Improving genomic predictions with inbreeding and nonadditive effects in two admixed maize hybrid populations in single and multienvironment contexts

Morgane Roth, Aurélien Beugnot, Tristan Mary-Huard, Laurence Moreau, Alain Charcosset, Julie Fiévet

Improving genomic predictions with inbreeding and non-additive effects in two admixed maize hybrid populations in single and multi-environment contexts

Morgane Roth[*,1], Aurélien Beugnot[†], Tristan Mary-Huard[†,‡], Laurence Moreau[†], Alain Charcosset[†], Julie B. Fievet[†]

* Plant Breeding Research Division, Agroscope, 8820 Wädenswil, Zurich, Switzerland

[†] Université Paris-Saclay, INRAE, CNRS, AgroParisTech, GQE-Le Moulon, 91190, Gif-sur-Yvette, France.

[‡] Université Paris-Saclay, INRAE, AgroParisTech, UMR MIA-Paris Paris, 75005, Paris, France

[1] Present address: INRAE, GAFL, F-84143, Montfavet, France

Running title: Genomic prediction for admixed hybrids

Keywords: Admixture, inbreeding, non-additive effects, genomic selection, genotype by environment interactions

Corresponding author: Morgane Roth, INRAE GAFL, 67 Allée des Chênes 84140 Montfavet, France, +33 (0) 4 32 72 26 79, morgane.roth@inrae.fr

ABSTRACT

Genetic admixture, resulting from the recombination between structural groups, is frequently encountered in breeding populations. In hybrid breeding, crossing admixed lines can generate substantial non-additive genetic variance and contrasted levels of inbreeding which can impact trait variation. This study aimed at testing recent methodological developments for the modelling of inbreeding and non-additive effects in order to increase prediction accuracy in admixed populations. Using two maize (*Zea mays*) populations of hybrids admixed between dent and flint heterotic groups, we compared a suite of five genomic prediction models incorporating (or not) parameters accounting for inbreeding and non-additive effects with the natural and orthogonal interaction approach (NOIA) in single and multi-environment contexts. In both populations, variance decompositions showed the strong impact of inbreeding on plant yield, height and flowering time which was supported by the superiority of prediction models incorporating this effect (+0.038 in predictive ability for mean yield). In most cases dominance variance was reduced when inbreeding was accounted for. The model including additivity, dominance, epistasis and inbreeding effects appeared to be the most robust for prediction across traits and populations (+0.054 in predictive ability for mean yield). In a multi-environment context, we found that the inclusion of non-additive and inbreeding effects was advantageous when predicting hybrids not yet observed in any environment. Overall, comparing variance decompositions was helpful to guide model selection for genomic prediction. Finally, we recommend the use of models including inbreeding and non-additive parameters following the NOIA approach to increase prediction accuracy in admixed populations.

INTRODUCTION

Hybrid breeding exploits fundamental principles of evolutionary biology for improving crop and animal performance. When performing intraspecific crosses between two natural populations, the fitness of F1 hybrids is expected to depend on the genetic distance between parents. When this distance is optimum, fitness is enhanced and progenies can outperform their parents, a phenomenon called hybrid vigor or heterosis (Shull 1914; Smith *et al.* 1990;

Vasseur *et al.* 2019). On the edges of this zone, suboptimal fitness originates from inbreeding depression when the distance is too low or outbreeding depression when the distance is too high (Waser and Price 1994). In animal and plant breeding, outbreeding depression is rarely a concern and a focus is made on the minimization of (negative) inbreeding effects and on the maximization of (positive) heterotic effects (Lin *et al.* 2017). Heterosis has thus been exploited for decades in allogamous crops such as maize, but it is also increasingly considered in autogamous plants such as wheat where more and more hybrid breeding programs are being developed (Labroo *et al.* 2021).

Maize has been the pioneer crop for the invention of hybrid breeding and is a model species for the study of heterosis. From the 1950's heterotic groups have been designed to maximize the performance of inter-group crosses. In breeding programs, the value of a hybrid between a pair of inbred lines is typically assessed through the general combining ability (GCA) of the lines, corresponding to the average value of their hybrid progeny and through their specific combining ability (SCA), corresponding to the interaction between the two parental lines (Sprague and Tatum 1942). It has been shown that the ratio between the variation of SCA and GCA decreases as dominant alleles at QTLs tend to fixation within one of the two heterotic groups. So structuring breeding populations into heterotic groups is a practical way to reduce the contribution of SCA to genetic variance (Reif *et al.* 2007; Larièpe *et al.* 2017). This increases correlatively the relative contribution of additive genetic variance (statistical effects captured via the GCA), which is beneficial for breeding efficiency. However, the near-complete reproductive isolation between heterotic groups leads to genetic erosion in the maize breeding population due to drift and selection which can be problematic for long term diversity management (Gerke *et al.* 2015; Allier *et al.* 2019). In this context, admixing plant material with different origins (different heterotic groups, different breeding programs, introgression of exotic material) can be of interest to retain or increase allelic diversity within groups while creating novel allelic combinations (Rincent *et al.* 2014; Rio *et al.* 2020). From a statistical point of view, this can result in a larger contribution of non-additive variance (statistical effects captured via the SCA) to trait variation of hybrids and to a degree of inbreeding depression.

From a biological perspective, dominance, overdominance, pseudo-overdominance, and epistasis are the driving mechanisms underlying heterosis (Shull 1908; Jones 1917; Garcia *et al.* 2008; Fiévet *et al.* 2010; Waller 2021). It is acknowledged that estimating non-additive effects at the statistical level is not sufficient to infer biological (functional) non-additive effects because statistical estimates are population-dependent (Schön *et al.* 2010, Varona *et al.* 2018a) and non-additive biological effects are partly captured by statistical additivity. Yet, statistical estimates of non-additivity in a given population can allow for a more accurate prediction of total genetic variance which is useful for genomic prediction. Genetic variance is usually measured by calculating the genetic covariance between individuals (proportional to their kinship or pedigree relationship) which can be assessed using variation across genome-wide genetic markers in a more precise way than when using pedigree information (Bernardo 1993; Muñoz *et al.* 2014; Legarra 2016). To assess non-additive variance, the total genetic variance needs to be dissected between statistical additive, dominant and epistatic terms. However classical models accounting for additive, dominant and epistatic effects do not efficiently handle the confounding beween these genetic effects. This generally results in an overestimation of the total genetic variance (Vitezica *et al.* 2013; Muñoz *et al.* 2014; Varona *et al.* 2018a), in incorrectly interpretating the contribution of non-additive effects to genetic variance and in ignoring their usefulness for genomic prediction (Vitezica *et al.* 2017). Variance decomposition and prediction models need to be improved to alleviate these issues. Álvarez-Castro and Carlborg (2007) developed the natural and orthogonal interaction (NOIA) framework to formally navigate between the functional and statistical modelling of non-additive effects. Their statistical model declares dominance and additive effects orthogonally and can be applied to populations deviating from Hardy-Weinberg equilibrium (HWE). More recently, Vitezica and collaborators (2017, 2018) further adapted this framework for the estimation of additive, dominance and epistatic covariances between individuals in a genomic (multilocus) context for genomic prediction.

Dominance affects heterosis if the sum of dominance effects is different from zero, a phenomenon called directional dominance (Frankel 1983; Lynch and Walsh 1998). In maize

breeding, it has been long recognized that the correlation between phenotypic performance and genetic distance between parents (or heterozygosity = 1-homozygosity) holds in the presence of directional dominance only for intra-heterotic hybrids but not for inter-heterotic hybrids (Charcosset *et al.* 1991; Bernardo 1992; Charcosset and Essioux 1994). We hypothesize that this correlation is restored in populations recombining heterotic groups. Indeed, the inbreeding effect of a genomic segment inherited from the same group by the two parents of one hybrid is expected to be predictable with markers thanks to the linkage disequilibrium (LD) within this group. Also, new LD patterns are created between linked alleles that have contrasted frequencies across groups. Besides, in presence of directional dominance, the assumption that dominance effects are centered around zero does not hold. It was shown in pigs that statistical dominance variance is overestimated if inbreeding is not accounted for in variance decompositions and that using the inbreeding level (as measured by the proportion of homozygous marker loci of each individual) as a fixed covariate when estimating variance components can efficiently control for this artefact (Xiang *et al.* 2016; Vitezica *et al.* 2018; Varona *et al.* 2018b). Beyond variance component estimation, marker-based models also aim at performing genomic predictions of the value of new individuals based on their marker-based relatedness with individuals already phenotyped. In this context, incorparating both non-additive and inbreeding parameters in genomic prediction models thus appears to be an appropriate strategy when dealing with hybrids between admixed individuals Xiang *et al.* 2016; Vitezica *et al.* 2018; Varona *et al.* 2018b).

Predicting individuals in different environments is also a major stake in plant breeding. The ranking of individuals changes from one environment to another in the presence of genotype by environment interactions (G×E), which is very common in complex traits. Several methods have been proposed to take into account G×E in predictions across environments, for example modelling marker × environment interactions explicitly (Lopez-Cruz *et al.* 2015), using the Hadamard product between genetic and environmental covariance matrices (Jarquín *et al.* 2014), calculating genotypic sensitivities to environmental gradients (Millet *et al.* 2019) or using crop growth models for trait-assisted prediction (Robert *et al.* 2020). These methods

clearly outperform predictions based on main effects (G + E only), however little is known about the partition of G×E effects among additive and non-additive components and their potential use in genomic prediction in a hybrid context (Kadam *et al.* 2016; Acosta-Pech *et al.* 2017).

Here, we worked with two maize admixed populations with the objectives of i) studying the impact of inbreeding and the relative importance of statistical non-additive effects on trait variation and ii) finding the best genomic prediction models in single and multi-environment contexts. These two populations of hybrids are derived from admixed inbred lines between dent and flint heterotic groups and were chosen for their contrasted level of genetic diversity and linkage disequilibrium : the parental lines from the iF2 hybrid population ("immortalized F2", Hua *et al.* 2002) were obtained from an original cross between two lines only (one dent, one flint) while the parental lines from the Het2 population are admixed between a total of 604 lines (300 dent and 304 flint, Rio *et al.* 2020). To our knowledge, only one study assessed genomic prediction accuracy among admixed maize hybrids (Guo *et al.* 2013, based on an iF2 hybrid population). Here, four traits known for their agronomical relevance and contrasted genetic architecture were measured in the iF2 and Het2 populations. We assessed the effect of inbreeding and quantified non-additive variance in these populations and tested the corresponding models with genomic predictions in single and multiple environments.

MATERIAL AND METHODS

Plant material

Two different populations were considered in this study. The heterosis#2 hybrid population (Het2) consists in 291 single-cross hybrids obtained from an incomplete diallel design between 321 double haploid (DH) lines. These DH lines derived from F1 hybrids obtained by crossing dent lines to flint lines issued from two diversity panels of 300 and 304 individuals respectively (see Rio *et al.* 2020 for more details). As a consequence, DH lines were admixed lines carrying some chromosome segments with a dent origin and others with a flint origin (due to recombination events during the meiosis of F1s). To produce Het2 hybrids, admixed DH lines were chosen and crossed to limit the number of hybrids per line. In this study, a DH line can be the parent of four hybrids at the maximum. Due to seed production failures, some

admixed lines are involved in only one hybrid. The number of contributions per DH line in the Het2 hybrid panel ranges from 1 to 4. The Het2 panel was phenotyped in four different locations in France (Jargeau, Loiret; Aubiat, Puy de Dôme; Souprosse, Landes; Saint-Martin de Hinx, Landes) in 2016 and 2017 resulting in five environments (further referred to as jar16, aub17, sou17, smh16 and smh17). Each Het2 environment was composed of elementary two-row plots of 9.28 m². In addition to the 291 hybrids, repeated on average 1.3 times per environment in a partially-replicated design (p-rep), we dedicated 12 plots to controls (3 replicates for each of the 2 commercial hybrids Milesim and DKC4841 and each of the 2 additional controls B73 × UH007 and PH207 × UH007 hybrids).

