* 2022;**32**:R634–9. <https://doi.org/10.1016/j.cub.2022.04.070>
33. Fradgley N, Gardner KA, Bentley AR. et al. Multi-trait ensemble genomic prediction and simulations of recurrent selection highlight importance of complex trait genetic architecture for long-term genetic gains in wheat. *in silico Plants* 2023;**5**. <https://doi.org/10.1093/insilicoplants/diad002>
34. Henderson CR, Quaas RL. Multiple trait evaluation using relatives' records. *J Anim Sci* 1976;**43**:1188–97. <https://doi.org/10.2527/jas1976.4361188x>
35. Montesinos-López OA, Montesinos-López A, Crossa J. et al. A genomic Bayesian multi-trait and multi-environment model. *G3 Genes Genomes Genet* 2016;**6**:2725–44. <https://doi.org/10.1534/g3.116.032359>
36. Zhong S, Dekkers JCM, Fernando RL. et al. Factors affecting accuracy from genomic selection in populations derived from multiple inbred lines: a barley case study. *Genetics* 2009;**182**: 355–64. <https://doi.org/10.1534/genetics.108.098277>
37. Jia Y, Jannink J-L. Multiple-trait genomic selection methods increase genetic value prediction accuracy. *Genetics* 2012;**192**: 1513–22. <https://doi.org/10.1534/genetics.112.144246>
38. Neyhart JL, Lorenz AJ, Smith KP. Multi-trait improvement by predicting genetic correlations in breeding crosses. *G3: Genes Genomes Genet* 2019;**9**:3153–65. <https://doi.org/10.1534/g3.119.400406>

39. He D, Kuhn D, Parida L. Novel applications of multitask learning and multiple output regression to multiple genetic trait prediction. *Bioinformatics* 2016;**32**:i37–43. <https://doi.org/10.1093/bioinformatics/btw249>
40. Montesinos-López OA, Montesinos-López A, Mosqueda-Gonzalez BA. et al. Accounting for Correlation Between Traits in Genomic Prediction. *Methods in Molecular Biology*, 2022;**2467**:285–27. https://doi.org/10.1007/978-1-0716-2205-6_10.
41. Cheng H, Kizilkaya K, Zeng J. et al. Genomic prediction from multiple-trait Bayesian regression methods using mixture priors. *Genetics* 2018;**209**:89–103. <https://doi.org/10.1534/genetics.118.300650>
42. Song H, Zhang Q, Ding X. The superiority of multi-trait models with genotype-by-environment interactions in a limited number of environments for genomic prediction in pigs. *J Anim Sci Biotechnol* 2020;**11**:1–13. <https://doi.org/10.1186/s40104-020-00493-8>
43. Wang Z, Cheng H. Single-trait and multiple-trait genomic prediction from multi-class Bayesian alphabet models using biological information. *Front Genet* 2021;**12**:717457. <https://doi.org/10.3389/fgene.2021.717457>
44. Calus MPL, Veerkamp RF. Accuracy of multi-trait genomic selection using different methods. *Genet Select Evol* 2011;**43**. <https://doi.org/10.1186/1297-9686-43-26>
45. Janss L. Disentangling pleiotropy along the genome using sparse latent variable models. In: *10th World Congress on Genetics Applied to Livestock Production (WCGALP)*, 2014, Bayshore Drive, Vancouver, BC: The Westin Bayshore.
46. Montesinos-López OA, Montesinos-López A, Luna-Vázquez FJ. et al. An R package for bayesian analysis of multi-environment and multi-trait multi-environment data for genome-based prediction. *G3: Genes Genomes Genet* 2019;**9**:1355–69. <https://doi.org/10.1534/g3.119.400126>
47. Sirjan, Sapkota J, Boatwright L, Jordan K. et al. Multi-trait regressor stacking increased genomic prediction accuracy of sorghum grain composition. *Agronomy* 2020;**10**:1221. <https://doi.org/10.3390/agronomy10091221>
48. Spyromitros-Xioulis E, Tsoumakas G, Groves W. et al. Multi-target regression via input space expansion: treating targets as inputs. *Mach Learn* 2016;**104**:55–98. <https://doi.org/10.1007/s10994-016-5546-z>
49. Budhlakoti N, Mishra DC, Rai A. et al. A comparative study of single-trait and multi-trait genomic selection. *J Comput Biol* 2019;**26**:1100–12. <https://doi.org/10.1089/cmb.2019.0032>
50. Brault C, Doligez A, Cunff L. et al. Harnessing multivariate, penalized regression methods for genomic prediction and qtl detection of drought-related traits in grapevine. *G3* 2021;**11**. <https://doi.org/10.1093/g3journal/jkab248>
51. Rothman AJ, Levina E, Zhu J. Sparse multivariate regression with covariance estimation. *J Comput Graph Stat* 2010;**19**:947–62. <https://doi.org/10.1198/jcgs.2010.09188>
52. Mbebi AJ, Tong H, Nikoloski Z. L2, 1-norm regularized multivariate regression model with applications to genomic prediction. *Bioinformatics* 2021;**37**:2896–904. <https://doi.org/10.1093/bioinformatics/btab212>
53. Liu J, Ji S, Ye J. Multi-task feature learning via efficient l2,1-norm minimization. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*, 339–348, Arlington, Virginia, USA, 2009. AUAI Press.
