frontiers in Plant Science

TYPE Original Research
PUBLISHED 04 August 2026
DOI 10.3389/fpls.2026.1834912

OPEN ACCESS

EDITED BY
Lei Zhang,
Jiangsu Normal University, China

REVIEWED BY
Chensong Chen,
The University of Queensland, Australia
Raju Thada Magar,
Michigan State University, United States

*CORRESPONDENCE
Diego Jarquin
jhernandezjarqui@ufl.edu

RECEIVED 20 March 2026
REVISED 28 April 2026
ACCEPTED 26 May 2026
PUBLISHED 04 August 2026

CITATION
Proma S, Lubanga N, Sacks E, Leakey ADB, Zhao H, Ghimire BK, Lipka AE, Njuguna JN, Yu CY, Seong ES, Yoo JH, Nagano H, Anzoua KG, Yamada T, Chebukin P, Jin X, Clark LV, Petersen KK, Peng J, Sabitov A, Dzyubenko E, Dzyubenko N, Glowacka K, Nascimento M, Campana Nascimento AC, Dwiyanti MS, Bagment L, Shaik A, Garcia-Abadillo J and Jarquin D (2026) Optimizing resource allocation in *Miscanthus* breeding via sparse testing designs for genomic prediction.
Front. Plant Sci. 17:1834912.
doi: 10.3389/fpls.2026.1834912

Optimizing resource allocation in *Miscanthus* breeding via sparse testing designs for genomic prediction

Shatabdi Proma[1,2], Nelson Lubanga[3], Erik Sacks[2,4], Andrew D. B. Leakey[2,4,5,6,7], Hua Zhao[8], Bimal Kumar Ghimire[9], Alexander E. Lipka[4], Joyce N. Njuguna[2], Chang Yeon Yu[10], Eun Soo Seong[10], Ji Hye Yoo[10], Hironori Nagano[11], Kossonou G. Anzoua[11], Toshihiko Yamada[11], Pavel Chebukin[12], Xiaoli Jin[13], Lindsay V. Clark[14], Karen Koefoed Petersen[15], Junhua Peng[16], Andrey Sabitov[17], Elena Dzyubenko[17], Nicolay Dzyubenko[17], Katarzyna Glowacka[18], Moyses Nascimento[19], Ana Carolina Campana Nascimento[19], Maria S. Dwiyanti[20], Larisa Bagment[21], Ansari Shaik[1,2], Julian Garcia-Abadillo[1] and Diego Jarquin[1,2]*

[1]Agronomy Department, University of Florida, Gainesville, FL, United States, [2]Center for Advanced Bioenergy and Bioproducts Innovation, University of Illinois at Urbana Champaign, Urbana, IL, United States, [3]Insitute of Biological, Environmental and Rural Sciences, Aberystwyth University, Aberystwyth, United Kingdom, [4]Department of Crop Sciences, University of Illinois at Urbana Champaign, Urbana, IL, United States, [5]Institute for Genomic Biology, University of Illinois at Urbana Champaign, Urbana, IL, United States, [6]Department of Plant Biology, University of Illinois at Urbana Champaign, Urbana, IL, United States, [7]Center for Digital Agriculture, University of Illinois at Urbana Champaign, Urbana, IL, United States, [8]College of Plant Science and Technology, Huazhong Agricultural University, Wuhan, China, [9]Bio-Herb Research Center, Kangwon National University, Chuncheon, Republic of Korea, [10]Division of Bioresource Sciences, Kangwon National University, Chuncheon, Republic of Korea, [11]Field Science Center for Northern Biosphere, Hokkaido University, Sapporo, Japan, [12]FSBSI "FSC of Agricultural Biotechnology of the Far East named after A.K. Chaiki", Ussuriisk, Russia, [13]Zhejiang Key Laboratory of Crop Germplasm Innovation and Utilization, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou, China, [14]Research Scientific Computing, Seattle Children's Research Institute, Seattle, WA, United States, [15]Schroll Medical ApS, Arslev, Denmark, [16]Spring Valley Agriscience Co. Ltd., Jinan, Shandong, China, [17]Vavilov All-Russian Institute of Plant Genetic Resources, St. Petersburg, Russia, [18]Department of Biochemistry, University of Nebraska-Lincoln, Lincoln, NE, United States, [19]Laboratory of Intelligence Computational and Statistical Learning, Department of Statistics, Federal University of Viçosa (UFV), Viçosa, Brazil, [20]Research Faculty of Agriculture, Hokkaido University, Sapporo, Japan, [21]N.I. Vavilov All-Russian Institute of Plant Genetic Resources, St. Petersburg, Russia

Phenotyping high-biomass perennial crops is laborious and the rate of genetic gain in conventional perennial crop breeding programs is typically low. So, it is especially important to identify methods that produce efficiency gains in the breeding process. *Miscanthus* is a C4 perennial grass with favorable characteristics for producing biomass as a feedstock for biofuels and diverse bio-based products. Increasing biomass yield will increase profitability and environmental benefits, so it is a key target for *Miscanthus* breeding. In addition, the identification of well-adapted genotypes across a wide range of environmental conditions requires the establishment of multi-environment trials (METs). Sparse testing is a genomic prediction-based strategy that reduces the phenotyping costs in METs by selecting a subset of genotypes to evaluate in a subset of environments and then predicts the performance of the unobserved genotype-environment combinations. A *Miscanthus sacchariflorus* (MSA) population comprising 336

genotypes observed across three environments was analyzed implementing sparse testing designs. Three prediction models considering main effects (environments, genotypes, genomic) and interaction effects (genotype-by-environment; G×E interaction) were implemented for forecasting dry biomass yield (YDY), total culm (TCM), average internode length (AIL), and culm node number (CNN). Multiple calibration sets based on different compositions and sizes were considered to evaluate performance in terms of the predictive ability (PA) and the mean square error (MSE) for a fixed testing set size. The training set size ranged from 52 to 112 to predict a fixed set of 224 unobserved genotypes across all three environments. The results showed that the model accounting for G×E interaction consistently presented the highest PA and the lowest MSE: for CNN (PA: ~0.77, MSE: ~0.5) and YDY (PA: ~0.70, MSE: ~1.3) while for TCM and AIL these ranged from ~0.28 to 0.41 and ~1.3 to 4.3, respectively. Overall, varying training sets and allocation strategies did not affect PA and MSE, with 52 non-overlapping and 0 overlapping genotypes per environment as the optimal cost-effective allocation framework. This suggests that implementing sparse testing designs could significantly reduce phenotyping costs by fivefold, without compromising PA in breeding programs for perennial crops such as *Miscanthus*.

KEYWORDS

genomic prediction (GP), *Miscanthus sacchariflorus* (MSA), multi-environment trials (METs), resource allocation, sparse testing designs

1 Introduction

Biomass crops offer a source of feedstock that can be used to help meet demand for clean fuels and diverse biobased products, while also improving soil health and water quality (Metzger and Hüttermann, 2009; Ladanai and Vinterbäck, 2009; Saha et al., 2013; Kreig et al., 2019; Sutton et al., 2025). *Miscanthus* is a perennial species that holds the potential as a biomass grass for a wide range of applications (Chung and Kim, 2012). It can be used to produce lignocellulosic ethanol, generation of electricity and heat through gasification or direct combustion, and for the manufacture of paper, construction materials, biodegradable plastics, animal bedding, mulch, and livestock feed (Clifton-Brown et al., 2001; Clifton-Brown and Lewandowski, 2002).

Currently, commercial *Miscanthus* biomass production in Europe and North America relies predominantly on a single sterile triploid clone of *Miscanthus* × giganteus (M.× g), an interspecific cross between *Miscanthus sacchariflorus* (MSA) and *Miscanthus sinensis* (MSI) (Hodkinson and Renvoize, 2001). However, the standard clone M.× g has insufficient winter hardiness in northern Europe and the northern US Midwest (Clifton-Brown and Lewandowski, 2000; Dong et al., 2019). Moreover, cultivating a single clone on a large-scale basis faces risks associated with pests and diseases (Njuguna et al., 2023). Therefore, new high-yielding and climate-resilient *Miscanthus* varieties must be developed to withstand the ever-changing climate.

There is a set of challenges common to breeding highly productive perennial crops that must be addressed to fully exploit their beneficial characteristics, regardless of whether they are grasses or woody species. *Miscanthus* is a valuable model system to address these issues. Relative to annual crops, *Miscanthus* breeding is challenging because the crops has to establish for 2-3 years before high-quality yield data can be gathered (Chung and Kim, 2012). The large size of the crop makes phenotyping challenging, although high-throughput remote-sensing based techniques are being developed to alleviate this bottleneck (Varela et al., 2022, Varela et al., 2025).

The initial cost of establishing *Miscanthus* is also significant because the high yielding, commercial varieties are vegetatively propagated via rhizome (Lewandowski et al., 2003; Anzoua et al., 2011). Multi-environmental trials (METs) are important in *Miscanthus* breeding programs to identify both locally adapted and superior genotypes that perform well across environments (Njuguna et al., 2023). However, field evaluation of *Miscanthus* genotypes across multiple environments further increases costs and labor. To optimize both phenotyping cost and the accurate selection of improved cultivars, efficient approaches need to be implemented (Aoun et al., 2021).

Genomic selection (GS), i.e. employing genomic prediction (GP) models in practical selection, has emerged as a powerful tool in plant breeding owing to the advancement of high-throughput sequencing and the availability of genome-wide single-nucleotide polymorphisms (SNPs). The concept of GS was introduced by Meuwissen et al. (2001) to overcome the challenges posed by having a larger number of predictors (p) compared to the records available (n) for model fitting. However, Bernardo (1994) was the first who proposed to use genomic data in the form of Restriction Fragment Length Polymorphism (RLFP) for predicting the performance of unobserved maize hybrid crosses. GS uses a GP model utilizing whole genome marker data to predict the performance of the unobserved individuals via their genomic profiles only (McGaugh et al., 2021; Persa et al., 2021).

In GP, phenotypic and genotypic data are applied to calibrate and optimize various models (Jarquin et al., 2021). Usually, before

embarking on using GS in real scenarios, the levels of PA that can be reached with the available data are estimated by considering cross-validation studies mimicking realistic prediction scenarios (Cheng et al., 2021; McGaugh et al., 2021). For this, all the available data is split into two non-overlapping partitions, training and testing sets and the former is used to predict the latter partition. The training set is composed of individuals who have both phenotypic and genomic information, while the testing set contains untested individuals with only genomic information (Isidro et al., 2015). Then, the potential of untested genotypes in the testing set is evaluated via their predicted values.