The immortalized F2 population (iF2) consists in 265 hybrids obtained from an incomplete diallel design among 184 highly recombinant inbred lines of the so-called 'LHRF' population derived from an initial F1 cross between F252, an early dent line and F2, a European flint line (Falque *et al.* 2005; Ganal *et al.* 2011). The LHRF population has experienced four generations of intermating (*i.e.* four meiosis events) and is thus more admixed than the Het2 population. The iF2 population was evaluated in two environmentally contrasted locations in France, Saint-Martin de Hinx (smh, above-mentionned) and Le Moulon (mln, Essonne) in 2010 and 2011 resulting in three environments (smh10, smh11 and mln11). Each environment was made of elementary two-row plots of 9.28 m². In addition to the 265 hybrids repeated on average 1.28 times per environment in a partially-replicated design, we dedicated 56 plots to controls (28 replicates for the F2 line and 28 replicates for the F1 hybrid F2 × F252).

For both populations, planting density was settled according to the usual practice of each location (ranging from 70 to 95,000 plants per hectare). Hybrids were randomized within the environments.

Plant phenotyping and field data correction

Plant height (HT; from the soil to the tip of the tassel, in cm), grain moisture at harvest (GM; in %), grain yield at 15% moisture (GY; in q/ha) and flowering date (FLO) were measured. For FLO, male or female flowering date were recorded, corresponding to the day at which 50% of the plants exhibited mature tassels or silks, respectively. Male flowering was measured in Het2 (calendar days) and female flowering was measured in iF2 population

(expressed in growing degree days in degree days considering 6°C as the base temperature). In each environment, elementary plots with a final number of plants lower than the median density minus 15 plants were excluded from the dataset.

In both populations, phenotypic data were adjusted within each environment to correct for spatial environmental effects using the following model:

$$\begin{aligned} Y_{thrc} &= \mu + \alpha_t + G_h + R_r + C_c + E_{thrc} \quad M1 \\ G &\sim N(0, \sigma_g^2 I) \\ R &\sim N(0, \sigma_r^2 I) \\ C &\sim N(0, \sigma_c^2 I) \\ E &\sim N(0, \sigma_e^2 I) \\ G &\perp R \perp C \perp E\,, \end{aligned}$$

where Y_{thrc} is the phenotype of hybrid *h* measured at row *r* and column *c*, α_t is the effect of control *t* (a 5 level factor: 4 for the different control hybrids and one for non-control hybrids), G_h is the effect of hybrid *h* (one level for each experimental hybrid and one for each hybrid control), R_r the effect of row *r*, C_c the effect of column *c* and E_{thrc} the error. Except for control effects, all effects are assumed to be random. The sign ⊥ indicates independence between random effects.

Broad sense heritability in each environment was calculated according to the following formula,

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_e^2 / n_{rep}}$$

where σ_g^2 and σ_e^2 are respectively the genetic and error variance derived from the above model *M1*, and n_{rep} is the mean number of experimental hybrid repetitions per environment.

The best linear unbiased predictors (BLUPs) of the row and column effects were subtracted from the raw phenotypic values to obtain corrected field performances. Hereafter we note $\widetilde{Y_{her}}$ the corrected field performance of hybrid *h* in environment *e* and plot *r*, and $\overline{Y_{he.}}$ the average of $\widetilde{Y_{her}}$ values within environment e. Lastly, these corrected field plot performances

were used to calculate across-environment least square means $\overline{Y_{h..}}$ for each hybrid *h* with the following model: $\widetilde{Y_{her}} = \mu + \beta_h + \gamma_e + \varepsilon_r$

where β_h, γ_e and ε_r are the fixed effects of hybrid *h*, environment *e* and random effect plot *r* respectively, and $\overline{Y_{h..}} = \hat{\mu} + \widehat{\beta_h}$ with $\hat{\mu}$, the estimated intercept and $\widehat{\beta_h}$, the estimated effect of hybrid h.

Plant genotyping

For Het2 population, parental DH lines were genotyped using a private 15 K SNP-array provided by Limagrain (Chappes, France), including a subset of the 50K Illumina Maize SNP50 BeadChip array (Ganal *et al.* 2011). Genotyping data were expanded up to 600K SNP by imputation, using 600K SNP genotyping of the founder lines and pedigree information (see Rio *et al.* 2020 for more details). Het2 hybrid genotypes were reconstructed from their respective parental genotypic data. Monomorphic markers were eliminated, and remaining markers were filtered using a 5% threshold on minor allele frequency (MAF) based on frequencies in the hybrid population, resulting in a set of 462,247 markers.

For iF2 population, the 184 highly recombinant inbred lines were genotyped with the 50k SNP array (Ganal *et al.* 2011). Residual heterozygous data were treated as missing and all missing values were imputed using Beagle v.3.3.2 and default parameters (Browning and Browning 2007). The iF2 hybrid genotypes were reconstructed from their respective parental data. Monomorphic markers were eliminated, and remaining markers were filtered using a 5% threshold on MAF based on frequencies in the hybrid population, resulting in a set of 16,562 markers.

Variance decomposition

Variance decomposition was performed at three levels: (i) on across-environment least square means $\overline{Y_{h..}}$, (ii) separately on each environment e using corrected field plot performances $\widetilde{Y_{her}}$ and (iii) on multiple environments jointly using corrected field plot performances $\widetilde{Y_{her}}$.

The following baseline model was used to estimate variance components of across-environment least square means:

$$\begin{aligned} \overline{Y_{h..}} &= \mu + \beta Inb_h + A_h + D_h + AA_h + AD_h + DD_h + \varepsilon_h \qquad M2 \\ A &\sim N(0, \sigma_A^2 K_A),\ D \sim N(0, \sigma_D^2 K_D),\ \varepsilon \sim N(0, \sigma_\varepsilon^2 I) \\ AA &\sim N(0, \sigma_{AA}^2 K_{AA}),\ AD \sim N(0, \sigma_{AD}^2 K_{AD}),\ DD \sim N(0, \sigma_{DD}^2 K_{DD}) \\ A &\perp D \perp AA \perp AD \perp DD \perp \varepsilon \end{aligned}$$

where Inb_h is a quantitative variable representing the percentage of homozygosity (number of homozygote loci divided by total number of loci) of hybrid h and β is its associated regression coefficient. Including Inb_h in the model allows one to account for potential directional dominance effects in hybrid h and β is its associated regression coefficient, A_h and D_h are the additive and dominant random effects associated to hybrid h, AA_h, AD_h and DD_h are three random effects modeling epistasis, accounting for additive by additive, additive by dominant and dominant by dominant effects, respectively. ε_h is the random error. σ_A^2 ... σ_{DD}^2 are the variance terms associated to genetic effects, K_A ... K_{DD} are the kinship matrices corresponding to each genetic effect (see below for their estimation).

The following baseline model was used to estimate variance components in each environment separately using corrected field plot performance, which is declared for environment *e* as:

$$\begin{aligned} \widetilde{Y_{her}} &= \mu + \beta Inb_h + A_h + D_h + AA_h + AD_h + DD_h + \varepsilon_{hr} \qquad M3 \\ A &\sim N(0, \sigma_A^2 K_A),\ D \sim N(0, \sigma_D^2 K_D),\ \varepsilon \sim N(0, \sigma_\varepsilon^2 I) \\ AA &\sim N(0, \sigma_{AA}^2 K_{AA}),\ AD \sim N(0, \sigma_{AD}^2 K_{AD}),\ DD \sim N(0, \sigma_{DD}^2 K_{DD}) \\ A &\perp D \perp AA \perp AD \perp DD \perp \varepsilon \end{aligned}$$

using the same notations as in Model *M2*.

Different submodels of models *M2* and *M3* were considered by removing one or several effects. In the following we note ADI_Inb the full model including the additive effect (*A*), the dominance effect (*D*), the epistatic effects (I, corresponding to *AA*, *AD* and *DD*) and the inbreeding effect (*Inb*). Similarly, we note AD_Inb, the model containing no epistatic effects, A_Inb the model containing only the additive and inbreeding effects, AD the model containing only the additive and dominance effects and A the model containing only the additive effect.

The following baseline model was used to perform a variance component analysis jointly on all environments using corrected field plot performance:

$$
\begin{aligned}
\widetilde{Y_{her}} &= \mu + \alpha_e + \beta Inb_h + A_h + D_h + AA_h + AD_h + DD_h + \\
&\quad AE_{he} + DE_{he} + AAE_{he} + ADE_{he} + DDE_{he} + \varepsilon_{her} \qquad M4 \\
A &\sim N(0, \sigma_A^2 K_A),\ D \sim N(0, \sigma_D^2 K_D) \\
AA &\sim N(0, \sigma_{AA}^2 K_{AA}),\ AD \sim N(0, \sigma_{AD}^2 K_{AD}),\ DD \sim N(0, \sigma_{DD}^2 K_{DD}) \\
AE_e &\sim N(0, \sigma_{AE(e)}^2 K_A)\ IND,\ DE_e \sim N(0, \sigma_{DE(e)}^2 K_D)\ IND,\ e = 1 \dots n_{env} \\
AAE_e &\sim N(0, \sigma_{AAE(e)}^2 K_{AA})\ IND,\ ADE_e \sim N(0, \sigma_{ADE(e)}^2 K_{AD})\ IND, \\
&\quad DDE_e \sim N(0, \sigma_{DDE(e)}^2 K_{DD})\ IND,\ e = 1 \dots n_{env} \\
\varepsilon_e &\sim N(0, \sigma_{\varepsilon(e)}^2 I)\ IND, e = 1 \dots n_{env} \\
AE &\perp DE \perp AAE \perp ADE \perp DDE \perp \varepsilon
\end{aligned}
$$

where IND stands for the independence between the different vectors. Here AE_e corresponds to the vector of random effects (AE_{1e},…, AE_{ne}) that accounts for the additivity × environment interactions. These interaction terms are specific to environment *e* and are assumed to be normally distributed with a specific variance $\sigma_{AE(e)}^2$. Similar notations are used for *DEe*, AAE_e, ADE_e and DDE_e that correspond to interaction terms between environment and dominance and epistasis effects respectively. ε_e corresponds to the vector of error terms associated to environment e and is assumed to be normally distributed with an environment-specific variance ($\sigma_{\varepsilon(e)}^2$).

We note model *M4* as the ADI×E_{spec}_Inb model, from which we derive the four models AD×E_{spec} _Inb (no epistasis), AD×E_{spec} (no epistasis and no inbreeding), A×E_{spec}_Inb (only additive and inbreeding), A×E_{spec} (only additive effect). Together these models are referred to as G×E_{spec} models. We also considered five models where variance terms for G×E interactions were common to all environments, noted as A×E_{com}, A×E_{com}_Inb, AD×E_{com}, AD×E_{com}_Inb, ADI×E_{com}_Inb together referred to as G×E_{com} models. Thus, we derived nine models from *M4* for a total of ten G×E models. Models *M2*, *M3* and *M4* were also used to perform genomic predictions.