54. Zhou J, Chen J, Ye J. Clustered multi-task learning via alternating structure optimization. *Adv Neural Inform Process Syst* 2011;**24**.
55. Abernethy J, Bach F, Evgeniou T. et al. A new approach to collaborative filtering: operator estimation with spectral regularization. *J Mach Learn Res* 2009;**10** 803–826.
56. Hastie T, Qian J. Glmnet vignette. Retrieved June 2014;**9**:1–30.
57. Chiquet J, Mary-Huard T, Robin S. Structured regularization for conditional gaussian graphical models. *Stat Comput* 2017;**27**: 789–804. <https://doi.org/10.1007/s11222-016-9654-1>
58. Lyra DH, de Freitas, Mendonça GG. et al. Multi-trait genomic prediction for nitrogen response indices in tropical maize hybrids. *Mol Breed* 2017;**37**. <https://doi.org/10.1007/s11032-017-0681-1>
59. Alves RS, Rocha JR D AS D C, Teodoro PE. et al. Multiple-trait blup: a suitable strategy for genetic selection of eucalyptus. *Tree Genet Genomes* 2018;**14**:1–8.
60. Karaman E, Lund MS, Anche MT. et al. Genomic prediction using multi-trait weighted gblup accounting for heterogeneous variances and covariances across the genome. *G3: Genes Genomes Genet* 2018;**8**:3549–58. <https://doi.org/10.1534/g3.118.200673>
61. Montesinos-López OA, Montesinos-López A, Crossa J. et al. A singular value decomposition bayesian multiple-trait and multiple-environment genomic model. *Heredity* 2019;**122**: 381–401. <https://doi.org/10.1038/s41437-018-0109-7>
62. Ortiz R, Reslow F, Montesinos-López A. et al. Partial least squares enhance multi-trait genomic prediction of potato cultivars in new environments. *Sci Rep* 2023;**13**:9947 <https://doi.org/10.1038/s41598-023-37169-y>
63. Jie X, Yin J. Kernel least absolute shrinkage and selection operator regression classifier for pattern classification. *IET Comput Vis* 2013;**7**:48–55. <https://doi.org/10.1049/iet-cvi.2011.0193>
64. Runcie DE, Jiayi Q, Cheng H. et al. Megalmm: mega-scale linear mixed models for genomic predictions with thousands of traits. *Genome Biol* 2021;**22**:1–25.
65. Liang M, Cao S, Deng T. et al. Mak: a machine learning framework improved genomic prediction via multi-target ensemble regressor chains and automatic selection of assistant traits. *Brief Bioinform* 2023;**24**:bbad043.