Since the development of improved cultivars requires the establishment of METs to evaluate stability and local adaptation, it is necessary to consider models that accurately capture differences among genotypes in their response to environmental stimuli (Rincent et al., 2017). In most cases, the response patterns (rankings) change from one set of environmental conditions to another for the same group of genotypes (Crossa et al., 2004). This change in the response patterns, when it occurs, is referred as the presence of the genotype-by-environment (G×E) interaction (Jarquin et al., 2021). The breeding process can be sped up by ensuring an accurate selection of the best performing genotypes in advance while skipping field testing. Thus, integrating G×E into the GP models holds the potential to speed up the breeding process by ensuring an accurate selection of genotypes while decreasing the number of individuals to be evaluated in METs.

Previous studies have shown that the incorporation of G×E interactions into GP models could potentially improve predictive ability. For example, Jarquin et al. (2014a) proposed a model that implicitly allows the inclusion of the interaction between each maker SNP and each environmental covariate or environment via covariance structures. Meanwhile, Lopez-Cruz et al. (2015) presented a model that explicitly allows the incorporation of all contrasts between marker SNPs and environments. The prediction accuracies of the models are assessed using cross-validation (CV) schemes where the datasets are conveniently divided into training and testing populations mimicking real prediction problems of interest to breeders. Two of the main cross-validation schemes utilized are CV1 and CV2. CV1 predicts novel genotypes that were never tested in any environment, while CV2 predicts observed genotypes in incomplete field trials.

Sparse testing (Jarquin et al., 2020) is an innovative augmentation for field evaluation to leverage G×E in METs. It can increase the testing capacity of the genotypes at a fixed cost by allowing more genotypes to be assessed in target environments. Alternatively, it also offers the possibility of reducing phenotyping costs for a given set of genotypes and environments of interest (Garcia-Abadillo et al., 2024). In sparse testing, genotypes are split across environments and analyzed using different GP models to estimate their specific performance. The best design is determined by evaluating the relationship between the prediction accuracies of the models and the testing capacity of the designs. In other words, the best sparse testing design quantifies the accuracy of GP models with limited in-fields testing efforts to achieve acceptable predictive ability, ensuring efficient resource allocation without sacrificing model's performance.

The optimal designs in sparse testing balance the trade-off between high accuracy, reduced training set size and diverse composition. This is achieved by calibrating training sets either by using a large set of overlapped genotypes in a few environments or sets of few common genotypes across multiple environments. The accuracy of predicting unobserved genotypes in METs is influenced by *a*) the number of overlapping genotypes across environments, *b*) the number of environments where each genotype is observed, *c*) the implemented prediction model, and *d*) the average number of phenotypes per genotype (Garcia-Abadillo et al., 2024). Sparse testing has been implemented in different plant breeding programs in major crops including maize (Jarquin et al., 2020), wheat (Atanda et al., 2022), soybean (Persa et al., 2023) and sugarcane (Garcia-Abadillo et al., 2024).

The main aims of this study are to *1*) explore the potential of implementing sparse testing designs in multi-environment *Miscanthus* trials and, *2*) compare the prediction accuracies of three prediction models. The implemented metrics for model evaluation were the correlation between predicted and observed values (PA) and the mean square error (MSE). Earlier studies have focused on using traditional GP approaches in MSA and MSI germplasm panels (Dong et al., 2019; Njuguna et al., 2023). To our knowledge, the use of sparse testing designs has not yet been explored in *Miscanthus* breeding programs.

This study considered a diverse panel of MSA genotypes in different sparse testing designs similar to those proposed by Jarquin et al. (2020). Three types of designs including two extreme allocation strategies were considered *i*) non-overlapping genotypes (NO) where each genotype is uniquely observed across environments, *ii*) overlapping genotypes (O) where all genotypes in the training set are observed across all environments, *iii*) mixture of both NO and O scenarios. In addition, different reduced training set sizes were also considered along different compositions.

For all combinations between training set sizes and compositions, three prediction models were implemented: *i*) M1- includes the main effects of environment and genotype and it is based mainly in phenotypic information where no marker information is used, *ii*) M2- includes the main effects of environment, genotype, and markers, and *iii*) M3- includes the main effects of environment, genotype, and markers and the interaction between markers and environments.

Overall, the study shows that a GP-based sparse testing approach can be used to increase the selection accuracy and reduce the cost of evaluating *Miscanthus* genotypes in multiple environments. This was achieved using G×E model (M3), which returned the best performance, with the highest PA and the lowest MSE. Additionally, across all non-overlapping and overlapping design compositions, the performance of the model remained constant. This indicates that whether *Miscanthus* breeders choose to evaluate similar genotypes across various locations or different genotypes across locations, they would not have to compromise the accuracy of the model.

2 Materials and methods

2.1 Phenotypic data

The original phenotypic dataset consisted of 590 MSA accessions comprising a diversity panel collected from the wild in East Asia representing the known geographic range of this species. These genotypes were established in four locations [Sapporo, Japan by Hokkaido University (HU); Urbana, Illinois by the University of Illinois (UI); Chuncheon, South Korea by Kangwon National University (KNU) and Zhuji, China by Zhejiang University (ZJU)] between 2016 and 2017 as described by Njuguna et al. (2023). Briefly, productivity was phenotyped by assessing above-ground dry biomass and 15 yield component traits at every location between two and three years after planting. Only for dry biomass, a fourth-year data was also collected at all locations resulting in three years of phenotypic information for this trait. The best linear unbiased estimates (BLUES) of each trait for each genotype in every location (i.e., across years) were provided along with the dataset. These values were used to implement sparse testing designs. Further details of how the BLUES were obtained can be found in Njuguna et al. (2023). Briefly, at each location the linear model in Equation 1 was fitted it considers replicate (R_s, $s = 1, 2, \ldots, S$), genotype (L_i, i=1, 2,…, L), and year of harvest (H_k, k=1, 2,…, k) as fixed effects plus a measurement error term ϵ_{isk}

$$Y_{isk} = \mu + L_i + R_s + H_k + \epsilon_{isk} \quad (1)$$

The model was fitted implementing the *lm* function R (R Development Core Team, 2023) and the least squares means (LS means) for each genotype $\hat{L}_i$ across years were extracted using the R-package *lsmean* (Lenth, 2016).

Out of the 16 initial traits comprising yield and its component traits only four [dry biomass (YDY), culm node number (CNN), total culms (TCM), and average internode length (AIL)] were considered for analyses due to complete information of traits across environments. In sparse testing studies, at least for demonstration purposes only, a balanced dataset is initially required (a set of common genotypes observed across locations in this case). Due to a high frequency of missing values, one of the environments (UI) was discarded from the study. After discarding genotypes with missing information for at least one trait-in-environment combination, a total of 336 individuals remained in the analyses.

2.2 Genotyping and quality control

A detailed description of the genotyping process can be found in Clark et al. (2014). Briefly, genotyping was performed using restriction site-associated DNA sequencing (RAD-seq) following the library preparation. Hence, the library samples were sequenced with Illumina Hiseq 2500 and 4000. Raw sequence reads were aligned to the MSI reference genome and SNP calling was performed using TASSEL-GBS pipeline (Bradbury et al., 2007; Mitros et al., 2020). SNPs with more than 50% missing values and a minor allele frequencies (MAF) smaller than 3%, were removed resulting in a total of 136,814 genomic covariates available for analyses. Mising values were imputed implementing a naïve imputation by considering the mean per columns for each SNP.

2.3 Allocation scenarios

In a practical plant breeding program, only a subset of genotype-in-environment combinations can be evaluated in METs due to budget limitations and seed availability. Sparse testing designs contribute to implementing strategies for allocating genotypes across environments that extrapolate the whole population (Jarquin et al., 2020; Garcia-Abadillo et al., 2024). To implement the sparse testing strategies, the objective entails designing a trainin population to accurately predict the performance of the unobserved genotype-in-environment combinations. In general, according to its composition the calibration set can be classified in the following types:

A. *Overlapping (O) genotypes:* It resembles the original concept of genomic selection of predicting the performance of unobserved or newly developed genotypes in one or several environments where other genotypes were already observed. Usually, this manner to partition the available information in training and testing sets is also known as CV1 scheme.
B. *Non-overlapping (NO) genotypes:* This design fits perfectly in the METs implementation where some genotypes are observed in some environments but not in others in a sparse arrangement. In principle, this manner of composing training sets is regarded as CV2 scheme. The most extreme case considers observing only once each genotype across environments resulting in the non-overlapping design.
C. *No-overlapping/Overlapping genotypes:* Alternative designs can be form combining sets of overlapping and non-overlapping genotypes. Hence, the original two designs can be modified into different fractions of NO/O genotypes.

In the present study, considering a total population of 336 genotypes, assessed in three environments, results in 1,008 phenotypes (336×3). Conveniently, initial training and testing set sizes were selected. Later, the training set sizes were reduced to allow evaluate their impacts in predictive ability. A perfect design (all genotypes in all environments) allows to systematically evaluate predictive ability of the different training compositions. However, it is not necessary for implementing sparse testing designs since these will work similarly with unbalanced data. With this data, three different types of designs were implemented, all of which had a fixed testing set size of 224 genotypes per environment (i.e. 672 (224×3) phenotypes across the three environments). Correspondingly, the training set size varied from a maximum of 112 genotypes observed per environment (i.e. 336 (112×3) phenotypes across the three environments) to a minimum of 52 genotypes per environment (i.e. 156 (52×3) phenotypes across the three environments). training set sizes varied in step sizes of 10 between 112 and 52 (e.g. 102, 92, 82, 72, and 62).

1. *Completely non-overlapping allocation design (NO/O = 112/0 genotypes):* In this design, 112 non-overlapping genotypes were observed in each environment (i.e. 112 genotypes uniquely observed in one environment), with zero

10.3389/fpls.2026.1834912

overlapping genotypes (i.e. 0 genotypes assessed in all three environments).