261 *Kinship matrices*

262 Kinship coefficients were estimated for additive (K_A), dominant (K_D), additive × additive
263 (K_{AA}), additive × dominant (K_{AD}) and dominant × dominant (K_{DD}) effects according to the
264 Natural and Orthogonal Interaction Approach (NOIA, Álvarez-Castro and Carlborg 2007) as
265 expanded by Vitezica *et al.* (2017). Briefly, the additive and dominant coefficients were
266 calculated for a given hybrid *i* using genotypic frequencies as follows:

$$h_{A_{i,j}} = \begin{cases} -(-p_{Bb} - 2p_{bb}) \\ -(1 - p_{Bb} - 2p_{bb}) \\ -(2 - p_{Bb} - 2p_{bb}) \end{cases} for\ genotypes \begin{cases} BB \\ Bb \\ bb \end{cases}$$

$$h_{D_{i,j}} = \begin{cases} \dfrac{-2p_{Bb}p_{bb}}{p_{BB} + p_{bb} - (p_{BB} - p_{bb})^2} \\ \dfrac{4p_{BB}p_{bb}}{p_{BB} + p_{bb} - (p_{BB} - p_{bb})^2} \\ \dfrac{-2p_{BB}p_{Bb}}{p_{BB} + p_{bb} - (p_{BB} - p_{bb})^2} \end{cases} for\ genotypes \begin{cases} BB \\ Bb \\ bb \end{cases}$$

267

268 with p_{BB}, p_{Bb}, p_{bb} being the genotypic frequencies of BB, Bb and bb at locus *j*. These
269 coefficients were combined for *n* individuals and *m* markers in matrices H_A and H_D as follows:

$$H_A = \begin{pmatrix} h_{A_{1,1}} \cdots h_{A_{1,m}} \\ \vdots \cdots \vdots \\ h_{A_{n,1}} \cdots h_{A_{n,m}} \end{pmatrix} and\ H_D = \begin{pmatrix} h_{D_{1,1}} \cdots h_{D_{1,m}} \\ \vdots \cdots \vdots \\ h_{D_{n,1}} \cdots h_{D_{n,m}} \end{pmatrix}$$

270 Kinship matrices were obtained with the following formula:

$$K_A = \frac{H_A H_A'}{tr(H_A H_A')/n}$$

$$K_D = \frac{H_D H_D'}{tr(H_D H_D')/n}$$

$$K_{AA} = \frac{K_A \odot K_A}{tr(K_A \odot K_A)/n}$$

$$K_{AD} = \frac{K_A \odot K_D}{tr(K_A \odot K_D)/n}$$

$$K_{DD} = \frac{K_D \odot K_D}{tr(K_D \odot K_D)/n}$$

where *n* is the number of individuals and $\odot$ is the Hadamard product.

Model parameters were estimated via restricted maximum likelihood inference, using the MM4LMM R package (Laporte and Mary-Huard 2019). Models were compared using the BIC criterion. Fixed parameters were tested for significance using a Wald test procedure. Pairwise comparisons between the suites of models ADI×E_{spec}_Inb and ADI×E_{com}_Inb for each trait were made using a likelihood ratio test (LRT) with the same procedure as described in Yadav *et al.* 2021.

Estimation of predictive ability via cross-validations

In order to assess whether including inbreeding and non-additive effects could improve genomic prediction models, we calculated predictive ability using different validation strategies (Figure 1). The "Global" scenario was used to assess the predictive ability when predicting global means across environment, corresponding to the $\overline{Y_{h..}}$ values. In the "Within-environment" scenario, the goal was to calculate predictive ability within each environment and then compare predictive ability in the different environments. The corrected field plot performances $\widetilde{Y_{her}}$ were used for calibrating the model in the training set and $\overline{Y_{he.}}$ values were used for the validation set. In scenarios "Global" and "Within-environment", predictive ability was evaluated using a five-fold cross-validation design: the population was split into five equal parts, the training set (TS) consisting in 4/5 and the validation set (VS) of the remaining 1/5. In an iterative process, each part served as validation set for a given cross-

validation and predictive ability was calculated as the mean of the five correlation values (obtained for each part). Cross-validations were repeated 100 times.

Lastly, we considered two validation scenarios of across environment prediction (Figure 1) using *M4* and submodels. "G×E_new_env" aimed at evaluating the predictive ability for individuals which have been observed in all environments except one. In that scenario, only hybrids used in all environments were considered (this set differs depending on trait and population). The training set was constituted by removing for each hybrid of the validation set all the observations in one randomly selected environment. Removed values were used as the validation set. This means that the validation set contained a number of field observations equal to $\frac{n_{obs}}{n_{env}}$ and the training set consisted in $\frac{n_{obs} \times (n_{env}-1)}{n_{env}}$ field observations, with n_{obs} the total number of observations and n_{env} the total number of environments. "G×E_new_hyb" aimed at evaluating the predictive ability for hybrids which have not been observed in the TS: in this scenario, the TS was constituted by removing the observations of 1/5 of the hybrids from all environments. In these last two scenarios, $\widetilde{Y_{her}}$ values were used for assembling the TS, and averaged values over replicates of $\widetilde{Y_{her}}$ (per hybrid and environment) were used for validation.

All fixed effect and variance parameters were estimated within the TS and used to predict the genetic values in the VS Using models *M2, M3* and *M4* (and their submodels). Predictive ability was defined as the Pearson correlation between observed and predicted field performances in the VS.

RESULTS

Preliminary analyses of phenotypic data

Globally, Het2 individuals showed higher GY and higher HT than iF2 individuals (mean GY of 76.7 and 42.7q/ha, mean HT of 210.9 and 150.6 cm across environments, respectively Table 1, Table S1), indicating a better agronomical performance.

Heritability was high overall, ranging between 0.61 and 0.97 for the Het2 population and between 0.69 and 0.93 for the iF2 population and strongly depended on trait and environment

(Table 1). Lowest heritabilities for GY were found in smh16 and aub17 for the Het2 population (0.61 and 0.67) and in environment mln11 for the iF2 population (0.69, Table 1). Correlations between environments were highest for FLO (0.73 to 0.86 for Het2 and 0.63 to 0.78 for iF2) and lowest for GY (0.29 to 0.62 for Het2 and 0.31 to 0.59 for iF2, Figure S1).

Variance decomposition

Considering only polymorphic loci, mean homozygosity was 0.64 in Het2 and 0.51 in iF2 population, with a narrower distribution in Het2 when compared to iF2 population (respective standard deviations 0.017 and 0.061 Figure S2). We also observed a wider distribution of kinship values in iF2 population when compared to Het2, especially for additive and dominance kinship. In both populations, off diagonal epistasis kinship values had a narrow distribution around zero (Figure S2).

Variance decompositions on adjusted means across environments ($\overline{Y_{h..}}$ values) were obtained with the five models derived from *M2* (A, A_Inb, AD, AD_Inb and ADI_Inb) for each trait and population (Table 2). BIC values indicated that model A_Inb was the best for FLO and GY in both populations, and for HT in iF2 population. In those cases, the effect of inbreeding was significant whatever the model (Wald test, Table 2). The simplest model A had the lowest BIC value for GM in both populations and the effect of inbreeding was not significant for this trait (Table 2). Inbreeding had overall a negative effect on GY and HT and a positive effect on FLO. As an example, an increase of 0.1 of the inbreeding coefficient is expected to reduce global mean GY by 11,7q/ha in Het2 and by 3,3 to 3,5q/ha in iF2 population. Using model AD, the dominance variance represented up to a third of the additive variance, depending on the trait, the highest proportions being found for GY (Table 2). In contrast, dominance variance components were strongly reduced when adding the inbreeding parameter into the model (in some cases no dominance variance was detected, model AD_Inb Table 2). Among other non-additive effects, high proportions of genetic variance were assigned to epistatic AA effects for all traits in iF2, and for HT in Het2 population (using model ADI_Inb, Table 2). AD and DD effects remained weak across all traits in iF2 population, and were more pronounced in Het2 for FLO, GM and GY. Overall, the lowest

error variances were obtained with the model ADI_Inb, except for GY in Het2 population where the lowest error variance was found with model AD (Table 2).

Variance decompositions within each environment (*M3* and respective submodels) using corrected field performances $\widetilde{Y_{her}}$ confirmed the strong impact of inbreeding on phenotypic variation and indicated that variance components strongly vary from one environment to another (Table S2, Figure 2). A comparison of models for these single-environment variance decompositions using the BIC criterion suggests that A and AD were the best model choices in Het2 population while iF2 phenotypic data were best fitted with models A and A_Inb (Table S2**)**. The inbreeding effect was significant for GY in iF2 population using the three corresponding models (A_Inb, AD_Inb and ADI_Inb). It was consistently non-significant for GM in both populations. For other cases, the inbreeding effect was significant only for certain models or certain environments (Table S2). As observed with $\overline{Y_{h..}}$ values, the effect of inbreeding was negative on GY and HT and positive on FLO. Considering only significant inbreeding regression coefficients in both populations, an increase of inbreeding corresponded to a loss of yield (loss of 9.9 to 15.3q/ha in Het2 and of 2.1 to 4.4 q/ha in iF2 for an increase of 0.1 in inbreeding), to a reduction in height (reduction of 14.7 cm in Het2 and 3.0 to 3.3 cm in iF2 for an increase of 0.1 in inbreeding) to a later flowering (delays of 1.5 to 2 days in Het2 and 5.2 to 8.9 GDD in iF2 for an increase of 0.1 in inbreeding, Table S2). Further, we found that the relative proportions of genetic effects varied strongly between environments. For example when considering FLO in Het2 population, the genetic variance was purely additive for environment aub17, while strong epistatic effects were detected in other environments (Figure 2A, Table S2). Large dominant effects were observed for most traits in Het2 population using models AD and AD_Inb. Dominance effects were less pronounced in iF2, except for GY (Figure 2B). Compared to models AD and AD_Inb, dominance was mostly substituted by epistatic effects in the model ADI_Inb (both populations, Figure 2, Table S2). A tendency for a slight decrease in dominance effects from model AD to AD_Inb was also observed. AA effects were more pronounced in iF2 population, where for example the additive by additive variance term was larger than the additive one considering GY in environment smh11 (Figure 2B).

Using M4 models and submodels, variance decomposition across environments allowed the quantification of G×E variances considering either environment-specific variances (models G×E_{spec}) or variances common to all environments (models G×E_{com}, Table S3, Figure 3**,** Figure S3). We found a significant effect of inbreeding with all corresponding models for GY and FLO in both population and for HT in iF2 population only. Inbreeding had no significant effect on GM in both populations using these models (Table S3). In both populations the lowest BIC values were obtained with the incorporation of the fixed effect of inbreeding in all traits but GM (both populations) and HT (Het2 population,Table S3).