66. Montesinos-López A, Montesinos-López OA, Gianola D. et al. Multi-environment genomic prediction of plant traits using deep learners with dense architecture. *G3: Genes Genomes Genet* 2018;**8**:3813–28. <https://doi.org/10.1534/g3.118.200740>
67. Ma X, Wang H, Shengyang W. et al. Deepccr: large-scale genomics-based deep learning method for improving rice breeding. *Plant Biotechnol J* 2024;**22**:2691–3. <https://doi.org/10.1111/pbi.14384>
68. Montesinos-López OA, Montesinos-López A, Crossa J. et al. Multi-trait, multi-environment deep learning modeling for genomic-enabled prediction of plant traits. *G3: Genes Genomes Genet* 2018;**8**:3829–40. <https://doi.org/10.1534/g3.118.200728>
69. Mota LFM, Arikawa LM, Santos SWB. et al. Benchmarking machine learning and parametric methods for genomic prediction of feed efficiency-related traits in Nellore cattle. *Sci Rep* 2024;**14**:6404. <https://doi.org/10.1038/s41598-024-57234-4>
70. Sun Z, Li Q, Jin S. et al. Simultaneous prediction of wheat yield and grain protein content using multitask deep learning from time-series proximal sensing. *Plant Phenomics* 2022;**2022**:9757948. <https://doi.org/10.34133/2022/9757948>
71. Montesinos-López OA, Montesinos-López A, Tuberosa R. et al. Multi-trait, multi-environment genomic prediction of durum wheat with genomic best linear unbiased predictor and deep learning methods. *Front Plant Sci* 2019;**10**:1311. <https://doi.org/10.3389/fpls.2019.01311>

72. Chao D, Wang H, Wan F. et al. Mtcro: multi-task deep learning framework improves multi-trait genomic prediction of crops. *Plant Methods* 2025;**21**:12. <https://doi.org/10.1186/s13007-024-01321-0>
73. Montesinos-López OA, Montesinos-López JC, Singh P. et al. A multivariate poisson deep learning model for genomic prediction of count data. *G3: Genes Genomes Genet* 2020;**10**:4177–90. <https://doi.org/10.1534/g3.120.401631>
74. Mohsen Yoosefzadeh Najafabadi and Davoud Torkamaneh. Machine learning-enhanced multi-trait genomic prediction for optimizing cannabinoid profiles in cannabis. *Plant J* 2024;**121**:e17164. <https://doi.org/10.1111/tpj.17164>
75. Cappa EP, Chen C, Klutsch JG. et al. Multiple-trait analyses improved the accuracy of genomic prediction and the power of genome-wide association of productivity and climate change-adaptive traits in lodgepole pine. *BMC Genomics* 2022;**23**:536.
76. Alves AAC, Fernandes AFA, Lopes FB. et al. (quasi) multitask support vector regression with heuristic hyperparameter optimization for whole-genome prediction of complex traits: a case study with carcass traits in broilers. *G3: Genes Genomes Genet* 2023;**13**:jkad109.
77. Izquierdo P, Wright EM, Cichy K. GWAS-assisted and multitrait genomic prediction for improvement of seed yield and canning quality traits in a black bean breeding panel. *G3: Genes Genomes Genet* 2025;**15**:jkaf007.
78. Howard R, Carriquiry AL, Beavis WD. Parametric and nonparametric statistical methods for genomic selection of traits with additive and epistatic genetic architectures. *G3: Genes Genomes Genet* 2014;**4**:1027–46. <https://doi.org/10.1534/g3.114.010298>
79. López OAM. Abelardo Montesinos López, and José Crossa. Multivariate statistical machine learning methods for genomic prediction. Cham: Springer. 2022. <https://doi.org/10.1007/978-3-030-89010-0>.
80. VanRaden PM. Efficient methods to compute genomic predictions. *J Dairy Sci* 2008;**91**:4414–23. <https://doi.org/10.3168/jds.2007-0980>
81. Gupta AK, Varga T. Characterization of matrix variate normal distributions. *J Multivar Anal* 1992;**41**:80–8. [https://doi.org/10.1016/0047-259X\(92\)90058-N](https://doi.org/10.1016/0047-259X(92)90058-N)
82. Viroli C. On matrix-variate regression analysis. *J Multivar Anal* 2012;**111**:296–309. <https://doi.org/10.1016/j.jmva.2012.04.005>
83. Pollak EJ, Van der Werf, Quaas RL. Selection bias and multiple trait evaluation. *J Dairy Sci* 1984;**67**:1590–5. [https://doi.org/10.3168/jds.S0022-0302\(84\)81481-2](https://doi.org/10.3168/jds.S0022-0302(84)81481-2)
84. Geyer CJ. Practical Markov chain Monte Carlo. *Stat Science* 1992;**7**:473–83. <https://doi.org/10.1214/ss/1177011137>
85. Van Ravenzwaaij, Cassey P, Brown SD. A simple introduction to Markov chain Monte-carlo sampling. *Psychon Bull Rev* 2018;**25**:143–54. <https://doi.org/10.3758/s13423-016-1015-8>
86. Li X, Lund MS, Janss L. et al. The patterns of genomic variances and covariances across genome for milk production traits between chinese and nordic Holstein populations. *BMC Genet* 2017;**18**:1–12. <https://doi.org/10.1186/s12863-017-0491-9>
87. de Los G, Campos JM, Hickey RP-W. et al. Whole-genome regression and prediction methods applied to plant and animal breeding. *Genetics* 2013;**193**:327–45. <https://doi.org/10.1534/genetics.112.143313>
88. Avelar A, de Oliveira, Jr R. et al. Genomic prediction applied to multiple traits and environments in second season maize hybrids. *Heredity* 2020;**125**:60–72. <https://doi.org/10.1038/s41437-020-0321-0>
89. Wold H. Estimation of principal components and related models by iterative least squares. In Krishnajah PR. (ed.), *Multivariate Analysis* (pp. 391–420). NewYork: Academic Press 1966.