2. *Completely overlapping allocation design (NO/O = 0/112 genotypes):* In this design, 0 non-overlapping genotypes were observed in each environment (i.e. 0 genotypes uniquely observed in one environment), with between 52 and 112 overlapping genotypes (i.e. genotypes assessed in all three environments).
3. *Mixed allocation design (NO/O = 102/10 varying to 2/110 genotypes:* In these designs, data included a mixture of phenotypic information for non-overlapping genotypes that were observed in a single environment and overlapping genotypes that were observed in all three environments.

Many possible combinations of non-overlapping and overlapping genotypes could be used given the available dataset; however, sets of 10 common genotypes were conveniently traded to study prediction patterns of the training set composition in PA. Consider C represents the common set of genotypes across the three environments, U is the unique set of genotypes that are different across the three environments, and L is the number of environments. Then, the total of genotypes-in-environment combinations in the calibration set is $(L\times C) + (L\times U) = [L\times(C+U)]$ (Garcia-Abadillo et al., 2024).

In this study, the calibration sets were constructed to include the mixing extremes of XX/0, 0/XX, and mixed type XY/XW in function of different per environment sample sizes varying as follows 112, 102, 92, 82, 72, 62, and 52. The details of the different designs can be found in Table 1 where the rows represent the training set size and columns present different combinations or levels of non-overlapping/overlapping sets. For each specific calibration size and composition, 10 random re-samplings of genotype selections were considered. The different training set compositions can be visualized in Appendix Supplementary Figure 1.

2.4 Prediction models

A breeding value (BV) represent the genotype's potential despite the set of environmental conditions where a given genotype is observed. In principle, GS attempts to predict the BV of unobserved genotypes (GEBV) using genomic information as a proxy to connect with already observed ones. Considering the interest centered on only one environment, the GEBV represents our best guest of BV with the information given. In this case, the phenotypic performance is composed of a sum between the actual BV and a deviation due to the specific set of environmental conditions of the given environment (E) shaping the genotypic response.

However, breeders are interested in a wide range of environmental conditions, environments, locations, etc. Such that phenotypic performance depends on multiple factors (fixed: environment, management; non-additive: dominance, epistasis; and genotype-by-environment G×E interactions) beyond pure additive genetics. Since the GEBV will the same across environments, the goal becomes to estimate the best as possible the performance of each (i^{th}) genotype (phenotype) on each environment (j^{th}) y_{ij}. Hence, the genotype's performance will include structural variation (fixed effects), non-additive effects, and interaction effects. For this reason, the target changes from estimating GEBV to predict plant performance.

To predict the genotype's performance in different environments three random effects prediction models were considered, one based on phenotypic information only and the other two involving genomic information as main and interaction effects (Jarquin et al., 2020; Garcia-Abadillo et al., 2024). M1: E+L, is the base model that considers the genotype and environment main effects only. These effects are assumed to follow independent and identical (*iid*) normal distributions. Hence, in this case, there is not a mean to connect calibration and testing sets when the phenotypic information is absent for genotypes across environments.

M2: E+L+G, extends the M1 model by including the marker information as a main effect via covariance structures via **G**. This model allows the borrowing of information between training and testing for never-observed genotypes. One expected drawback of this model is that estimated genomic values of a particular genotype across environments are identical, impeding modeling of phenotypic variation across different environments.

M3: (E+L+G+GE) is the extension of M2 with the inclusion of the genomic-by-environment (G×E) interaction term. This model allows genotype-specific responses in each environment, and these are simulated with a common genomic effect or intercept plus a gradient slope modeled by the interaction effect.

TABLE 1 Combinations of non-overlapping/overlapping (NO/O) allocation strategies for different calibration set sizes.

Calibration sets	Allocation designs													Testing set
112 (33.3%)	112/0	102/10	92/20	82/30	72/40	62/50	52/60	42/70	32/80	22/90	12/100	2/110	0/112	224 (66.7%)
102 (30.4%)		102/0	92/10	82/20	72/30	62/40	52/50	42/60	32/70	22/80	12/90	2/100	0/102	224 (66.7%)
92 (27.4%)			92/0	82/10	72/20	62/30	52/40	42/50	32/60	22/70	12/80	2/90	0/92	224 (66.7%)
82 (24.4%)				82/0	72/10	62/20	52/30	42/40	32/50	22/60	12/70	2/80	0/82	224 (66.7%)
72 (21.4%)					72/0	62/10	52/20	42/30	32/40	22/50	12/60	2/70	0/72	224 (66.7%)
62 (18.5%)						62/0	52/10	42/20	32/30	22/40	12/50	2/60	0/62	224 (66.7%)
52 (15.5%)							52/0	42/10	32/20	22/30	12/40	2/50	0/52	224 (66.7%)

The column in the far left presents different sampling sizes (and percentages with respect to all the potential genotype-in-environment combinations). The column in the far right presents the fixed within-environments prediction set size (and percentage). The columns in the middle show the different allocation designs as a function of the training set size. Completely overlapping designs are shown in orange cells. The completely non-overlappings designs are shown in blue cells. Mixed allocation designs are in green cells.

10.3389/fpls.2026.1834912

The performance of the models was evaluated for all training set compositions and sizes. The metrics considered are the PA [as the Pearson correlation coefficient between observed and predicted values] and the MSE [mean of the squared of the differences between observed and predicted values]. Further details of these models are provided below.

M1: E+L [Environment + genotype].

Consider that y_{ij} represents the adjusted genotypic effect $\hat{L}_{i,j}$ (obtained in eq. 1) of the i^{th} $(i = 1, 2, \ldots, L)$ genotype in the j^{th} $(j = 1, 2, \ldots, J)$ environment and it is explained as the sum of the common mean μ, the random effect of the j^{th} environment (E_j), the random effect of the i^{th} genotype (L_i), plus a random error term (ε_{ij}). Here, the environment effect E_j was modeled considering only the environment ID as factor rather than including environmental information. For this reason, we assumed environments were independent between them. Hence, the only information we implemented to model relationships was the phenotypic information between common genotypes. The same assumption was considered for the more complex models here presented involving this term. Further assumptions are made regarding the relationships among the random terms E_j, L_i and ε_{ij} by considering these independent between them.

In addition, for each random term their corresponding levels are considered independent and identical normal distributed (*iid*). Hence, the model Equation 2 can be written as follows.

$$y_{ij} = \mu + E_j + L_i + \varepsilon_{ij} \tag{2}$$

$$\begin{pmatrix} \boldsymbol{E} \\ \boldsymbol{L} \\ \boldsymbol{\varepsilon} \end{pmatrix} \sim N\left(\begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix}, \begin{bmatrix} \boldsymbol{I}_J\sigma_j^2 & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_L\sigma_i^2 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}_{L\times J}\sigma_\varepsilon^2 \end{bmatrix} \right)$$

where $\boldsymbol{I}_J$, and $\boldsymbol{I}_L$ are the identity matrices according to the levels of environments and genotypes.

M2: E+L+G [environment + genotype + genomic (markers)].

M2 is an extension of M1 where the genomic information is introduced via a covariance structure. Here, g_i represents the genomic effect of the i^{th} genotype and it is computed as the linear combination between p markers (X_{im}) and their respective effects (b_m) $\{m = 1,2,\ldots,p\}$ such that $g_i = \sum_{m=1}^{p} X_{im} b_m$, where $b_m \sim N(\boldsymbol{0}, \sigma_b^2)$ with σ_b^2 representing the corresponding variance component. Stacking all genomic into a single vector $g = \{g_i\}$and by properties of the multivariate normal distribution we have that $\boldsymbol{g} \sim N(\boldsymbol{0}, \boldsymbol{G}\sigma_g^2)$ where $\boldsymbol{G} = \frac{\boldsymbol{X}\boldsymbol{X}'}{p}$ is the genomic relationship matrix describing genomic similarities between pairs of genotypes, X is the (standardized by columns) matrix of marker SNPs, and $\sigma_g^2 = p \times \sigma_b^2$ is the associated variance component (VanRaden, 2008). The resulting model Equation 3 can be represented as follows

$$y_{ij} = \mu + E_j + L_i + g_i + \varepsilon_{ij} \tag{3}$$

$$\begin{pmatrix} \boldsymbol{E} \\ \boldsymbol{L} \\ \boldsymbol{g} \\ \boldsymbol{\varepsilon} \end{pmatrix} \sim N\left(\begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix}, \begin{bmatrix} \boldsymbol{I}_J\sigma_j^2 & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_L\sigma_i^2 & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{G}\sigma_g^2 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}_{L\times J}\sigma_\varepsilon^2 \end{bmatrix} \right)$$

This model can be used to predict the non-phenotyped individuals as it allows the borrowing of information between individuals through the genomic relationship matrix.

M3: E+L+G+GE [environment + genotype + genomic + genomic markers-by-environment (G×E) interaction].

In this model, M2 is extended by adding the random term gE_{ij} which corresponds to the interaction between the i^{th} genotype and j^{th} environment. Stacking all the interaction effects into a single vector and following Jarquin et al. (2014a) we have that $\boldsymbol{gE} = \{gE_{ij}\} \sim N(\boldsymbol{0}, \boldsymbol{Z_g}\boldsymbol{G}\boldsymbol{Z_g}' \circ \boldsymbol{Z}_E\boldsymbol{Z}_E' \ \sigma_{gE}^2)$ where σ_{gE}^2 is the corresponding variance component, $\boldsymbol{Z}_g$ and $\boldsymbol{Z}_E$ are the design matrices connecting phenotypes to genotypes and environments, respectively, and "$\circ$" is the Hadamard product indicating cell-by-cell multiplication between two matrices of the same order. Thus, the model Equation 4 can be written as follows

$$y_{ij} = \mu + E_j + L_i + g_i + gE_{ij} + \varepsilon_{ij} \tag{4}$$

$$\begin{pmatrix} \boldsymbol{E} \\ \boldsymbol{L} \\ \boldsymbol{g} \\ \boldsymbol{gE} \\ \boldsymbol{\varepsilon} \end{pmatrix} \sim N\left(\begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix}, \begin{bmatrix} \boldsymbol{I}_J\sigma_j^2 & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_L\sigma_i^2 & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{G}\sigma_g^2 & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{Z}_E\boldsymbol{Z}_E' \circ \boldsymbol{Z}_g\boldsymbol{G}\boldsymbol{Z}_g' \ \sigma_{gE}^2 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}_{L\times J}\sigma_\varepsilon^2 \end{bmatrix} \right)$$

The previously mentioned model M2 returns common genetic effects for the same individuals across environments while with M3 the effects are particular to each environment.