Although G×E interactions (A×E, D×E, AA×E, AD×E, DD×E variance terms, represented by lighter colors on Figure 3**,** Figure S3) were always weaker than main genetic effects (A, D, AA, AD and DD variance terms, represented by darker colors on Figure 3**,** Figure S3), substantial G×E variances were found using *M4* (and submodels). The highest G×E effects were observed for GY and GM in both populations (maximum ratio of G×E variance over G variance of 0.89 and 0.91 respectively, Table S3). Of note, ADI_Inb detected larger G×E effects than other models for a given trait. Variance terms of genetic effects (G effects) were of similar magnitude in G×E_{com} and G×E_{spec} but G×E effects tended to be higher with G×E_{spec} models which contributed a reduction of the error variances. This was particularly visible considering GM in both populations (Figure S3). As observed with *M2* and *M3* models, the additive variance was by far the largest genetic effect identified across traits and populations and the dominance contribution was slightly reduced in the models AD_Inb (AD×E_{com}_Inb and AD×E_{spe}_Inb) when compared to the models AD (AD×E_{com} and AD×E_{spe}). This decrease was more evident for FLO in Het2 and GY in iF2, where dominance effects were strongest (Table S3, Figure 3B, Figure S3). Among epistatic effects, AA was strongest for all traits in iF2 and for HT only in Het2. DD effects were detected only for FLO in iF2 population. AD effects were negligible across all traits and populations (Table S3, Figure S3).

We observed different trends in the composition of G×E interactions terms between models G×E_{com} and G×E_{spec}. Overall, larger D×E variances were obtained using models G×E_{com} (light

orange color Figure 3**,** Figure S3**)** and larger A×E variances were obtained using the models $G\times E_{spec}$ (light red color Figure 3**,** Figure S3**,** Table S3).

When focusing on models $G\times E_{com}$, larger proportions of D×E variance were found in Het2 than in iF2 population using models $AD\times E_{com}$ and $AD\times E_{com}$_Inb (Figure S3). However, when assessed with the model $ADI\times E_{com}$_Inb, D×E variance terms were much smaller and mostly substituted by epistatic by environment interactions (Epi×E effects). For both populations these Epi×E effects consisted mainly in AD×E effects, although strong DD×E effects were found for GM. The strongest cumulated epistatic interactions, as assessed with the model $ADI\times E_{com}$_Inb, were found for GY in iF2 population (Figure 3B, Figure S3).

When focusing on models $G\times E_{spec}$, G×E variance terms were very contrasted across environments (Table S3). For example, considering GY in Het2 population and using model $A\times E_{spec}$, A×E variance ranged from 8.3 to 144.9 across the five environments **(**Table S3). We also found substantial D×E effects with models $AD\times E_{spec}$ and $AD\times E_{spec}$_Inb, especially for GY (both populations) and GM in Het2 population (Figure 3**,** Figure S3). However, we identified weaker A×E and D×E when using model $ADI\times E_{spec}$_Inb, with D×E effects being almost entirely replaced by epistatic by environment effects (Figure 3 **,** Figure S3). There, the three epistatic interactions (AA×E, AD×E and DD×E) were detected with variable proportions depending on trait and population. Of note, AD×E was nearly absent in iF2 population when assessed with $ADI\times E_{spec}$_Inb which is in contrast with the results obtained with $ADI\times E_{com}$_Inb, where AD×E were the main G×E effects identified (Figure 3B, Figure S3).

To gain further insights into differences between these model declinations, we tested the superiority of $G\times E_{spec}$ models (having more parameters) over $G\times E_{com}$ models and found no significant differences between them based on likelihood values (LRT values were always smaller than χ^2 values at $p<0.05$).

Genomic predictions

We predicted $\overline{Y_h}$ phenotypic values in cross-validations using model *M3* and submodels (scenario "Global", see Methods and Figure 1) and obtained mean predictive abilities ranging from 0.59 to 0.75 in Het2, and from 0.56 to 0.80 in iF2 population (Table 1, Figure 4, Table S4). Predictive abilities were higher in iF2 than in Het2 population, especially for HT (mean across models 0.80 vs. 0.66). In both populations, including the effect of inbreeding in the model improved predictions for all traits except for GM, which is the only trait on which inbreeding had no significant effect (Figure 4, Table 2). Predictive abilities increased of 0.02 on average when comparing either model A_Inb to A or model AD_Inb to AD (traits FLO, HT and GY, both populations, Table S4). Including the effect of inbreeding had the highest impact on predictions of GY in iF2 population (+0.08 from model A to A_Inb, +0.04 from model AD to AD_Inb). Regarding GM, adding inbreeding into the model seemed to have an overall negative impact on predictive abilities when considering models A_Inb and AD_Inb (Table S4**)**. Adding epistatic effects into the model allowed improving further predictions in iF2 population (mean increase of 0.01 between models AD_Inb and ADI_Inb (Figure 4B, Table S4).

Predictive abilities in the "Within-environment" scenario (see Methods and Figure 1) obtained with the five models ranged from 0.34 to 0.72 in Het2 and from 0.43 to 0.75 in iF2 population (Figure 5, Table S5). In Het2 population, slightly increased predictive abilities were observed using models including inbreeding as fixed effect for GY and with models of increasing complexity for FLO (Figure 5A). We also observed overall lower accuracies for environment smh16, comparatively to other environments (Figure 5A, Table S5). In iF2 population, a marked increase in predictive ability was obtained for GY and FLO with the three models including inbreeding as a fixed effect (Figure 5B). There, ADI_Inb performed overall best when compared to model A, resulting in +0.02 to +0.08 in mean predictive ability for GY (all environments) and in +0.02 to 0.03 for FLO (mln11 and smh10 Figure 5B, Table S5). For traits HT and GM model choice had a more limited impact on predictive abilities; changes in predictive abilities were not consistent across models and environments in Het2

456 population (Figure 5A) while in iF2 population model ADI_Inb tended to perform overall
457 better than other models (Figure 5B).

458 Last, we calibrated the model over all environments using model *M4* and submodels
459 and performed across-environment predictions either by predicting hybrids in a single new
460 environment chosen randomly ("G×E_new_env") or by predicting the phenotypes of new
461 hybrids in all environments ("G×E_new_hybrid", see Methods and Figure 1). When
462 predicting hybrids in a new environment ("G×E_new_env") we obtained mean predictive
463 abilities per environment ranging between 0.41 and 0.92 for Het2, and between 0.43 and 0.84
464 for the iF2 population (Table S7, Figure S4). In some cases, predictive ability varied strongly
465 depending on the validation run, resulting in a larger distribution of predictive ability values
466 (especially when considering population Het2 in sou17 for FLO and in smh17 for HT and
467 GM, Figure S4). When comparing the performance across all models, the largest differences
468 were observed between models with common vs. environment-specific G×E variance terms
469 ($G\times E_{com}$ vs. $G\times E_{spec}$ models). These differences were generally consistent across the five
470 model variations A×E, A×E_Inb, AD×E, AD×E_Inb and ADI×_Inb. For example, when
471 considering predictions of iF2 hybrids for GY in environment mln11, all $G\times E_{spec}$ models
472 performed significantly better than all $G\times E_{com}$ models (Figure 6A). Whether $G\times E_{com}$ or
473 $G\times E_{spec}$ models performed best was highly depending on the combination of trait and
474 environment (Figure S4). When comparing performances among models A×E, A×E_Inb,
475 AD×E, AD×E_Inb within each class of models $G\times E_{com}$ and $G\times E_{spec}$, we found that
476 ADI×E_Inb models allowed for a modest but consistent increase in predictive ability
477 compared to other models (mean increase of 0.006 to 0.010 when compared to A×E, Figure
478 S4, Table S6). Of note, model ADI_{spec}_Inb provided the highest predictive abilities for GM
479 and GY in iF2 population, which was particularly visible for GY in environment smh11
480 (mean pairwise difference of 0.07 when compared to other $G\times E_{spec}$ models, Figure 6A, Figure
481 S4, Table S6).

482 When applying the prediction scenario 'G×E_new_hyb' (Figure 1), we obtained mean
483 predictive abilities per environment in the range of 0.28 to 0.74 for Het2 and of 0.37 to 0.78
484 for iF2 population (Table S7). Comparing models within $G\times E_{com}$ and $G\times E_{spec}$ (five model

485 variations A×E, A×E_Inb, AD×E, AD_Inb×E, ADI_Inb×E), we found prediction
486 improvements when incorporating inbreeding into the model (all traits except GM).
487 ADI×E_Inb models were most frequently associated with the highest predictive abilities
488 (Figure 6B, Figure S4, Table S6). The maximum improvement was obtained using model
489 ADI×E_{com}_Inb for GM in environment smh17 (increase of 0.059 compared to A×E_{com}, Table
490 S7). We can note that in Het2 population, as observed in within environment predictions
491 (Figure 5), a gradual increase in predictive ability was observed with increasing model
492 complexity for FLO (Figure S4, Table S7). G×E_{spec} models performed overall better than
493 G×E_{com} models in iF2 population, the highest contrast being visible for GY, especially when
494 considering environment mln11 (pairwise differences of 0.14 on average between each
495 G×E_{spec} and each corresponding G×E_{com} model Figure 6A, Table S7). Considering Het2,
496 although we cannot conclude that either G×E_{com} or G×E_{spec} class of models performs overall
497 better, we can note that in most trait-environment combinations one class of models was
498 outperforming the other (Figure S4). For example, G×E_{com} models allows predicting GY
499 much more accurately than G×E_{spec} in environment smh16 (pairwise differences of 0.18 on
500 average), while G×E_{spec} tend to perform better in environment jar16 (pairwise differences of
501 0.02 on average). Differences were generally consistent across the five model variations A×E,
502 A×E_Inb, AD×E, AD×E_Inb and ADI×_Inb (at the exception of predictions of GY in sou17,
503 Het2 population, Figure S4**)**.

504 DISCUSSION

505 *Impact of inbreeding on phenotypic variation in Het2 and iF2 populations*

506 It is known that inbreeding level correlates with heterosis within heterotic groups, while
507 correlation is generally weak when considering inter-group hybrids (Charcosset and Essioux
508 1994; Burstin and Charcosset 1997, Melchinger 1999). First, loci contributing to heterosis are
509 expected to be fixed differentially among groups, leaving variation in inbreeding coefficient
510 among hybrids mostly in other regions. Second, differences in LD between heterotic groups
511 decrease the correlation even if variation within groups persists in genomic regions involved
512 in heterosis. Here, we dissected phenotypic variation in two populations admixed between
513 dent and flint heterotic groups where variation at each locus is due to both inter- and intra-

group genetic variation. We found in both populations that inbreeding, measured as the proportion of homozygous loci, had a significant effect on GY and FLO, on HT (only in iF2) and no effect on GM. This was shown by variance decompositions and validated by genomic predictions in all contexts. Our results are consistent with previous reports finding a significant correlation between inbreeding and GY, FLO and HT (Larièpe *et al.* 2012; Ramstein *et al.* 2020) but not GM (Larièpe *et al.* 2012). Further, we found that increasing inbreeding of 0.01 reduced GY from -1.53 to -0.21 q/ha, a range overlapping with previously reported values using pedigree (from -0.78 to -0.2 q/ha in Hallauer *et al.* 2010) or genomic data (-0.69 q/ha Ramstein *et al.* 2020). To note, MAF filtering could have an impact on the assessment of inbreeding, as deleterious mutations are expected to be present at low frequencies. In this context, a recent study showed that MAF can have a significant impact on the partition of genetic variance, with a tendency for lower MAF classes to be associated with larger polygenic SNP effects (Ramstein *et al.* 2020). In our case, a 5% MAF threshold was chosen to ensure a high quality of the genotyping dataset but the impact of MAF filtering on variance decompositions and prediction could be further investigated with comparative analyses. The detrimental impact of homozygosity proportions on trait values indicates a positive directional dominance effect, a feature underlying heterosis which may evolve via directional selection on fitness-related traits (Lynch and Walsh 1998). Delayed flowering and shorter development of hybrid plants with increasing inbreeding could reflect a decreased developmental rate with inbreeding (Charlesworth and Charlesworth 1987; Charlesworth and Willis 2009). Regarding GM, we found no clear interpretation for the absence of directional selection. Dynamic G×E and epistatic genetic interactions during kernel development might prevail in the establishment of grain moisture at seed maturity (Li *et al.* 2021).