90. Montesinos-López OA, Montesinos-López A, Sandoval DAB. et al. Multi-trait genome prediction of new environments with partial least squares. *Front Genet* 2022;**13**:966775. <https://doi.org/10.3389/fgene.2022.966775>
91. Pérez-Enciso M, Zingaretti LM. A guide on deep learning for complex trait genomic prediction. *Genes* 2019;**10**:553. <https://doi.org/10.3390/genes10070553>
92. Montesinos-López OA, Martín-Vallejo J, Crossa J. et al. New deep learning genomic-based prediction model for multiple traits with binary, ordinal, and continuous phenotypes. *G3: Genes Genomes Genetics* 2019;**9**:1545–56. <https://doi.org/10.1534/g3.119.300585>
93. de Los Campos, Gianola D, Rosa GJM. et al. Semi-parametric genomic-enabled prediction of genetic values using reproducing kernel hilbert spaces methods. *Genet Res* 2010;**92**:295–308. <https://doi.org/10.1017/S0016672310000285>
94. Guo G, Zhao F, Wang Y. et al. Comparison of single-trait and multiple-trait genomic prediction models. *BMC Genet* 2014;**15**:1–7. <https://doi.org/10.1186/1471-2156-15-30>
95. Lozada DN, Carter AH. Accuracy of single and multi-trait genomic prediction models for grain yield in us pacific northwest winter wheat. *Crop Breed Genet Genom* 2019;**1**:e190012.
96. Chen W, Gao Y, Xie W. et al. Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism. *Nat Genet* 2014;**46**:714–21. <https://doi.org/10.1038/ng.3007>
97. Xie W, Wang G, Yuan M. et al. Breeding signatures of rice improvement revealed by a genomic variation map from a large germplasm collection. *Proc Natl Acad Sci* 2015;**112**:E5411–9. <https://doi.org/10.1073/pnas.1515919112>
98. Yu SB, Xu WJ, Vijayakumar CHM. et al. Molecular diversity and multilocus organization of the parental lines used in the international rice molecular breeding program. *Theor Appl Genet* 2003;**108**:131–40. <https://doi.org/10.1007/s00122-003-1400-3>
99. Yan WG, Li Y, Agrama HA. et al. Association mapping of stigma and spikelet characteristics in rice (*Oryza sativa* L.). *Mol Breed* 2009;**24**:277–92. <https://doi.org/10.1007/s11032-009-9290-y>
100. Zhang H, Zhang D, Wang M. et al. A core collection and mini core collection of *Oryza sativa* L. in China. *Theor Appl Genet* 2011;**122**:49–61. <https://doi.org/10.1007/s00122-010-1421-7>
101. Zhao H, Yao W, Ouyang Y. et al. Ricevarmap: a comprehensive database of rice genomic variations. *Nucleic Acids Res* 2015;**43**:D1018–22. <https://doi.org/10.1093/nar/gku894>
102. Purcell S, Neale B, Todd-Brown K. et al. Plink: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;**81**:559–75. <https://doi.org/10.1086/519795>
103. Chang CC, Chow CC, Tellier LCAM. et al. Second-generation plink: rising to the challenge of larger and richer datasets. *Gigascience* 2015;**4**:s13742–015.
104. Yang W, Guo Z, Huang C. et al. Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. *Nat Commun* 2014;**5**:5087. <https://doi.org/10.1038/ncomms6087>
105. Kruijer W, Boer MP, Malosetti M. et al. Marker-based estimation of heritability in immortal populations. *Genetics* 2015;**199**:379–98. <https://doi.org/10.1534/genetics.114.167916>
106. Chen W, Gong L, Guo Z. et al. A novel integrated method for large-scale detection, identification, and quantification of widely targeted metabolites: application in the study of rice metabolomics. *Mol Plant* 2013;**6**:1769–80. <https://doi.org/10.1093/mp/sst080>

107. Jang S, Hur J, Kim S-J. et al. Ectopic expression of *osyab1* causes extra stamens and carpels in rice. *Plant Mol Biol* 2004;**56**:133–43. <https://doi.org/10.1007/s11103-004-2648-y>
108. Bradbury PJ, Zhang Z, Kroon DE. et al. Tassel: software for association mapping of complex traits in diverse samples. *Bioinformatics* 2007;**23**:2633–5. <https://doi.org/10.1093/bioinformatics/btm308>
109. Gilmour AR, Thompson R, Cullis BR. Average information reml: an efficient algorithm for variance parameter estimation in linear mixed models. *Biometrics* 1995;**51**:1440–50. <https://doi.org/10.2307/2533274>
110. Kruijer W, Flood P, Kooke R. Heritability: marker-based estimation of heritability using individual plant or plot data R. Available online at: <http://CRAN.R-project.org/packages/heritability>, 2016.
111. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2021.
112. Rutkoski J, Poland J, Mondal S. et al. Canopy temperature and vegetation indices from high-throughput phenotyping improve accuracy of pedigree and genomic selection for grain yield in wheat. *G3: Genes Genomes Genet* 2016;**6**:2799–808. <https://doi.org/10.1534/g3.116.032888>
113. Yang Y, Saand MA, Huang L. et al. Applications of multi-omics technologies for crop improvement. *Front Plant Sci* 2021;**12**:563953. <https://doi.org/10.3389/fpls.2021.563953>
114. Zhang J, Xie Y, Zhang H. et al. Integrated multi-omics reveals significant roles of non-additively expressed small RNAs in heterosis for maize plant height. *Int J Mol Sci* 2023;**24**:9150. <https://doi.org/10.3390/ijms24119150>
115. Juliana P, Montesinos-López OA, Crossa J. et al. Integrating genomic-enabled prediction and high-throughput phenotyping in breeding for climate-resilient bread wheat. *Theor Appl Genet* 2019;**132**:177–94. <https://doi.org/10.1007/s00122-018-3206-3>
116. Atanda SA, Steffes J, Yang Lan M. et al. Multi-trait genomic prediction improves selection accuracy for enhancing seed mineral concentrations in pea. *Plant Genome* 2022;**15**:e20260. <https://doi.org/10.1002/tpg2.20260>
117. Burgueño J, Gustavo de los Campos, Kent Weigel, and José Crossa. Genomic prediction of breeding values when modeling genotype × environment interaction using pedigree and dense molecular markers. *Crop Sci* 2012;**52**:707–19. <https://doi.org/10.2135/cropsci2011.06.0299>
118. Muneeb M, Feng S, Henschel A. Transfer learning for genotype–phenotype prediction using deep learning models. *BMC Bioinform* 2022;**23**:511. <https://doi.org/10.1186/s12859-022-05036-8>
119. Li J, Zhang D, Yang F. et al. Trg2p: a transfer learning-based tool integrating multi-trait data for accurate prediction of crop yield. *Plant Commun* 2024;**5**:100975.
120. García MV, Aznarte JL. Shapley additive explanations for NO₂ forecasting. *Eco Inform* 2020;**56**:101039. <https://doi.org/10.1016/j.ecoinf.2019.101039>
121. Niu Z, Zhong G, Hui Y. A review on the attention mechanism of deep learning. *Neurocomputing* 2021;**452**:48–62. <https://doi.org/10.1016/j.neucom.2021.03.091>
122. Zhang Y, Zhang M, Ye J. et al. Integrating genome-wide association study into genomic selection for the prediction of agronomic traits in rice (*Oryza sativa* L.). *Mol Breed* 2023;**43**:81. <https://doi.org/10.1007/s11032-023-01423-y>
123. Wray NR, Yang J, Hayes BJ. et al. Pitfalls of predicting complex traits from snps. *Nat Rev Genet* 2013;**14**:507–15. <https://doi.org/10.1038/nrg3457>
124. Jansen RC. Complex plant traits: time for polygenic analysis. *Trends Plant Sci* 1996;**1**:89–94. [https://doi.org/10.1016/S1360-1385\(96\)80040-9](https://doi.org/10.1016/S1360-1385(96)80040-9)
125. Scott MF, Fradgley N, Bentley AR. et al. Limited haplotype diversity underlies polygenic trait architecture across 70 years of wheat breeding. *Genome Biol* 2021;**22**:137. <https://doi.org/10.1186/s13059-021-02354-7>
126. Mahmoud M, Tost M, Ha N-T. et al. Ghat: an R package for identifying adaptive polygenic traits. *G3 Genome Genet* 2023;**13**:jkac319.
127. Stone M. Cross-validatory choice and assessment of statistical predictions. *J R Stat Soc B Methodol* 1974;**36**:111–33. <https://doi.org/10.1111/j.2517-6161.1974.tb00994.x>
128. Efron B, Gong G. A leisurely look at the bootstrap, the jack-knife, and cross-validation. *Am Stat* 1983;**37**:36–48. <https://doi.org/10.1080/00031305.1983.10483087>