2.5 Cross-validations

To evaluate the prediction ability (PA) of the different sparse testing designs (Table 1) two principal cross-validation schemes (CV1 and CV2) were used as reference. The phenotypic data for the genotypes across environments was available, however, a fraction of the data was masked intentionally as missing values (NA) following different designs. As explained earlier, non-overlapping genotypes follow the basis of the CV2 scheme while overlapping genotypes follow the basis of the CV1 scheme. To elaborate, in the CV2 scheme, PA is driven by utilizing genomic similarities among genotypes and correlations between environments, thus borrowing phenotypic information from genotypes across environments. CV1 presents the situation where some genotypes are never tested in any of the environments and predicting their performance relies on genomic information and on other genotypes that were tested in these fields.

The allocation designs (combinations of NO/O genotypes) in Table 1 show the gradual transitions between CV2 and CV1 modifying the numbers of NO and O genotypes in environments through random cross-validations. For example, the design 112/0 in Table 1 is a pure CV2 scheme while 0/112 is a pure CV1 scheme. In addition, with more NO and fewer O genotypes, the design closely aligns with the pure CV2 scheme. As the number of O genotypes increases and NO genotypes decrease across designs (from the left-hand side toward the right-hand side) the design gets closer to the CV1 scheme. Hence, Table 1 exhibits situations of both CV1 and the pure CV2 schemes based on the numbers of NO/O genotypes across designs.

Here, the effects of the training compositions set sizes of the different sparse testing designs were assessed using the three above-mentioned prediction models. In all cases, PA was evaluated as the within-environment Pearson correlation between the predicted and the observed values of the 224 genotypes in the testing set. In addition, MSE was also calculated on a trial basis where lower MSE values reflected a better performance of the models.

2.6 Heritability

Since all genotypes were observed in all environments, when conducting the full data analysis, we combined both the line σ^2_L and genomic σ^2_G variance components as proxy of the genetic component. For each model we considered different methods to compute the broad sense heritability. For the model based only on main effects (M1) we computed the broad sense heritability as follows:

$$H^2 = \frac{\sigma^2_L + \sigma^2_G}{\sigma^2_L + \sigma^2_G + \sigma^2_\varepsilon}$$

While for M2 (E+L+G+GE) the entry-means heritability was computed similarly as in Jarquin et al. (2014b):

$$H^2 = \frac{\sigma^2_L + \sigma^2_G}{\sigma^2_L + \sigma^2_G + \sigma^2_{GE}/e + \sigma^2_\varepsilon}$$

where σ^2_{GE} is the variance of the interaction between genotypes and environments; σ^2_ε is the residual variance and e is the number of environments.

2.7 Software

All the models were fitted using the BGLR package (Pérez and de Los Campos, 2014) in R statistical programming language (R Core Team, 2023). A total of 12,000 iterations and a bur-in process of 2,000, thinning of five and default hyper-parameters were considered to fit the models.

3 Results

The average results for PA and MSE across the three environments, 10 replicates of the different allocation designs, three prediction models, and four traits are presented in Figures 1, 2. For each trait, the PA (Appendix Supplementary Figures 2-5) and MSE (Appendix Supplementary Figures 6–9) values were computed between observed and predicted values within each environment. Afterward, the mean average PA and MSE were calculated across environments by taking the mean of the within-environment values.

3.1 Variance components and entry-means heritability

Table 2 shows the percentage of explained variance for two GP models across three environments for each of the four traits. Both models used complete phenotypic records without missing data across all environments. The results indicated that the environment accounted for the largest proportion of phenotypic variance (σ^2_E), ranging from 25% to 43% across different models and traits. The M1 model consistently exhibited higher environmental variance than the M2 model for all traits. For instance, under the M1 model, YDY and TCM it explained approximately 35% and 43% of the phenotypic variance, whereas under the M2 model, it explained 34% and 40%, respectively. The line effect (σ^2_L) explained the least variance among all variance components (between 5% and 7% across traits). While the marker main effects (σ^2_G) explained between 6% and 15% of the variance. The interaction term $\sigma^2_{G\times E}$ explained notably higher variance in all traits (YDY: 33%; TCM: 26%; AIL: 28%, and CNN: 26%). The percentage of variability addressed by the residual error (σ^2_ε) was significantly smaller in M2 model, ranging from 42% to 63% across traits compared to the corresponding from M1. For instance, in TCM, the unexplained variance accounted for by the error term was 43% in M1, while it was only 23% in M2, representing a 45% decrease in M2.

The amount of phenotypic variability $\sigma^2_{G\times E}$ explained by the G × E component with respect to the genetic main effects combined (σ^2_L, and σ^2_G) was always superior across traits. For YDY the amount of variability explained by the interaction term compared to the genetic signal was 107.3% higher, while was it 137.8% for TCM, 105.5% for AIL, and 46.8% for CNN.

The estimated heritability values computed with M1 were 0.300 for YDY, 0.253 for TCM, 0.203 for AIL, and 0.344 for CNN. While the resulting entry means heritability computed with the variance components from M2 were 0.358 (YDY), 0.254 (TCM), 0.245 (AIL), and 0.442 (CNN). The estimated heritability obtained with M2 was always superior with M2 respect to M1. For YDY the estimated heritability increased by 19.2%, while by 0.45% for TCM, 20.6% for AIL, and 28.5% for CNN.

3.2 Predictive ability

For YDY, models M3 (E+L+G+GE) and M2 (E+L+G) always outperformed M1 (E+L) with M3 obtaining the highest PA (~ 0.7) at the 112/0 design (Figure 1A). This result was higher around 13% and 23% compared to M2 (~0.62) and M1 (~0.58). The superiority of M3 remained consistent across all designs (from the left end at 112/0 toward the right end at 0/112). In other words, M3 was not influenced by the increasing number of overlapping genotypes across designs (middle as well as right sides of the plots). The PA in M2 slowly increased as more overlapping genotypes were added. For example, at designs 102/10, 92/20, and 72/40, the PA in M2 was ~ 0.65, ~0.66, and ~0.68, respectively, which then plateaued to ~0.70 at the 40/72 design. M1 started with a PA ~0.58 at 112/0; however, as expected, it significantly dropped toward 0 as the number of overlapping genotypes increased.

Considering the reduced training sets, M2 and M3 were not significantly impacted by decreasing the sample size of the training sets. As expected, the PA in M1 reduced as the training set sizes decreased. In this case, the results maintained a decreasing trend rising number of overlapping genotypes across designs as well.

Similarly to the previous trait, there was a superiority of M3 concerning the PA over M2 and M1 for TCM (Figure 1B). The PA in M1 significantly reduced with the training set sizes. For M3 and

10.3389/fpls.2026.1834912

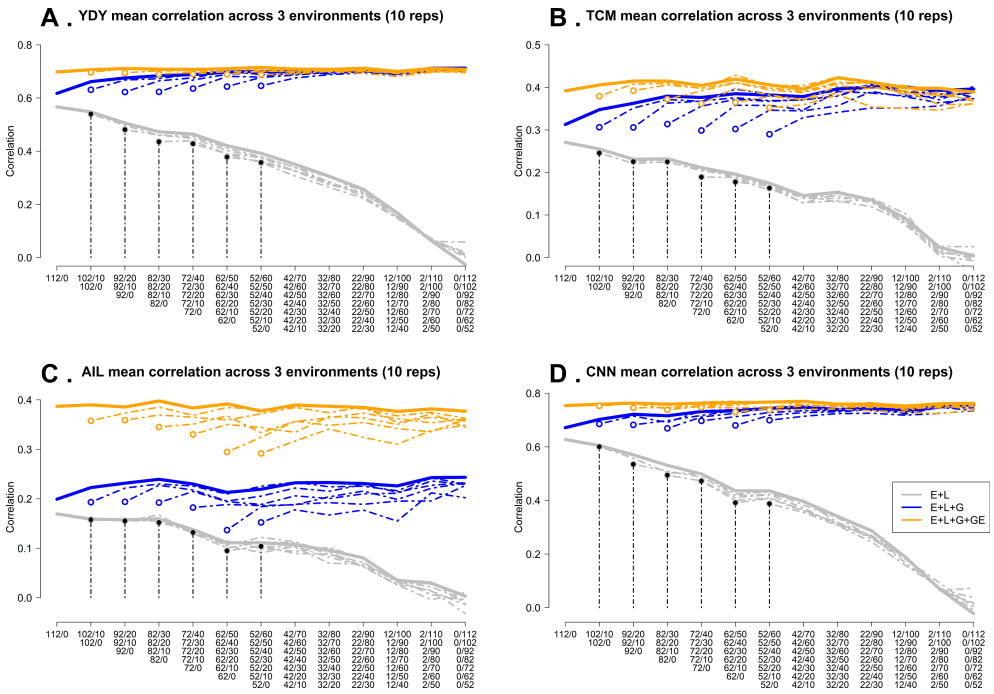

FIGURE 1
Mean PA (correlation between predicted and observed values) for four traits in varying training sets and allocation designs. **(A)** dry biomass yield (YDY), **(B)** total culm numbers (TCM), **(C)** average internode lengths (AIL), **(D)** culm node numbers(CNN) across three environments in three genomic prediction models, (M1: E+L; M2: E+L+G; M3: E+L+G+GE). Bold lines show the average PA for the largest calibration set (112 genotypes). Dotted lines show the reduced training sets (102, 92, 82, 72, 62, and 52 genotypes). X-axis shows the compositions of different non-overlapping (NO)/ overlapping (O) allocation designs. Leftmost side of the X-axis shows compositions for completely NO genotypes with 0 overlaps across environments for different testing set sizes (e.g., 112/0 and 52/0). From left to right, the NO/O ratio decreases with increasing overlapping genotypes. Vertical, black-dashed lines indicate PA for different reduced training sets (e.g., 112/0, 92/0, and 52/0).

M2, the PA ranged between ~0.35 and ~0.40, and between ~0.28 and ~0.39, respectively [range between the smallest training set and the largest training set] for varying calibration set sizes. Moreover, for M2, the PA gradually increased as the number of overlapping genotypes. For example, the PAs in M2 at designs 112/0, 82/30, 32/80, and 12/100 were ~0.31, ~0.37, ~0.38, and ~0.39, respectively, for the full training set size. In contrast, M1 noticeably dropped after more overlapping genotypes were added. The reduced training set sizes slightly influenced both M2 and M3 results. In nearly all situations with a few exceptions, M3 delivered the best results over M2 under the same designs.