Orthogonal estimation of statistical non-additive effects and relation with inbreeding

This study aimed at testing recent methodological improvements for quantifying non-additive genetic variance in two maize populations while preventing for a potential inflation of non-additive variance terms caused by the non-orthogonality of genetic effects. Accordingly, we found that the total variance was relatively stable across models (Table 1, Table S2-S3).

Incorporating dominance effects reduced the error variance overall with only slight reductions in the additive variance (for example from 30.72 to 29.03 for GY in iF2 considering $\overline{Y_{h..}}$ values). This is congruent with the low correlations found between kinship matrices for additivity and dominance (-0.005 in iF2 and –0.070 in Het2, off-diagonal values, Table S8). However, when including epistatic effects with the ADI_Inb model, epistatic variance replaced some of the additive and almost all dominance variance for most traits. This indicates that, despite being declared as independent in the model, the epistatic effects are not decorrelated from dominance and additive effects using this method on our populations. In fact, even when considering only off-diagonal values, some correlations between additive or dominant and epistatic kinship matrices were above 0.30 and the highest values were found between A and AA kinship matrices (0.71 for Het2, 0.34 for iF2, Table S8), a range overlapping with that observed in maize by González-Diéguez *et al.* (2021). Overall, our results indicate that orthogonality is not always guaranteed with the NOIA approach. As illustrated in our and other studies, it seems that the presence of LD in breeding populations prevents the orthogonalization of kinship matrices using this method (Vitezica *et al.* 2017; González-Diéguez *et al.* 2021). In fact, LD generates non-independence between genetic effects by essence (Hill and Mäki-Tanila 2015) and LD can mimic epistatic effects (Wade *et al.* 2001; de los Campos *et al.* 2019). Although LD hinders the precise dissection of genetic variance into its components, estimates are expected to be more consistent with the NOIA approach than with orthogonal models which are valid under more restrictive conditions (Vitezica *et al.* 2017; Álvarez-Castro and Crujeiras 2019; Yadav *et al.* 2021). Also, predictive ability should not be affected because the total genetic variance should be correctly estimated with the NOIA approach (Vitezica *et al.* 2017).

It is known that in the presence of directional dominance, dominance variance may be inflated if inbreeding is not included in the model (de Boer and Hoeschele 1993; Aliloo *et al.* 2016; Vitezica *et al.* 2018; Varona *et al.* 2018b). Similarly, an overestimation of the SCA variance has been reported when the distance between parents is not accounted for (Larièpe *et al.* 2017). We confirmed this trend, as using model AD_Inb led in some cases to a significant reduction of the dominance variance when compared to model AD. The variance reduction was more pronounced in analyses on mean values than in within-environment or across-

environment analyses (Table 2, Figure 2-3). Overall, this indicates that our genome-wide indicator of inbreeding (the homozygosity proportion) captures mean dominance effects and thus contains part of the information present in the marker-based dominance kinship (Aliloo *et al.* 2016, Yadav *et al.* 2021). Adding inbreeding as a cofactor in the model thus appears to be necessary in the presence of directional dominance, which also contributes to more realistic estimates of the variance caused by statistical dominance.

Declaring a residual hybrid effect could capture genetic effects which are not declared in the model or not captured by SNP markers (see for example "residual genetic" in González-Diéguez *et al.* 2021). To check this we performed additional analyses by adding a random "hybrid" effect with no covariance structure to models A, A_Inb, AD, AD_Inb and ADI_Inb in within environment variance decompositions. We found that this can decrease BIC values when considering models A and A_Inb (for example FLO jar16, GM mln11, mean decrease of – 4.55 in iF2, marginal increase of 0.544 in Het2 Figure S5). We observed a drop of the variance allocated to the residual hybrid effect effect to near-zero with model ADI_Inb, which also corresponds to the greatest increases in BIC when compared to the model ADI_Inb without residual hybrid effect effect (5.7 in Het2 and 5.1 in iF2, Figure S6). Thus, it seems that epistatic effects are partially captured by additive and residual hybrid variance terms and that declaring a residual hybrid effect should be avoided when quantifying non-additive variance.

Comparison of inbreeding and non-additive effects between iF2 and Het2 populations

The smaller average inbreeding level observed in iF2 seems, at first sight, to contrast with the clearly poorer agronomical performance of iF2 hybrids when compared to Het2 hybrids (roughly halved GY values, Table 1, Figure S2). However, it is worth mentioning that we cannot compare inbreeding values between both populations in an absolute way because they were genotyped at different densities and only polymorphic markers were retained in each population (see Methods). In Het2, a larger number of independent genetic segments and a greater allelic diversity were expected by design, when compared to iF2. Indeed, Het2 was originally founded with admixed individuals resulting from crosses between 300 dent and 304

flint parents whereas iF2 hybrids have been generated with only one dent and one flint parents. This led to more polymorphic markers being retained in Het2, each with an increased probability for homozygosity. Another point is that in genomic prediction, inbreeding must be sufficiently variable in the population to be exploited. For example, this was not the case in the inter-group hybrid population studied by González-Diéguez *et al.* 2021, where no strong inbreeding effects were found and thus did not contribute to improved genomic prediction accuracies. Here, standard deviation of homozygosity proportions was 3.6 times larger in iF2 than in Het2 (Figure S2) which might explain why improvements in predictive ability with models considering inbreeding (A_Inb, AD_Inb and ADI_Inb) were larger in iF2 than in Het2.

Het2 hybrids were characterized by lower epistatic and Epi×E effects when compared to iF2 (Table 1, Figure 2-3). Epistasis kinships values tended to be more variable in iF2, which could explain this contrast (Figure S2). Trait variation was also underlined by larger dominant and D×E effects in Het2 than in iF2 population, but dominance kinship variation seemed to be larger in iF2 population (Figure S2) and models taking dominance and/or D×E effects into account did not improve genomic predictions in Het2 as much as in iF2 (Figure 4-**6**). This could imply that dominance effects have been overestimated in Het2 relative to the additive variance.

Non-additive and G×E effects reflect differences in trait complexity

Complex traits are typically underlined by significant non-additive and G×E effects which makes them less heritable than simple traits (Stranger *et al.* 2011). Because breeding targets traits of varying complexity level, quantifying these effects and their relative proportions is useful for defining breeding strategies. Although variance decomposition does not allow for an accurate estimation of biological effects or gene-action modes (Huang and Mackay 2016), it is believed that orthogonal models offer the potential to get closer to the most biologically meaningful scenario (Alvarez-Castro et al. 2019). We explored four standard agronomical traits and found (unsurprisingly) that dominance and epistasis were strongest in GY, a highly complex trait displaying strong heterosis (Table 2, Figure 2-3, Hill *et al.* 2008; Li *et al.* 2013;

631 Larièpe *et al.* 2017; Samayoa *et al.* 2017). Large dominance and epistasis variances were also
632 found for FLO, GM and HT but they were milder and less consistent across environments and
633 populations than for GY (Table 2, Figure 2). To note, different measurements were done for
634 FLO in Het2 and iF2 (male flowering in calendar days in Het2 and female flowering in GDD
635 for iF2), which is expected to affect more G×E interactions than genetic components. Among
636 epistatic variance terms, AA variance prevailed across traits (Table 2, Figure 2-3). Owing to
637 the difficulties to detect epistasis and to contradicting results across studies, the significance
638 of epistasis for plant phenotypic variation is still debated (Lynch and Walsh 1998; Wade *et al.*
639 2001; Buckler *et al.* 2009; Jiang and Reif 2015). An epistatic control for FLO has been
640 demonstrated in several species such as barley (Mathew *et al.* 2018), rice (Lin *et al.* 2000; Li
641 *et al.* 2014) and *Arabidopsis* (Juenger *et al.* 2005). In maize, allelic effects at QTLs linked to
642 flowering date seem to strongly depend on genetic background (Durand *et al.* 2012; Rio *et al.*
643 2020) and non-linear effects of inbreeding on flowering suggest complex gene interactions
644 (Ramstein *et al.* 2020).

645 G×E effects were predominant for GY, and weaker for the three other traits. Also, dominance
646 and epistasis effects were usually stronger when considering their interaction with the
647 environment than taken alone, which was not the case for additive effects (Figure 3**,** Table
648 S3). This indicates that, independent of trait and population, additive genetic variance appears
649 to be less sensitive to environmental effects. This contrasts with former studies having found
650 either a prevalence of GCA×E over SCA×E effects (linseed : Bhateria *et al.* 2006; maize :
651 Acosta-Pech *et al.* 2017) or varying proportions of these effects depending on the trait
652 (cassava : Parkes *et al.* 2013). Interestingly, declaring G×E variance terms as environment
653 specific ($G\times E_{spec}$ models) increased the proportions of A×E at the cost of the additive, D×E
654 and Epi×E variance terms for all traits, which deserves further investigation. It was also
655 difficult to conclude which of the three Epi×E effects (AA×E, AD×E and DD×E) was largest
656 across traits and populations. It has been argued that Epi×E variance contributes to phenotypic
657 plasticity and is an important component of plant adaptation (Lukens and Doebley 1999; Fethi
658 *et al.* 2011; Kerwin *et al.* 2017). We could not find a study having quantified altogether A×E,
659 D×E, AA×E, AD×E and DD×E effects on plant phenotypes. Previous works have found
660 significant AA×E effects on HT (rice : Cao *et al.* 2001; wheat : Zhang *et al.* 2008), GY

(cotton : McCarty *et al.* 2004; rice : Liu *et al.* 2006; maize : Ma *et al.* 2007). Also, epistatic effects could be amplified under stressed conditions which could add to Epi×E effects when studying jointly highly contrasted environments (reviewed in Fethi *et al.* 2011).

Improving genomic predictions with inbreeding and non-additive genetic parameters

Our goal was to test whether we can efficiently control for inbreeding and exploit non-additive effects using adequate prediction models. Having seen that the strongest inbreeding and epistatic effects were observed in iF2 population, we found that prediction models incorporating these effects did perform better than models A and AD (model ADI_Inb generally performing at best: Figure 4B, 5B, 6B). Regarding population Het2, improvements with models incorporating dominance and inbreeding (A_Inb and AD_Inb) were clear for FLO and GY and less consistent for HT and GM, with both models AD_Inb and ADI_Inb yielding overall the best performances (Figure 4A, 5A, Figure S4). In summary, the ADI_Inb model appeared to be the most reliable model across traits and populations in terms of prediction. This result might be surprising, as we have seen that the presence of LD causes imperfect estimations of each genetic effect. However, simulations suggests that LD does not impair the higher effectiveness of models including epistasis when compared to standard G-BLUP for GY in maize and wheat (Jiang and Reif 2015).

In many studies, including dominance effects in prediction models did not improve accuracies (Pégard *et al.* 2020 and references therein, maize : González-Diéguez *et al.* 2021, maize, pig, cattle : Zhang *et al.* 2019) but we also found evidence that dominance can clearly contribute to improve hybrid prediction (maize and wheat: Jiang and Reif 2015; maize : Ramstein *et al.* 2020; sorghum : Ishimori *et al.* 2020 in sorghum; dairy cattle : Aliloo *et al.* 2019; *Eucalyptus pellita* : Thavamanikumar *et al.* 2020). In our work, adding dominance without accounting for inbreeding improved predictions only in some cases (model AD, *e.g.* predictions for GY in iF2, Figure 4), which reflects that dominance may have been inflated in AD models (discussed above). In contrast, we found strong evidence for directional dominance and a large improvement of predictions accuracies when using models A_Inb, AD_Inb and ADI_Inb in all context and both populations (except for trait GM). Including inbreeding in

genomic prediction models was advocated in the presence of dominance (*e.g.* Iversen *et al.* 2019 for pig) but led to inconclusive results in maize: Ramstein *et al.* (2020) report a significant directional dominance but no improvements of accuracies and González-Diéguez (2021) concluded that inbreeding was not enough contrasted within the population to contribute to better accuracies. Yet, it was shown recently in a population of sugar cane elite clones that incorporating both heterozygosity as fixed parameter and non-additive (dominant and additive by additive) variances using the NOIA framework improved significantly prediction accuracy for cane yield (Yadav *et al.* 2021). To summarize, it appears that failure to improve prediction accuracy using non-additive or inbreeding effects could either stem from the intrinsic low contribution of these effects in the population of study (biological origin, *e*.g. mating system; study design, *e.g.* population size or crossing design, Bolormaa *et al.* 2015) or from an inappropriate modelling of these effects. For example when non-orthogonality between kinship matrices generates strong confounding between genetic effects or when dominance but not epistasis is declared in the model, non-additive effects might be estimated inaccurately (Vitezica *et al.* 2017, Bolormaa *et al.* 2015, Ramstein *et al.* 2020). Despite having identified some degree of correlation between our kinship matrices, we have seen that total genetic variance was stabilized across models. The improvements in genomic prediction we observed with the ADI_Inb model thus not only reflect that non-additive effects (dominance and epistasis) were pronounced in our populations but also that with an appropriate modelling, they contributed to a better estimation of the genetic variance. Our cross-validations were performed within dent × flint backgrounds and with relatively high levels of relatedness between TS and VS. It remains to be tested whether the ADI_Inb model proves more efficient than simple additive models for genomic prediction across populations. Increasing diversity in the training set could allow for weaker levels of LD and for larger contrasts in terms of additive and non-additive kinship and inbreeding coefficients, but as is generally acknowledged, relationship must be maintained at a sufficiently high level between TS and VS (Roth et *al.* 2020, Rio et al.2019).

Choosing the best prediction model *a priori* (*i.e.* before training and evaluating experimentally the predictive ability of models) could lead to significant gains in time and efficiency. We tried to evaluate whether goodness of fit could be a reliable criterion for model

choice. Overall, although lower BIC values were often associated with the best performing prediction model, correlation between mean predictive ability and BIC value across all traits was weak (Figure S7). Our results confirm that the BIC should not be the only model selection criterion (Lebarbier and Mary-Huard 2004). An alternative could be to consider the Akaike Information Criterion (AIC, Lebarbier and Mary-Huard 2004), however in our results AIC and BIC led to the selection of the exact same models (data not shown).

Choosing an appropriate model for multi-environment genomic predictions

The large G×E interactions found in our variance decompositions (Figure 3**,** Figure S3) indicated that these effects can be exploited to improve genomic predictions across different environments. We tested two alternative ways to declare G×E in the models ($G\times E_{spec}$ *vs.* $G\times E_{com}$) and found that environment-specific variance terms performed overall better than global variance terms. Our cross-validation design also reflected two practical settings where predicted hybrids have already been observed in two to four environments (G×E_new_env) or not (G×E_new_hybrid, Figure 1). Unsurprisingly, predictive ability was much higher in the former than in the latter scenario. Including parameters controlling for non-additive and inbreeding effects led to marginal improvements in the G×E_new_env context, which could indicate that environment-specific (fixed) and A×E (random) effects may be sufficient to adjust the phenotypic values when the predicted hybrid has already been characterized in other environments. In contrast, the inclusion of non-additive and inbreeding effects in the G×E_new_hybrids scenario did contribute to higher predictive abilities (Figure 6B) in a similar trend to that observed in the prediction of global means ($\overline{Y_{h..}}$ values, Figure 4) or within-environment phenotypic values ($\widetilde{Y_{her}}$ values, Figure 5).

In the light of these results, we wondered whether it is worth using multi-environment prediction models to predict unobserved hybrids when calibration can be done in single environments in a simpler manner. We addressed this question by comparing predictions between G×E_new_hybrids and within-environment scenarios. Multi-environment predictions performed overall better than single-environment predictions (mean increase in

pairwise comparisons 0.034 for Het2, 0.022 for iF2, Figure S8). A considerable increase in predictive ability was observed for Het2 in environment smh16 when using G×E modelling (for example +0.16 in predictive ability for GY when considering $G\times E_{com}$ models, Figure S4). In fact, the lower heritability values found in smh16 may have been partially compensated by adding environments with higher heritability in the training set, which contributed to a more accurate estimation of genotypic effects. Besides, declaring G×E variances as environment-specific ($G\times E_{spec}$) allowed more precise predictions in environments which were more weakly correlated to the other environments from the training set (see predictions of GY in mln11 for iF2, Figure 6**A and Fig S1)**. Hence, it appears that multi-environment calibration with models incorporating G×E can be very useful to improve predictions of unobserved hybrids in environments with lower heritability, provided that phenotypes in the target environment are sufficiently correlated to those of environments used for calibration.

Potential use of admixture and long-term diversity management

Maize breeding traditionally relies on crosses between two inbred lines from different heterotic groups to maximize the contribution of statistical additive effects to total genetic variance (Larièpe *et al.* 2017). The reproductive isolation between groups causes a reduction of genetic diversity via drift which is stronger when considering groups individually than the whole population (Gerke *et al.* 2015). Indeed, as drift affects the genome in a random way, different regions are fixed in the different groups (Reif *et al.* 2005; Schön *et al.* 2010). However, drift remains a problem when considering diversity management in the long term. Here, iF2 and Het2 hybrids have overall low average performances but some Het2 hybrids performed as well as commercial hybrids, for example Het_354152 has a mean GY of 112.5 q/ha. In an iF2 population also derived from a dent and a flint lines, Guo and collaborators (2013) found only very few iF2 hybrids that outperformed the ‘reference’ parental hybrid for yield. This illustrates that a favourable disruption of the long-established heterotic patterns is possible but rare. Thus, admixed lines should be rather regarded overall as a reservoir for pre-breeding, than a reservoir of breeding lines (see also “external program” in Allier *et al.* 2020). In these admixed populations, non-additive variance and variations in inbreeding levels

translate into differences in hybrid vigour, which we were able to exploit in predictions. Importantly, the models we developed for single and multi-environment genomic prediction can be applied to other types of breeding material (admixed or not) and species to obtain a better estimation of genetic variance, and are expected to be particularly powerful in the presence of directional dominance. In fact, while our populations can be regarded as extreme cases of admixture (especially in iF2 where genetic diversity is reduced at maximum), our models could be applied to other contexts: (i) starting a program from a commercial hybrid due to the restricted access to elite inbred lines, a strategy frequently used in developing countries (Guo *et al.* 2013), (ii) introgressing material outside from the heterotic pool (ex. other heterotic pool, open pollinated varieties, exotic material) in an elite genetic background, (iii) merging long-established breeding programs when merging companies. Our results highlight that genomic prediction modelling can serve not only for improving evaluation accuracy in established breeding programs but also to develop new breeding strategies with the introduction of new breeding material for cultivar improvement (Voss-Fels *et al.* 2019; Liu *et al.* 2019).

Data and code availability statement

Raw genotyping, phenotypic data and custom R code used for data formatting, analysis and result plotting are stored in the Data INRAE dataset "Genomic prediction in admixed dent x flint hybrids" with following DOI number: https://doi.org/10.15454/ZGP766. The physical position of markers used for the genotyping of Het2 and iF2 parents are available in Rio *et al.* 2020 and Ganal *et al.* 2011 respectively. Supplemental Material available at figshare: https://doi.org/10.25386/genetics.19030004.

Acknowledgements

We thank Valerie Combes, Delphine Madur, and Stéphane Nicolas and Simon Rio (GQE, Le Moulon) for DNA extraction, analysis, and assembly of genotypic data. We thank Cyril Bauland (GQE, Le Moulon), Carine Palaffre, Bernard Lagardère, and Jean-René Loustalot

(INRAE Saint-Martin de Hinx) for the panel assembly and the coordination of seed production and the phenotyping of the two panels; all the breeding companies that are partners of the Amaizing project (Caussade Semences, Euralis, Limagrain, MAS Seeds, KWS, Syngenta) for the production of admixed parental lines of the Het2 panel, Euralis, Limagrain and MAS Seeds for their contribution to the phenotyping of this panel, and the company Limagrain for the genotyping of the admixed parental lines. We are grateful to the partners of the R2D2 Selgen project for helpful discussions.

We thank two anonymous reviewers for their helpful comments on earlier versions of the manuscript.

Funding

This research was supported by the ”Investissement d’Avenir” project ”Amaizing” (Amaizing, ANR-10-BTBR-0001) and the INRAE metaprogram SelGen. We thank Agroscope for granting further financial support to this study.

Conflicts of Interest

The authors declare no conflict of interest

REFERENCES

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

Aliloo, H., J. E. Pryce, O. González-Recio, B. G. Cocks, and B. J. Hayes, 2016 Accounting for dominance to improve genomic evaluations of dairy cows for fertility and milk production traits. Genet. Sel. Evol. 48:8. https://doi.org/10.1186/s12711-016-0186-0

Allier, A., S. Teyssèdre, C. Lehermeier, B. Claustres, S. Maltese, *et al.*, 2019 Assessment of

831 breeding programs sustainability: application of phenotypic and genomic indicators to a
832 North European grain maize program. Theor. Appl. Genet. 132: 1321–1334.
833 https://doi.org/10.1007/s00122-019-03280-w

834 Álvarez-Castro, J. M., and Ö. Carlborg, 2007 A unified model for functional and statistical
835 epistasis and its application in quantitative trait loci analysis. Genetics 176: 1151–1167.
836 https://doi.org/10.1534/genetics.106.067348

837 Álvarez-Castro J. M., and R. M. Crujeiras, 2019 Orthogonal Decomposition of the Genetic
838 Variance for Epistatic Traits Under Linkage Disequilibrium—Applications to the
839 Analysis of Bateson-Dobzhansky-Müller Incompatibilities and Sign Epistasis. Front.
840 Genet. 10: 1–12. https://doi.org/10.3389/fgene.2019.00054

841 Bernardo, R., 1992 Relationship between single-cross performance and molecular marker
842 heterozygosity. Theor. Appl. Genet. 83: 628–634. https://doi.org/10.1007/BF00226908

843 Bernardo, R., 1993 Estimation of coefficient of coancestry using molecular markers in maize.
844 Theor. Appl. Genet. 85: 1055–1062. https://doi.org/10.1007/BF00215047

845 Bhateria, S., S. P. Sood, and A. Pathania, 2006 Genetic analysis of quantitative traits across
846 environments in linseed (*Linum usitatissimum* L.). Euphytica 150: 185–194.
847 https://doi.org/10.1007/s10681-006-9106-7

848 Boer, I. J. M. de, and I. Hoeschele, 1993 Genetic evaluation methods for populations with
849 dominance and inbreeding. Theor. Appl. Genet. 86: 245–258.
850 https://doi.org/10.1007/BF00222086