For AIL, M3 clearly outperformed both M2 and M1 across designs (Figure 1C). As more common genotypes were added, the PA did not significantly change for M2 and M3, while it gradually dropped for M1. In this case, the reduced training set sizes impacted the results of both M2 and M3 models. However, M3 was always superior for the same compositions of training sets and allocation designs over M2. The PA between the smallest and the largest training set ranged from ~0.28 to ~0.41 for M3 and from ~0.13 to ~0.23 for M2. As expected and similar than for the other traits, M1 was significantly impacted with reduced training set sizes, as well as for the increasing number of overlapping genotypes.

Finally, similar patterns regarding the performance of M1, M2, and M3 for all compositions of training sets and allocation designs were observed for CNN (Figure 1D). Both M2 and M3 performed better compared to M1; however, M3 achieved the highest PA ~0.77 at 112/0 and remained unchanged as the number of NO/O ratios reduced (as the number of overlapping genotypes increased). For M2, the PA gradually increased with decreasing NO/O. For example, M2 returned a PA of ~0.67 at 112/0, ~0.70 at 72/40, and it reached a maximum plateau of ~0.77 at 2/100. However, although the best results of M2 are comparable to those from M3, the latter do not require a high number of overlapping genotypes. In contrast, M1 always returned the lowest PA across all designs, showing a gradual drop as the number of overlapping genotypes was increased.

In summary, consistently M3 returned the best results while as expected M1 showed the worst results for all traits in all compositions of calibration set sizes and allocation designs. For YDY and CNN, M2 and M3 showed smaller differences between those from the other traits, with M3 outperforming M2 considering the larger number of non-overlapping genotypes in the designs. In addition, for all traits the PA remained almost unchanged with M3 as more genotypes were added into the designs. For M2, the PA slowly increased as the number of overlapping genotypes increased as well. While for M1 it was consistently reduced as the ratio of NO/O decreased. In general, accuracy was lower in AIL and TCM compared to YDY and CNN. For AIL, the PA obtained with M3 and M2 was negatively affected by reducing the training set sizes but remained stable for YDY, TCM and CNN.

3.3 Mean square error

Figures 2A–D display the average MSE for YDY, TCM, AIL, and CNN respectively for M1 (E+L), M2 (E+L+G), and M3 (E+L+G

10.3389/fpls.2026.1834912

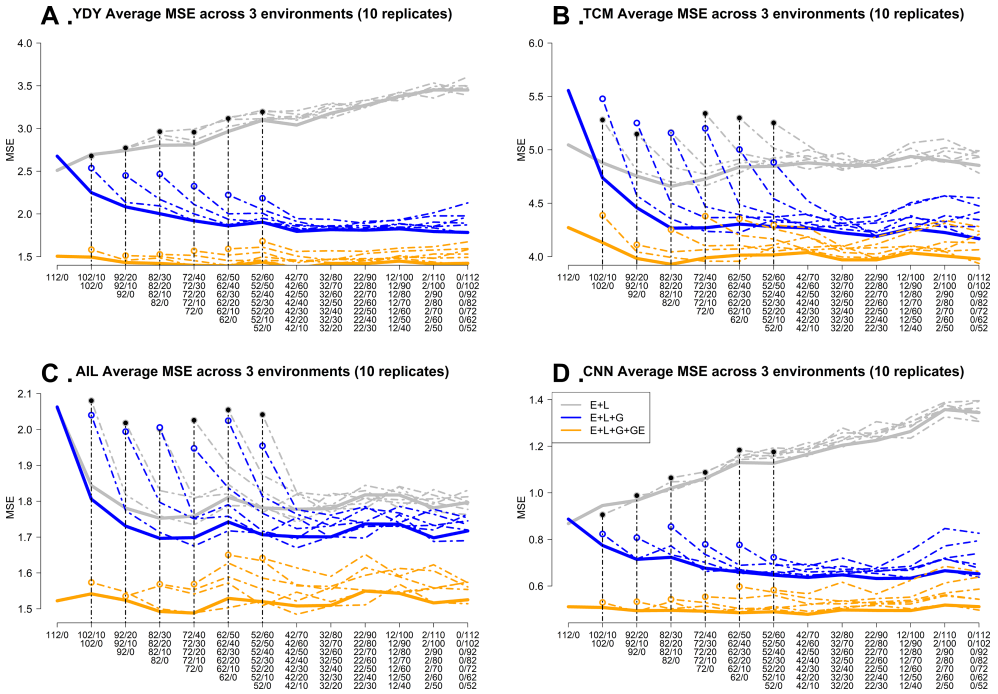

FIGURE 2
Mean MSE values for four traits in varying training sets and allocation designs. **(A)** dry biomass yield (YDY), **(B)** total culm numbers (TCM), **(C)** average internode lengths (AIL), **(D)** culm node numbers (CNN) across three environments in three genomic prediction models, (M1: E+L; M2: E +L+G; M3: E+L+G+GE). Bold lines show the average PA for the largest calibration set (112 genotypes). Dotted lines show the reduced training sets (102, 92, 82, 72, 62, and 52 genotypes). X axis shows the compositions of different non-overlapping (NO)/overlapping (O) allocation designs. Leftmost side of the X axis shows compositions for completely NO genotypes with 0 overlaps across environments for different testing set sizes (e.g., 112/0 and 52/0). From left to right, the NO/O ratio decreases with increasing overlapping genotypes. Vertical, black-dashed lines indicate PA for different reduced training sets (e.g., 112/0, 92/0, and 52/0).

+GE) models across all environments and 10 replicates. Here, the smallest the MSE the best. For YDY, both M3 (~1.5) and M2 (~2.4) performed better than M1 (2.5) with M3 consistently returning the lowest MSE values (Figure 2A). Moreover, M3 remained constant as the ratio of NO/O reduced across designs. At the beginning (112/0) when the genotypes were tested only once, the MSE value in M1 was slightly lower (~2.5) compared to M2 (~2.7) (Figure 2A). However, as more common genotypes were added to the designs, the values gradually increased for M1 while decreased for M2. With the reduced calibration set sizes, the MSE values in M3 increased slightly to ~1.7, while M1 significantly increased. Interestingly, with M2 the MSE values slightly reduced. Nevertheless, M3 always returned the lowest values compared to M2 and M1 across all NO/O genotype compositions (from the left-hand side toward the right-hand side).

Similar to the previous trait, M3 achieved the smallest MSE value for TCM. The values slightly varied between ~3.8 and ~4.3 across designs (Figure 2B). M1 had a slightly lower value (~5.1) at 112/0 compared to M2 (~5.6). However, for different compositions of NO/O genotypes, the MSE reduced in both M2 (~4.8) and M1 (~4.3) with M2 experiencing a more pronounced reduction. When the calibration set sizes reduced, all models were impacted; M1 and M2 showed a decreasing pattern. On the other hand, despite M3 presented a slightly increasing pattern in MSE, it constantly yielded the best results with lower MSE than M1 and M2.

For AIL, the MSE was always low with M3 compared to M1 and M2 (Figure 2C). The values for M3 ranged between ~1.48 to ~1.52 for different combinations of NO/O genotypes. Models M1 and M2 both started with a MSE value of ~2.08 at 112/0; however, for M2 it declined for the subsequent allocation schemes (designs in the middle and on the right-hand side of the plots) as the number of common genotypes increased. Although under both M1 and M2 the

TABLE 2 **Percentage of variability explained explained for two models (E +L+G and E+L+G+GE) for a total of 336 *Miscanthus sacchariflorus* population across three environments (Chuncheon, South Korea CH, Sapporo, Japan by Hokkaido University SP, and Zhuji, China ZJU) for four traits (YDY, TCM, AIL and CNN).**

		Percentage of explained variability				
Trait	Model	σ^2_E	σ^2_L	σ^2_G	$\sigma^2_{G\times E}$	σ^2_ε
YDY	E+L+G	35.3	4.6	14.8		45.3
YDY	E+L+G+GE	33.8	5.4	10.4	32.9	17.5
TCM	E+L+G	43.0	5.0	9.4		42.6
TCM	E+L+G+GE	40.2	5.3	5.5	25.7	23.2
AIL	E+L+G	28.6	6.1	8.4		56.9
AIL	E+L+G+GE	25.4	7.4	6.2	28.1	32.8
CNN	E+L+G	43.2	4.6	14.9		37.3
CNN	E+L+G+GE	42.9	5.4	12.2	25.8	13.6

σ^2_E, σ^2_L, and σ^2_G represent the variance components of the main effects of the environment, lines, and markers, respectively, $\sigma^2_{G\times E}$ corresponds to the variance component of the interaction between markers and environments, and σ^2_ε is the unexplained variance addressed by the residual error term.

MSE decreased, the reduction was more pronounced in M2 compared to M1. With reduced training set sizes, all models were impacted for different compositions of NO/O. The MSE values slightly increased in all three models with M1 and M2 showing a decreasing pattern while M3 showing a slightly increasing pattern. Despite the increasing trajectory observed in M3 it always returned the best results.

Finally for CNN, the models demonstrated similar trends as observed in YDY (Figure 2D). M3 constantly produced the lowest MSE (~0.5) compared to both M2 and M1. The prominence of M3 was shown for all combinations of NO/O ratio. On the other hand, M1 obtained a MSE value of ~0.88 while M2 of ~0.9 at the design 112/0. With the reduced NO/O ratio, the value kept increasing in M1, whereas decreasing in M2. Reduced training set sizes resulted in a marginal increase of the MSE value for M3 (~0.62). However, despite the minimal change observed in M3, it always outpaced both M1 and M2 across the comparable combinations of training sets and designs.

Although, in general, M2 and M3 both performed better than M1 in all traits, M3 always achieved the lowest MSE values for all compositions of calibration set sizes and allocation designs. Also, the MSE values varied minimally in M3 when more overlapping genotypes were added across designs.

Overall, the results in PA and MSE suggest that M3 was the most superior model in all traits for different compositions of calibration sets and sizes. In addition, under M3 the MSE remained constant as the number of overlapping genotypes increased across designs. Although in some cases, the results in M3 were slightly influenced by the reduced sample sizes based on traits, it still outperformed both M2 and M1. On the other hand, M1 returned the worst results; while, for M2 the results improved with the addition of common genotypes.