851 Bolormaa, S., J. E. Pryce, Y. Zhang, A. Reverter, W. Barendse, *et al.*, 2015 Non-additive
852 genetic variation in growth, carcass and fertility traits of beef cattle. Genet. Sel. Evol. 47:
853 1–12. https://doi.org/10.1186/s12711-015-0114-8

854 Browning, S. R., and B. L. Browning, 2007 Rapid and accurate haplotype phasing and
855 missing-data inference for whole-genome association studies by use of localized
856 haplotype clustering. Am. J. Hum. Genet. 81: 1084–1097.
857 https://doi.org/10.1086/521987

Buckler, E. S., J. B. Holland, P. J. Bradbury, C. B. Acharya, P. J. Brown, *et al.*, 2009 The genetic architecture of maize flowering time. Science. 325: 714–718. https://doi.org/10.1126/science.1174276

Burstin, J., and A. Charcosset, 1997 Relationship between phenotypic and marker distances: theoretical and experimental investigations. Heredity. 79: 477–483. https://doi.org/10.1038/hdy.1997.187

de los Campos, G. de, D. A. Sorensen, and M. A. Toro, 2019 Imperfect linkage disequilibrium generates phantom epistasis (& perils of big data). G3 Genes, Genomes, Genet. 9: 1429–1436. https://doi.org/10.1534/g3.119.400101

Cao, G., J. Zhu, C. He, Y. Gao, J. Yan, *et al.*, 2001 Impact of epistasis and QTL×environment interaction on the developmental behavior of plant height in rice (*Oryza sativa* L.). Theor. Appl. Genet. 103. https://doi.org/10.1007/s001220100536

Charcosset, A., M. Lefort-Buson, and A. Gallais, 1991 Relationship between heterosis and heterozygosity at marker loci: a theoretical computation. Theor. Appl. Genet. 81: 571–575. https://doi.org/10.1007/BF00226720

Charcosset, A., and L. Essioux, 1994 The effect of population structure on the relationship between heterosis and heterozygosity at marker loci. Theor. Appl. Genet. 89: 336–343. https://doi.org/10.1007/BF00225164

Charlesworth, D., and B. Charlesworth, 1987 Inbreeding depression and its evolutionary consequences. Annu. Rev. Ecol. Syst. 18: 237–268. https://doi.org/10.1146/annurev.es.18.110187.001321

Charlesworth, D., and J. H. Willis, 2009 The genetics of inbreeding depression. Nat. Rev. Genet. 10: 783–796. https://doi.org/10.1038/nrg2664

Durand, E., S. Bouchet, P. Bertin, A. Ressayre, P. Jamin, *et al.*, 2012 Flowering time in maize: Linkage and epistasis at a major effect locus. Genetics 190: 1547–1562. https://doi.org/10.1534/genetics.111.136903

Falque, M., L. Décousset, D. Dervins, A.-M. Jacob, J. Joets, *et al.*, 2005 Linkage mapping of

1454 new maize candidate gene loci. Genetics 170: 1957–1966. https://doi.org/10.1534/genetics.104.040204

Fethi, B., C. Hanbary, and E. G. Mohamed, 2011 Genetic adaptability of inheritance of resistance to biotic and abiotic stress level on crop: Role of epistasis. African J. Biotechnol. 10: 19913–19917. https://doi.org/10.5897/AJBX11.067

Fiévet, J. B., C. Dillmann, and D. de Vienne, 2010 Systemic properties of metabolic networks lead to an epistasis-based model for heterosis. Theor. Appl. Genet. 120: 463–473. https://doi.org/10.1007/s00122-009-1203-2

Frankel, R., 1983 *Heterosis - Reappraisal of theory and practice*. Berlin Heidelberg New York Tokyo.

Ganal,, M. W., G. Durstewitz, A. Polley, A. Bérard, E. S. Buckler, *et al.*, 2011 A marge maize (*Zea mays* L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. PLoS One 6: e28334. https://doi.org/10.1371/journal.pone.0028334

Garci, A. A. F., S. Wang, A. E. Melchingerand, and Z. B. Zeng, 2008 Quantitative trait loci mapping and the genetic basis of heterosis in maize and rice. Genetics 180: 1707–1724. https://doi.org/10.1534/genetics.107.082867

Gerke, J. P., J. W. Edwards, K. E. Guill, J. Ross-Ibarra, and M. D. McMullen, 2015 The genomic impacts of drift and selection for hybrid performance in maize. Genetics 201: 1201–1211. https://doi.org/10.1534/genetics.115.182410

González-Diéguez, D., A. Legarra, A. Charcosset, L. Moreau, C. Lehermeier, *et al.*, 2021 Genomic prediction of hybrid crops allows disentangling dominance and epistasis, (J. B. Endelman, Ed.). Genetics 218: iyab026. https://doi.org/10.1093/genetics/iyab026

Gordillo, G. A., and H. H. Geiger, 2008 Alternative recurrent selection strategies using doubled haploid lines in hybrid maize breeding. Crop Sci. 48: 911–922. https://doi.org/10.2135/cropsci2007.04.0223

Guo T., H. Li, J. Yan, J. Tang, J. Li, et al., 2013 Performance prediction of F1 hybrids

between recombinant inbred lines derived from two elite maize inbred lines. Theor. Appl. Genet. 126: 189–201. https://doi.org/10.1007/s00122-012-1973-9

Hallauer, A. R., J. B. M. Filho, and M. J. Carena, 2010 Inbreeding, pp. 425–475 in *Quantitative Genetics in Maize Breeding*, Springer New York.

Hill, W. G., M. E. Goddard, and P. M. Visscher, 2008 Data and theory point to mainly additive genetic variance for complex traits. PLoS Genet. 4: e1000008. https://doi.org/10.1371/journal.pgen.1000008

Hill, W. G., and A. Mäki-Tanila, 2015 Expected influence of linkage disequilibrium on genetic variance caused by dominance and epistasis on quantitative traits. J. Anim. Breed. Genet. 132: 176–186. https://doi.org/10.1111/jbg.12140

Hua J. P., Y. Z. Xing, C. G. Xu, X. L. Sun, S. B. Yu, et al., 2002 Genetic dissection of an elite rice hybrid revealed that heterozygotes are not always advantageous for performance. Genetics 162: 1885–1895. https://doi.org/10.1093/genetics/162.4.1885

Ishimori, M., T. Hattori, K. Yamazaki, H. Takanashi, M. Fujimoto, et al., 2020 Impacts of dominance effects on genomic prediction of sorghum hybrid performance. Breed. Sci. 70: 605–616. https://doi.org/10.1270/jsbbs.20042

Iversen M. W., Ø. Nordbø, E. Gjerlaug-Enger, E. Grindflek, M. S. Lopes, et al., 2019 Effects of heterozygosity on performance of purebred and crossbred pigs. Genet. Sel. Evol. 51: 1–13. https://doi.org/10.1186/s12711-019-0450-1

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

Jiang, Y., and J. C. Reif, 2015 Modeling epistasis in genomic selection. Genetics 201: 759–768. https://doi.org/10.1534/genetics.115.177907

Jones, D. F., 1917 Dominance of linked factors as a means of accounting for heterosis. Genetics 2: 609–609. https://doi.org/10.1093/genetics/2.6.609a

Juenger, T. E., S. Sen, K. A. Stowe, and E. L. Simms, 2005 Epistasis and genotype-

environment interaction for quantitative trait loci affecting flowering time in *Arabidopsis thaliana*, pp. 87–105 in *Genetics of Adaptation*, Springer-Verlag, Berlin/Heidelberg.

Kadam, D. C., S. M. Potts, M. O. Bohn, A. E. Lipka, and A. J. Lorenz, 2016 Genomic prediction of single crosses in the early stages of a maize hybrid breeding pipeline. G3 Genes, Genomes, Genet. 6: 3443–3453. https://doi.org/10.1534/g3.116.031286

Kerwin, R. E., J. Feusier, A. Muok, C. Lin, B. Larson, *et al.*, 2017 Epistasis × environment interactions among *Arabidopsis thaliana* glucosinolate genes impact complex traits and fitness in the field. New Phytol. 215: 1249–1263. https://doi.org/10.1111/nph.14646

Labroo, M. R., A. J. Studer, and J. E. Rutkoski, 2021 Heterosis and hybrid crop breeding: A multidisciplinary review. Front. Genet. 12: 1–19. https://doi.org/10.3389/fgene.2021.643761

Laporte, F., and T. Mary-Huard, 2019 MM4LMM: Inference of Linear Mixed Models Through MM Algorithm

Larièpe, A., B. Mangin, S. Jasson, V. Combes, F. Dumas, *et al.*, 2012 The genetic basis of heterosis: Multiparental quantitative trait loci mapping reveals contrasted levels of apparent overdominance among traits of agronomical interest in maize (*Zea mays* L.). Genetics 190: 795–811. https://doi.org/10.1534/genetics.111.133447

Larièpe, A., L. Moreau, J. Laborde, C. Bauland, S. Mezmouk, *et al.*, 2017 General and specific combining abilities in a maize (*Zea mays* L.) test-cross hybrid panel: relative importance of population structure and genetic divergence between parents. Theor. Appl. Genet. 130: 403–417. https://doi.org/10.1007/s00122-016-2822-z

Lebarbier, E., and T. Mary-Huard, 2004 Le critère BIC : fondements théoriques et interprétation. Technical Report R-5315 INRIA 1–17.

Legarra, A., 2016 Comparing estimates of genetic variance across different relationship models. Theor. Popul. Biol. 107: 26–30. https://doi.org/10.1016/j.tpb.2015.08.005

Li, C., Y. Li, B. Sun, B. Peng, C. Liu, *et al.*, 2013 Quantitative trait loci mapping for yield components and kernel-related traits in multiple connected RIL populations in maize.