Therefore, the results indicate that evaluating different fractions of non-overlapping *Miscanthus* genotypes in all environments or repeating some overlapping genotypes in all environments would generate almost similar accuracy. The performance of M3 shows that either approach would be effective. However, the accuracy would also depend on the genetic architecture of the traits including the types of environments where these were evaluated. The observed results show that there is no optimum fraction for the number of overlapping genotypes. However, a small set of 10 to 20 overlapping genotypes could be better for research purposes to estimate environmental variability that is not confounded with genetic variability (Jarquin et al., 2020).

4 Discussion

The development of improved varieties requires establishing multi-environment trials (METs) to evaluate their performance under a wide range of environmental conditions (Brown et al., 2020; Smith et al., 2021). METs play a crucial role in effectively capturing G×E interaction by evaluating multiple genotypes in different environments (Verbyla, 2023; Argaw et al., 2025). However, setting up METs in *Miscanthus* is resource-intensive as the establishment cost due to rhizomatous breeding is high (Arnoult and Brancourt-Hulmel, 2015). Additionally, prolonged field testing that requires *Miscanthus* crops to wait for three years to yield full potential is required for phenotyping (O'Loughlin et al., 2017), which lengthens the release of commercial *Miscanthus* varieties. Hence, it is difficult to determine a strategic selection and allocation of superior *Miscanthus* genotypes across target environments.

Genomic selection has emerged as a powerful tool to inform breeders about the selection of genotypes in a more effective way, with the initial goal of boosting genetic gain (Budhlakoti et al., 2022; Alemu et al., 2024). In addition, it plays a vital role in addressing challenges with complex traits and G×E confronted in METs. Previous studies have explored different ways of incorporating G×E into the GP models (Jarquin et al., 2014a; Acosta-Pech et al., 2017). González-Barrios et al. (2019) explored resource optimization across multiple environments by using mega-environments while including G×E into the model. In the *Miscanthus* population, only a few studies have leveraged the use of GP models in METs (Slavov et al., 2014; Clark et al., 2019; Dong et al., 2019; Njuguna et al., 2023).

Clark et al. (2012) and Olatoye et al. (2020) showed the importance of relatedness between *Miscanthus* individuals while designing a training population to use in GP models. However, no previous studies have explored the effects of sparse testing allocation strategies on the accuracy of GP models for *Miscanthus* METs using different combinations of non-overlapping and overlapping genotypes. As with other perennial crops, GP-based sparse testing is of particular interest in *Miscanthus* METs because initial evaluation of a large number of *Miscanthus* genotypes requires high establishment and phenotyping costs across multiple years (Lewandowski et al., 2016). Optimum allocation of the superior *Miscanthus* genotypes can reduce the total costs of the initial trials.

In this study that considers 336 genotypes, three environments, and 1,008 phenotypic records, it was shown that GP-based sparse testing strategies could maintain high accuracy at different compositions of calibration sets. Hence, a substantial reduction of phenotyping costs can be accomplished, especially when using model M3. In addition, the accuracy remained practically unchanged as more overlapping genotypes were added to the same model. Thus, the incorporation of G×E into prediction models has the potential to borrow information from genotypes observed in target environments. However, maintaining consistent PA while adding more overlapping genotypes has less significance in *Miscanthus* breeding in North America, considering costs. This is because *Miscanthus* relies on vegetative propagation where the cost is initially associated with establishments, demanding intensive labor, management, and maintenance. Therefore, in *Miscanthus* breeding, the focus on cost-saving mainly shifts toward observing fewer genotypes while maintaining high PA in the target environments by optimizing resource allocation. In addition, the inclusion of the interaction term helps to increase the genetic signal which is leveraged to increase predictive ability.

The results regarding the decreasing calibration set sizes showed that PA remained almost unchanged across all traits, especially while using model M3. This model provided an opportunity to

increase the testing capacity of evaluating more genotypes by ensuring similar accuracy, thus saving costs associated with resource management. For example, evaluating 52 genotypes in the smallest training set in each of the three environments resulted in almost the same accuracy level as with 112 genotypes in the largest training set (336/3 = 112), when using model M3. Thus, instead of observing a total of 1,008 phenotypes for 336 genotypes-in-environment combinations (336×3 = 1,008), evaluating only 52 genotypes (52×3 = 156 phenotypes) in each of the three environments would reduce the phenotyping costs by ~85% [= (1,008-156)/1,008]. Obviously, the actual cost in real *Miscanthus* breeding programs would vary from this estimate since also differences in cost, labor, management, etc. vary among environments/locations. However, evaluating fewer genotypes would still save the costs associated with phenotyping data collection, and environmental inputs such as water use, fertilizer use, etc.

Albeit overlapping genotypes have minor importance in *Miscanthus* breeding, adding a small fraction could be useful for research intent. It could act as a baseline for comparison while estimating environmental variance across different environments. This is because, in the models, the overlapping genotypes behave as a control while separating environmental variance from the genetic variance. Another reason for adding some overlapping genotypes could be logistical restraints that limit testing genotypes in some environments but allow evaluation of them in others (Jarquin et al., 2020). Hence, including a fraction of 10 to 20 overlapping genotypes with the rest of the non-overlapping genotypes (42) distributed in three environments would be an effective approach for substantial cost savings in *Miscanthus* breeding.

4.1 Comparison of the prediction models

An important objective of this study was to investigate the effects of three GP models for varying training set sizes and training compositions across three environments and four traits. The model that returned maximum PA and minimum MSE for different combinations of NO/O genotypes for varying training sets was considered the best. Results indicated that model M1 had the lowest PA and highest MSE compared to M2 and M3. When the genotypes were observed in at least one environment, the PA was better in M1 (112/0 or 52/0). This is because, when a genotype is observed in at least one environment, its effect is blended with the effect of the environments.

The phenotypic information of the observed genotypes had therefore a strong impact on GP accuracy. However, the accuracy rapidly dropped across designs when the ratio of NO/O reduced (from the left extreme to the right extreme). It is because when the ratio of NO/O reduced, the design slowly approached toward the true CV1 scheme by 0/112. In CV1, there is no observed information available on the genotypes, and it is impossible for M1 to borrow information for the observed genotypes. As no marker information was included in M1, it could not borrow information on the genotypes through genomic similarities. Similar patterns were found in Jarquin et al. (2020) and Garcia-Abadillo et al. (2024) for maize and sugarcane datasets respectively where prediction accuracy decreased for M1 in the situation where no phenotypic information was available for the genotypes. On the other hand, MSE in M1 highlighted the opposite pattern, increasing with reduced NO/O fractions.

For all traits, M3 had the highest PA. The main reason is that M3 included G×E interaction in the model which could borrow information about the behavior of the genetic value of the genotypes across and within environments. Although M2 and M3 had nearly similar accuracies in YDY and CNN, the MSE value was lower in M3 for the same compositions of training sets and NO/O combinations. This may be attributable to the G×E term used in M3, which could significantly reduce the unexplained proportion of the total variance compared to M2. Hence, constant low MSE values in M3 compared to M2 suggested that it was the best prediction model for *Miscanthus* sparse testing designs.

Another notable trend in M3 was that it consistently returned almost the same accuracy across designs in all traits. In other words, adding more overlapping genotypes did not change the accuracy in M3. To elaborate, with the addition of overlapping genotypes, the designs shift toward CV1 where there is a lack of phenotyping information on the untested genotypes, hence depending on genomic marker data to fill missing information. In this circumstance, the G term in M3 model could borrow information of the genotypes via genomic similarities whereas the G×E term could capture the interaction more effectively. This finding also suggests that for the *Miscanthus* population, it does not matter whether complete non-overlapping genotypes are tested in each environment or if some overlapping sets are used across environments. In both cases, it is possible to reduce the phenotyping cost; however, the accuracy could vary depending on the traits and how correlated the environments are between them. Also, as discussed for same training set size similar results were obtained with the varying compositions, some of these are more plausible to occur in the real application in breeding programs where availability of materials could be a bottleneck for propagation.

These findings regarding constant PA across designs in *Miscanthus* differ from the observations reported in sugarcane (Garcia-Abadillo et al., 2024). In sugarcane, the PA decreased in the M3 model as more overlapping genotypes were added. This could be due to a strong environmental correlation or due to including small sample sizes in the training sets.

Considering the PA across traits, it was observed that M3 returned higher PA in YDY and CNN compared to AIL and TCM. Moreover, decreasing the calibration set size did not influence the accuracies in YDY and CNN but negatively impacted TCM and AIL. Also, the MSE varied minimally for YDY and CNN but varied significantly for TCM and AIL with reduced training set size in M3. With full calibration set size, the PA in TCM and AIL was ~0.41. This value was reduced by ~18 to ~ 42% respectively in TCM and AIL as the training set was reduced. Njuguna et al. (2023) reported prediction accuracies in the same *Miscanthus* population for YDY, CNN, AIL, and TCM that were 0.56, 0.53, 0.15, and 0.46, respectively. The key difference between this study and our study is that we used sparse testing designs with 33% of the genotypes in the training sets and 66% of the genotypes for the testing set.

In contrast, Njuguna et al. (2023) used more phenotypic records for their analysis. The author also stated the importance of a large

training population for high prediction accuracies in the *M. sacchariflorus* population. Hence, TCM and AIL could require a large training set size (> 112) to capture the variability for improved prediction accuracy. In contrast, for YDY and CNN, the G×E interaction term could account for most of the variances in a smaller training set size (52) and return similar accuracies as with the largest training set size (112). Another key aspect of the results obtained in this study was the use of high-density markers. For the present study, a total of around 136k was available for performing GP analysis. In practical situations, it is important to account for budget constraints that could be associated with genotyping costs along with allocation schemes. Jarquin et al. (2020) used ~62k makers for deploying sparse testing designs in the maize dataset and found reasonable accuracies yielded by M3.

In sugarcane, which is a closest species of *Miscanthus*, ~22k markers were enough for M3 to return PA varying between ~0.34 and 0.61 (Garcia-Abadillo et al., 2024). However, reducing marker density could lower accuracy for conducting sparse testing designs in *Miscanthus*. Thus along with allocation strategies, it is important to balance between marker density and PA to maximize genetic gain and cost-efficiency. If accuracy is reduced with fewer markers, alternative approaches will need to be considered in the genomic prediction models.

4.2 Practical implementation of sparse testing in *Miscanthus* breeding

Miscanthus has a broad geographical distribution related to a diverse set of germplasms with multiple adaptation ranges across Asia to Siberia (Huang et al., 2019). It is an excellent source for establishing *Miscanthus* breeding programs in North America. However, there are various challenges to identifying and selecting the best variety for the breeding programs (Clifton-Brown et al., 2019). Besides finding the best crosses, evaluating *Miscanthus* genotypes requires expensive rhizomatous or *in vitro* propagation for the establishment (Xue et al., 2015). Given the breeding trials, previous studies in *Miscanthus* focused mainly on the yield and yield component traits (Kaiser et al., 2015; Clark et al., 2019; Njuguna et al., 2023). However, extensive multilocation field trials are expensive as most yield and yield component traits cannot be measured before two to three years of establishment. There are additional costs for labor, land (Lewandowski et al., 2016), and genotyping for optimizing molecular marker use. Therefore, finding the superior genotypes and reducing the associated costs are important issues to consider while increasing genetic gain in *Miscanthus* breeding programs.

Considering *Miscanthus* breeding programs with fixed budgets and the long establishment phases, the breeders must carefully decide at the initial stages how many genotypes they should evaluate in fields. On the other hand, they should also optimize genotyping procedures for the rest of the genotypes and aim for efficient resource allocations to maximize genetic gain. The practical implementation of GP-based sparse testing in *Miscanthus* breeding shows how to optimize resource allocation while maintaining high PA and decreasing phenotyping costs by adjusting varying compositions of NO/O genotypes. Using the saved costs, more elite genotypes can be evaluated, thus increasing the selection intensity and eventually optimizing genetic gain.

This study in sparse testing also highlights how enhancing selection intensity by increasing the evaluation of genotypes eventually leads to improved genetic gain. Although this study did not assess how an increase in selection intensity would maximize genetic gain, it still showed the potential of evaluating more genotypes by allocating resources as a factor in optimizing genetic gain. Overall, this study aligns with this goal by illustrating how G×E components in the GP model behave under different compositions of non-overlapping and overlapping genotypes. By finding compositions that maximize accuracy, breeders can better target superior *Miscanthus* genotypes with well-adaptability, increasing final genetic gains.

5 Conclusions

Genome-based sparse testing design is an approach to accelerate breeding programs cost-effectively. It allows evaluate candidate genotypes in many environments and provides a detailed evaluation regarding their broad adaptability and stability. In the present study, different compositions of training sets and allocation schemes were assessed in *M. sacchariflorus* multi-environment trials. The GP M3 (E +L+G+GE) model that incorporated the G×E interaction term, showed the best PA with low MSE across allocation designs for all the traits. It showed that incorporating G×E in GP model improved the PA and can be utilized in sparse testing.

Overall, M3 provided an opportunity to maintain high accuracy even after reducing the calibration set sizes. This could significantly reduce phenotyping costs in multi-environment *Miscanthus* breeding trials. Also, it can increase the testing capacity of the *Miscanthus* genotypes in METs. From this study, an ideal design strategy for multi-environment *Miscanthus* trials could involve testing a few common genotypes along with more non-overlapping genotypes assigned to all environments.

Data availability statement

The datasets analyzed for this study can be found in the figshare repository at https://doi.org/10.6084/m9.figshare.31796794.

Author contributions

SP: Data curation, Writing – original draft, Methodology, Investigation, Visualization, Validation, Formal analysis. NL: Investigation, Conceptualization, Writing – review & editing, Methodology, Supervision. ES: Writing – review & editing, Resources. ADBL: Funding acquisition, Writing – review & editing, Resources. HZ: Resources, Writing – review & editing. BG: Resources, Writing – review & editing. AEL: Resources, Data curation, Writing –

review & editing. JN: Data curation, Resources, Writing – review & editing. CY: Writing – review & editing, Resources. ESS: Resources, Writing – review & editing. JY: Writing – review & editing, Resources. HN: Writing – review & editing, Resources. KA: Writing – review & editing, Resources. TY: Writing – review & editing, Resources. PC: Resources, Writing – review & editing. XJ: Writing – review & editing, Resources. LC: Data curation, Writing – review & editing. KP: Writing – review & editing, Resources. JP: Resources, Writing – review & editing. ASa: Writing – review & editing, Resources. ED: Resources, Writing – review & editing. ND: Writing – review & editing, Resources. KG: Resources, Writing – review & editing. MN: Writing – review & editing, Supervision. AC: Supervision, Writing – review & editing. MS: Resources, Writing – review & editing. LB: Resources, Writing – review & editing. ASh: Data curation, Writing – review & editing, Supervision. JG-A: Supervision, Writing – review & editing. DJ: Conceptualization, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Writing – Review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This work was funded by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Biological and Environmental Research Program under Award Number DE-SC0018420).

Conflict of interest

Author KP was employed by company Schroll Medical ApS. Author JP was employed by company Spring Valley Agriscience Co. Ltd.

The remaining author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author DJ declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

Author disclaimer

Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2026.1834912/full#supplementary-material

References

Acosta-Pech, R., Crossa, J., de Los Campos, G., Teyssèdre, S., Claustres, B., Pérez-Elizalde, S., et al. (2017). Genomic models with genotype× environment interaction for predicting hybrid performance: an application in maize hybrids. *Theor. Appl. Genet.* 130, 1431–1440. doi: 10.1007/s00122-017-2898-0

Alemu, A., Åstrand, J., Montesinos-Lopez, O. A., y Sanchez, J. I., Fernandez-Gonzalez, J., Tadesse, W., et al. (2024). Genomic selection in plant breeding: Key factors shaping two decades of progress. *Mol. Plant* 17, 552–578. doi: 10.1016/j.molp.2024.03.007

Anzoua, K.G., Yamada, T., and Henry, R.J. (2011). Miscanthus. In: C. Kole (ed.). *Wild Crop Relatives: Genomic and Breeding Resources*. Berlin, Heidelberg: Springer. doi: 10.1007/978-3-642-21102-7_9

Aoun, M., Carter, A., Thompson, Y. A., Ward, B., and Morris, C. F. (2021). Environment characterization and genomic prediction for end-use quality traits in soft white winter wheat. *Plant Genome* 14, e20128. doi: 10.1002/tpg2.20128

Argaw, T., Fenta, B. A., Zegeye, H., Azmach, G., and Funga, A. (2025). Multi-environment trials data analysis: linear mixed model-based approaches using spatial and factor analytic models. *Front. Res. Metrics Anal.* 10, 1472282. doi: 10.3389/frma.2025.1472282

Arnoult, S., and Brancourt-Hulmel, M. (2015). A review on Miscanthus biomass production and composition for bioenergy use: genotypic and environmental variability and implications for breeding. *Bioenergy Res.* 8, 502–526. doi: 10.1007/s12155-014-9524-7

Atanda, S. A., Govindan, V., Singh, R., Robbins, K. R., Crossa, J., and Bentley, A. R. (2022). Sparse testing using genomic prediction improves selection for breeding targets in elite spring wheat. *Theor. Appl. Genet.* 135, 1939–1950. doi: 10.1007/s00122-022-04085-0

Bernardo, R. (1994). Prediction of maize single-cross performance using RFLPs and information from related hybrids. *Crop Sci.* 34, 20–25. doi: 10.2135/cropsci1994.0011183x003400010003x

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308

Brown, D., Van den Bergh, I., de Bruin, S., Machida, L., and van Etten, J. (2020). Data synthesis for crop variety evaluation. A review. *Agron. Sustain. Dev.* 40, 25. doi: 10.1007/s13593-020-00630-7

Budhlakoti, N., Kushwaha, A. K., Rai, A., Chaturvedi, K. K., Kumar, A., Pradhan, A. K., et al. (2022). Genomic selection: A tool for accelerating the efficiency of molecular breeding for development of climate-resilient crops. *Front. Genet.* 13, 832153. doi: 10.3389/fgene.2022.832153

Cheng, J., Dekkers, J. C., and Fernando, R. L. (2021). Cross-validation of best linear unbiased predictions of breeding values using an efficient leave-one-out strategy. *J. Anim Breed. Genet.* 138, 519–527. doi: 10.1111/jbg.12545

Chung, J.-H., and Kim, D.-S. (2012). Miscanthus as a potential bioenergy crop in East Asia. *J. Crop Sci. Biotechnol.* 15, 65-77. doi: 10.1007/s12892-012-0023-0

Clark, L. V., Brummer, J. E., Głowacka, K., Hall, M. C., Heo, K., Peng, J., et al. (2014). A footprint of past climate change on the diversity and population structure of Miscanthus sinensis. *Ann. Bot.* 114, 97–107. doi: 10.1093/aob/mcu084

Clark, L. V., Dwiyanti, M. S., Anzoua, K. G., Brummer, J. E., Ghimire, B. K., Głowacka, K., et al. (2019). Genome-wide association and genomic prediction for biomass yield in a genetically diverse Miscanthus sinensis germplasm panel phenotyped at five locations in Asia and North America. *GCB Bioenergy* 11, 988–1007. doi: 10.1111/gcbb.12620

Clark, S. A., Hickey, J. M., Daetwyler, H. D., and van der Werf, J. H. (2012). The importance of information on relatives for the prediction of genomic breeding values and the implications for the makeup of reference data sets in livestock breeding schemes. *Genet. Sel Evol.* 44, 1–9. doi: 10.1186/1297-9686-44-4

Clifton-Brown, J. C., and Lewandowski, I. (2000). Water use efficiency and biomass partitioning of three different Miscanthus genotypes with limited and unlimited water supply. *Ann. Bot.* 86, 191–200. doi: 10.1006/anbo.2000.1183

Clifton-Brown, J. C., and Lewandowski, I. (2002). Screening Miscanthus genotypes in field trials to optimise biomass yield and quality in Southern Germany. *Eur. J. Agron.* 16, 97–110. doi: 10.1016/s1161-0301(01)00120-4

Clifton-Brown, J. C., Lewandowski, I., Andersson, B., Basch, G., Christian, D. G., Kjeldsen, J. B., et al. (2001). Performance of 15 Miscanthus genotypes at five sites in Europe. *Agron. J.* 93, 1013–1019. doi: 10.17104/9783406783432-63

Clifton-Brown, J., Schwarz, K. U., Awty-Carroll, D., Iurato, A., Meyer, H., Greef, J., et al. (2019). Breeding strategies to improve Miscanthus as a sustainable source of biomass for bioenergy and biorenewable products. *Agronomy* 9, 673. doi: 10.3390/agronomy9110673

Crossa, J., Yang, R. C., and Cornelius, P. L. (2004). Studying crossover genotype× environment interaction using linear-bilinear models and mixed models. *J. Agric. Biol. Environ. Stat.* 9, 362–380. doi: 10.1198/108571104x4423

Dong, H., Clark, L. V., Lipka, A. E., Brummer, J. E., Głowacka, K., Hall, M. C., et al. (2019). Winter hardiness of Miscanthus (III): Genome-wide association and genomic prediction for overwintering ability in Miscanthus sinensis. *GCB Bioenergy* 11, 930–955. doi: 10.1111/gcbb.12587

Garcia-Abadillo, J., Adunola, P., Aguilar, F. S., Trujillo-Montenegro, J. H., Riascos, J. J., Persa, R., et al. (2024). Sparse testing designs for optimizing predictive ability in sugarcane populations. *Front. Plant Sci.* 15, 1400000. doi: 10.3389/fpls.2024.1400000

González-Barrios, P., Díaz-García, L., and Gutiérrez, L. (2019). Mega-environmental design: Using genotype× environment interaction to optimize resources for cultivar testing. *Crop Sci.* 59, 1899–1912. doi: 10.2135/cropsci2018.11.0692

Hodkinson, T. R., and Renvoize, S. (2001). Nomenclature of Miscanthus x giganteus (Poaceae). *Kew Bull.* 56, 759. doi: 10.2307/4117709

Huang, L. S., Flavell, R., Donnison, I. S., Chiang, Y. C., Hastings, A., Hayes, C., et al. (2019). Collecting wild Miscanthus germplasm in Asia for crop improvement and conservation in Europe whilst adhering to the guidelines of the United Nations' Convention on Biological Diversity. *Ann. Bot.* 124, 591–604. doi: 10.1093/aob/mcy231

Isidro, J., Jannink, J. L., Akdemir, D., Poland, J., Heslot, N., and Sorrells, M. E. (2015). Training set optimization under population structure in genomic selection. *Theor. Appl. Genet.* 128, 145–158. doi: 10.1007/s00122-014-2418-4

Jarquín, D., Crossa, J., Lacaze, X., Du Cheyron, P., Daucourt, J., Lorgeou, J., et al. (2014a). A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theor. Appl. Genet.* 127, 595–607. doi: 10.1007/s00122-013-2243-1

Jarquin, D., De Leon, N., Romay, C., Bohn, M., Buckler, E. S., Ciampitti, I., et al. (2021). Utility of climatic information via combining ability models to improve genomic prediction for yield within the genomes to fields maize project. *Front. Genet.* 11, 592769. doi: 10.3389/fgene.2020.592769

Jarquin, D., Howard, R., Crossa, J., Beyene, Y., Gowda, M., Martini, J. W., et al. (2020). Genomic prediction enhanced sparse testing for multi-environment trials. *G3: Genes Genomes Genet.* 10, 2725–2739. doi: 10.1534/g3.120.401349

Jarquín, D., Kocak, K., Posadas, L., Hyma, K., Jedlicka, J., Graef, G., et al. (2014b). Genotyping by sequencing for genomic prediction in a soybean breeding population. *BMC Genomics* 15, 740. doi: 10.1186/1471-2164-15-740

Kaiser, C. M., Clark, L. V., Juvik, J. A., Voigt, T. B., and Sacks, E. J. (2015). Characterizing a Miscanthus germplasm collection for yield, yield components, and genotype× environment interactions. *Crop Sci.* 55, 1978–1994. doi: 10.1109/ner.2017.8008436

Kreig, J. A., Ssegane, H., Chaubey, I., Negri, M. C., and Jager, H. I. (2019). Designing bioenergy landscapes to protect water quality. *Biomass Bioenergy* 128, 105327. doi: 10.1016/j.biombioe.2019.105327

Ladanai, S., and Vinterbäck, J. (2009). *Global potential of sustainable biomass for energy* Vol. 13 (Uppsala: Department of Energy and Technology, Swedish University of Agricultural Sciences).

Lenth, R. V. (2016). Least-squares means: The R package lsmeans. *J. Stat. Software* 69, 1–33. doi: 10.18637/jss.v069.i01

Lewandowski, I., Clifton-Brown, J., Trindade, L. M., Van der Linden, G. C., Schwarz, K. U., Müller-Sämann, K., et al. (2016). Progress on optimizing Miscanthus biomass production for the European bioeconomy: Results of the EU FP7 project OPTIMISC. *Front. Plant Sci.* 7, 1620. doi: 10.3389/fpls.2016.01620

Lewandowski, I., Scurlock, J. M., Lindvall, E., and Christou, M. (2003). The development and current status of perennial rhizomatous grasses as energy crops in the US and Europe. *Biomass Bioenergy* 25, 335–361. doi: 10.1016/s0961-9534(03)00030-8

Lopez-Cruz, M., Crossa, J., Bonnett, D., Dreisigacker, S., Poland, J., Jannink, J. L., et al. (2015). Increased prediction accuracy in wheat breeding trials using a marker× environment interaction genomic selection model. *G3: Genes Genomes Genet.* 5, 569–582. doi: 10.1534/g3.114.016097

McGaugh, S. E., Lorenz, A. J., and Flagel, L. E. (2021). The utility of genomic prediction models in evolutionary genetics. *Proc. R. Soc. B.* 288, 20210693. doi: 10.1098/rspb.2021.0693

Metzger, J. O., and Hüttermann, A. (2009). Sustainable global energy supply based on lignocellulosic biomass from afforestation of degraded areas. *Naturwissenschaften* 96, 279–287. doi: 10.1007/s00114-008-0479-4

Meuwissen, T. H., Hayes, B. J., and Goddard, M. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829. doi: 10.1093/genetics/157.4.1819

Mitros, T., Session, A. M., James, B. T., Wu, G. A., Belaffif, M. B., Clark, L. V., et al. (2020). Genome biology of the paleotetraploid perennial biomass crop Miscanthus. *Nat. Commun.* 11, 5442. doi: 10.1038/s41467-020-18923-6

Njuguna, J. N., Clark, L. V., Lipka, A. E., Anzoua, K. G., Bagmet, L., Chebukin, P., et al. (2023). Genome-wide association and genomic prediction for yield and component traits of Miscanthus sacchariflorus. *GCB Bioenergy* 15, 1355–1372. doi: 10.1111/gcbb.13097

O'Loughlin, J., Finnan, J., and McDonnell, K. (2017). Accelerating early growth in Miscanthus with the application of plastic mulch film. *Biomass Bioenergy* 100, 52–61. doi: 10.1016/j.biombioe.2017.03.003

Olatoye, M. O., Clark, L. V., Labonte, N. R., Dong, H., Dwiyanti, M. S., Anzoua, K. G., et al. (2020). Training population optimization for genomic selection in Miscanthus. *G3: Genes Genomes Genet.* 10, 2465–2476. doi: 10.1534/g3.120.401402

Pérez, P., and de Los Campos, G. (2014). Genome-wide regression and prediction with the BGLR statistical package. *Genetics* 198, 483–495. doi: 10.1534/genetics.114.164442

Persa, R., Canella Vieira, C., Rios, E., Hoyos-Villegas, V., Messina, C. D., Runcie, D., et al. (2023). Improving PA in sparse testing designs in soybean populations. *Front. Genet.* 14, 1269255. doi: 10.3389/fgene.2023.1269255

Persa, R., Grondona, M., and Jarquin, D. (2021). Development of a genomic prediction pipeline for maintaining comparable sample sizes in training and testing sets across prediction schemes accounting for the genotype-by-environment interaction. *Agriculture* 11, 932. doi: 10.3390/agriculture11100932

R Core Team (2023). *R: A language and environment for statistical computing* (Vienna, Austria: R Foundation for Statistical Computing).

Rincent, R., Kuhn, E., Monod, H., Oury, F. X., Rousset, M., Allard, V., et al. (2017). Optimization of multi-environment trials for genomic selection based on crop models. *Theor. Appl. Genet.* 130, 1735–1752. doi: 10.1007/s00122-017-2922-4

Saha, M. C., Bhandhari, H. S., and Bouton, J. H. (2013). *Bioenergy feedstocks: breeding and genetics*. Eds. M. C. Saha, H. S. Bhandhari and J. H. Bouton (Hoboken, NJ: John Wiley & Sons).

Slavov, G. T., Nipper, R., Robson, P., Farrar, K., Allison, G. G., Bosch, M., et al. (2014). Genome-wide association studies and prediction of 17 traits related to phenology, biomass and cell wall composition in the energy grass Miscanthus sinensis. *New Phytol.* 201, 1227–1239. doi: 10.1111/nph.12621

Smith, A., Ganesalingam, A., Lisle, C., Kadkol, G., Hobson, K., and Cullis, B. (2021). Use of contemporary groups in the construction of multi-environment trial datasets for selection in plant breeding programs. *Front. Plant Sci.* 11, 623586. doi: 10.3389/fpls.2020.623586

Sutton, D. J., Veum, K. S., Davis, M., Lord, S., Ransom, C., and Sudduth, K. (2025). Soil health benefits of perennial biofuel crops on claypan soils. *Soil Secur* 19, 100193. doi: 10.1016/j.soisec.2025.100193

Varela, S., Zheng, X., Njuguna, J. N., Sacks, E. J., Allen, D. P., Ruhter, J., et al. (2022). Deep convolutional neural networks exploit high-spatial-and-temporal-resolution aerial imagery to phenotype key traits in Miscanthus. *Remote Sens* 14, 5333. doi: 10.3390/rs14215333

Varela, S., Zheng, X., Njuguna, J., Sacks, E., Allen, D., Ruhter, J., et al. (2025). Breaking the barrier of human-annotated training data for machine learning-aided plant research using aerial imagery. *Plant Physiol.* 197, kiaf132. doi: 10.1093/plphys/kiaf132

Verbyla, A. (2023). On two-stage analysis of multi-environment trials. *Euphytica* 219, 121. doi: 10.1007/s10681-023-03248-4

Xue, S., Kalinina, O., and Lewandowski, I. (2015). Present and future options for the improvement of Miscanthus propagation techniques. *Renewable Sustain. Energy Rev.* 49, 1233–1246. doi: 10.1016/j.rser.2015.04.168