966 Euphytica 193: 303–316. https://doi.org/10.1007/s10681-013-0901-7

967 Li, G., A. Chen, X. Liu, W. Wang, H. Ding, *et al.*, 2014 QTL detection and epistasis analysis
968 for heading date using single segment substitution lines in rice (*Oryza sativa* L.). J.
969 Integr. Agric. 13: 2311–2321. https://doi.org/10.1016/S2095-3119(13)60615-2

970 Li W., Y. Yu, L. Wang, Y. Luo, Y. Peng, et al., 2021 The genetic architecture of the dynamic
971 changes in grain moisture in maize. Plant Biotechnol. J. 19: 1195–1205.
972 https://doi.org/10.1111/pbi.13541

973 Lin, H. X., T. Yamamoto, T. Sasaki, and M. Yano, 2000 Characterization and detection of
974 epistatic interactions of 3 QTLs, Hd1, Hd2, and Hd3, controlling heading date in rice
975 using nearly isogenic lines. Theor. Appl. Genet. 101: 1021–1028.
976 https://doi.org/10.1007/s001220051576

977 Lin, Z., F. Shi, B. J. Hayes, and H. D. Daetwyler, 2017 Mitigation of inbreeding while
978 preserving genetic gain in genomic breeding programs for outbred plants. Theor. Appl.
979 Genet. 130: 969–980. https://doi.org/10.1007/s00122-017-2863-y

980 Liu, G.-F., J. Yang, and J. Zhu, 2006 Mapping QTL for biomass yield and its components in
981 rice (*Oryza sativa* L.). Acta Genet. Sin. 33: 607–616. https://doi.org/10.1016/S0379-
982 4172(06)60090-5

983 Liu, X., H. Wang, X. Hu, K. Li, Z. Liu, *et al.*, 2019 Improving genomic selection with
984 quantitative trait loci and nonadditive effects revealed by empirical evidence in maize.
985 Front. Plant Sci. 10: 1129. https://doi.org/10.3389/fpls.2019.01129

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

990 Lukens, L. N., and J. Doebley, 1999 Epistatic and environmental interactions for quantitative
991 trait loci involved in maize evolution. Genet. Res. 74: 291–302.
992 https://doi.org/10.1017/S0016672399004073

993 Lynch, M., and B. Walsh, 1998 *Genetics and analysis of quantitative traits*. Sunderland, MA,
994 USA.

995 Ma, X. Q., J. H. Tang, W. T. Teng, J. B. Yan, Y. J. Meng, *et al.*, 2007 Epistatic interaction is
996 an important genetic basis of grain yield and its components in maize. Mol. Breed. 20:
997 41–51. https://doi.org/10.1007/s11032-006-9071-9

998 Mathew, B., J. Léon, W. Sannemann, and M. J. Sillanpää, 2018 Detection of epistasis for
999 flowering time using Bayesian multilocus estimation in a barley MAGIC population.
1000 Genetics 208: 525–536. https://doi.org/10.1534/genetics.117.300546

1001 McCarty, J. C., J. N. Jenkins, and J. Wu, 2004 Primitive accession derived germplasm by
1002 cultivar crosses as sources for cotton improvement: II. Genetic effects and genotypic
1003 values. Crop Sci. 44: 1231–1235. https://doi.org/10.2135/cropsci2004.1231

1004 Melchinger A. E., 2009 Genetic diversity and heterosis, pp. 99–118 in *The Genetics and*
1005 *Exploitation of Heterosis in Crops*, edited by J. G. Coors and S. Pandey. American
1006 Society of Agronomy, Crop Science Society of America, and Soil Science Society of
1007 America, Madison, WI.

1008 Millet, E. J., W. Kruijer, A. Coupel-Ledru, S. Alvarez Prado, L. Cabrera-Bosquet, *et al.*, 2019
1009 Genomic prediction of maize yield across European environmental conditions. Nat.
1010 Genet. 51: 952–956. https://doi.org/10.1038/s41588-019-0414-y

1011 Muñoz, P. R., M. F. R. Resende, S. A. Gezan, M. D. V. Resende, G. de los Campos, *et al.*,
1012 2014 Unraveling additive from nonadditive effects using genomic relationship matrices.
1013 Genetics 198: 1759–1768. https://doi.org/10.1534/genetics.114.171322

1014 Parkes, E. Y., M. Fregene, A. Dixon, B. Boakye-Peprah, and M. T. Labuschagne, 2013
1015 Combining ability of cassava genotypes for cassava mosaic disease and cassava bacterial
1016 blight, yield and its related components in two ecological zones in Ghana. Euphytica
1017 194: 13–24. https://doi.org/10.1007/s10681-013-0936-9

1018 Pégard, M., V. Segura, F. Muñoz, C. Bastien, V. Jorge, *et al.*, 2020 Favorable conditions for
1019 genomic evaluation to outperform classical pedigree evaluation highlighted by a proof-

of-concept study in poplar. Front. Plant Sci. 11: 1–23. https://doi.org/10.3389/fpls.2020.581954

Ramstein, G. P., S. J. Larsson, J. P. Cook, J. W. Edwards, E. S. Ersoz, *et al.*, 2020 Dominance effects and functional enrichments improve prediction of agronomic traits in hybrid maize. Genetics 215: 215–230. https://doi.org/10.1534/genetics.120.303025

Reif, J. C., S. Hamrit, M. Heckenberger, W. Schipprack, H. P. Maurer, *et al.*, 2005 Trends in genetic diversity among European maize cultivars and their parental components during the past 50 years. Theor. Appl. Genet. 111: 838–845. https://doi.org/10.1007/s00122-005-0004-5

Reif, J. C., F.-M. Gumpert, S. Fischer, and A. E. Melchinger, 2007 Impact of interpopulation divergence on additive and dominance variance in hybrid populations. Genetics 176: 1931–1934. https://doi.org/10.1534/genetics.107.074146

Rio, S., T. Mary-Huard, L. Moreau, and A. Charcosset, 2019 Genomic selection efficiency and a priori estimation of accuracy in a structured dent maize panel. Theor. Appl. Genet. 132: 81–96. https://doi.org/10.1007/s00122-018-3196-1

Rio, S., T. Mary-Huard, L. Moreau, C. Bauland, C. Palaffre, *et al.*, 2020 Disentangling group specific QTL allele effects from genetic background epistasis using admixed individuals in GWAS: An application to maize flowering. PLoS Genet. 16: e1008241. https://doi.org/10.1371/journal.pgen.1008241

Robert, P., J. Le Gouis, and R. Rincent, 2020 Combining crop growth modeling with trait-assisted prediction improved the prediction of genotype by environment interactions. Front. Plant Sci. 11: 1–11. https://doi.org/10.3389/fpls.2020.00827

Samayoa, L. F., R. A. Malvar, and A. Butrón, 2017 QTL for maize midparent heterosis in the heterotic pattern american dent × European flint under corn borer pressure. Front. Plant Sci. 8: 573. https://doi.org/10.3389/fpls.2017.00573

Schön C. C., B. S. Dhillon, F. F. Utz, and A. E. Melchinger, 2010 High congruency of QTL positions for heterosis of grain yield in three crosses of maize. Theor. Appl. Genet. 120:

321–332. https://doi.org/10.1007/s00122-009-1209-9

Shull, G. H., 1908 The composition of a field of maize. J. Hered. 4: 296–301. https://doi.org/10.1093/jhered/os-4.1.296

Shull, G. H., 1914 Duplicate genes for capsule form in *Bursa pastoris* Zeitscher. Induktica Abstammu Vererbunglehra 12: 97–149.

Smith, O. S. C., J. S. C. Smith, S. L. Bowen, R. A. Tenborg, and S. J. Wall, 1990 Similarities among a group of elite maize inbreds as measured by pedigree, F1 grain yield, grain yield, heterosis, and RFLPs. Theor. Appl. Genet. 80: 833–840. https://doi.org/10.1007/BF00224201

Sprague, G. F., and L. A. Tatum, 1942 General vs. specific combining ability in single crosses of corn. Agron. J. 34: 923–932. https://doi.org/10.2134/agronj1942.00021962003400100008x

Stranger, B. E., E. A. Stahl, and T. Raj, 2011 Progress and promise of genome-wide association studies for human complex trait genetics. Genetics 187: 367–383. https://doi.org/10.1534/genetics.110.120907

Thavamanikumar, S., R. J. Arnold, J. Luo, and B. R. Thumma, 2020 Genomic studies reveal substantial dominant effects and improved genomic predictions in an open-Pollinated breeding population of eucalyptus pellita. G3 Genes, Genomes, Genet. 10: 3751–3763. https://doi.org/10.1534/g3.120.401601

Varona, L., A. Legarra, M. A. Toro, and Z. G. Vitezica, 2018a Non-additive effects in genomic selection. Front. Genet. 9: 1–12. https://doi.org/10.3389/fgene.2018.00078

Varona, L., A. Legarra, W. Herring, and Z. G. Vitezica, 2018b Genomic selection models for directional dominance: an example for litter size in pigs. Genet. Sel. Evol. 50: 1. https://doi.org/10.1186/s12711-018-0374-1

Vasseur, F., L. Fouqueau, D. De Vienne, T. Nidelet, C. Violle, *et al.*, 2019 Nonlinear phenotypic variation uncovers the emergence of heterosis in *Arabidopsis thaliana*. PLoS Biol. 17: 1–26. https://doi.org/10.1371/journal.pbio.3000214

1074 Vitezica, Z. G., L. Varona, and A. Legarra, 2013 On the additive and dominant variance and
1075 covariance of individuals within the genomic selection scope. Genetics 195: 1223–1230.
1076 https://doi.org/10.1534/genetics.113.155176

1077 Vitezica, Z. G., A. Legarra, M. A. Toro, and L. Varona, 2017 Orthogonal estimates of
1078 variances for additive, dominance, and epistatic effects in populations. Genetics 206:
1079 1297–1307. https://doi.org/10.1534/genetics.116.199406

1080 Vitezica, Z. G., A. Reverter, W. Herring, and A. Legarra, 2018 Dominance and epistatic
1081 genetic variances for litter size in pigs using genomic models. Genet. Sel. Evol.
1082 https://doi.org/10.1186/s12711-018-0437-3

1083 Voss-Fels, K. P., M. Cooper, and B. J. Hayes, 2019 Accelerating crop genetic gains with
1084 genomic selection. Theor. Appl. Genet. 132: 669–686. https://doi.org/10.1007/s00122-
1085 018-3270-8

1086 Wade, M. J., R. G. Winther, A. F. Agrawal, and C. J. Goodnight, 2001 Alternative definitions
1087 of epistasis: Dependence and interaction. Trends Ecol. Evol. 16: 498–504.
1088 https://doi.org/10.1016/S0169-5347(01)02213-3

1089 Waller, D. M., 2021 Addressing Darwin's dilemma: Can pseudo-overdominance explain
1090 persistent inbreeding depression and load? Evolution (N. Y). 75: 779–793.
1091 https://doi.org/10.1111/evo.14189

1092 Waser, N. M., and M. V. Price, 1994 Crossing-distance effects in *Delphinium nelsonii* :
1093 outbreeding and inbreeding depression in progeny fitness. Evolution. 48: 842–852.
1094 https://doi.org/10.1111/j.1558-5646.1994.tb01366.x

1095 Xiang, T., O. F. Christensen, Z. G. Vitezica, and A. Legarra, 2016 Genomic evaluation by
1096 including dominance effects and inbreeding depression for purebred and crossbred
1097 performance with an application in pigs. Genet. Sel. Evol. 48: 92.
1098 https://doi.org/10.1186/s12711-016-0271-4

1099 Yadav S., X. Wei, P. Joyce, F. Atkin, E. Deomano, et al., 2021 Improved genomic prediction
1100 of clonal performance in sugarcane by exploiting non - additive genetic effects. Theor.

Appl. Genet. 134: 2235–2252. https://doi.org/10.1007/s00122-021-03822-1

Zhang, K., J. Tian, L. Zhao, and S. Wang, 2008 Mapping QTLs with epistatic effects and QTL × environment interactions for plant height using a doubled haploid population in cultivated wheat. J. Genet. Genomics 35: 119–127. https://doi.org/10.1016/S1673-8527(08)60017-X

Zhang, H., L. Yin, M. Wang, X. Yuan, and X. Liu, 2019 Factors affecting the accuracy of genomic selection for agricultural economic traits in maize, cattle, and pig populations. Front. Genet. 10: 189. https://doi.org/10.3389/fgene.2019.00189

1111 TABLES

1112

1113

1114 **Table 1.** Mean trait values and heritabilities per environment in Het2 and iF2 populations.
1115 Heritabilities are shown between bracket. FLO, female and male flowering time; HT, plant
1116 height; GM, grain moisture; GY, grain yield.

Population	Environment	GY [q/ha]	FLO [f, GDD]	FLO [m, days]	HT [cm]	GM [%]
Het2	jar16	69.5 (0.75)		204.0 (0.97)	175.5 (0.90)	30.3 (0.89)
	aub17	93.3 (0.67)		192.8 (0.92)	217.5 (0.87)	28.5 (0.92)
	sou17	73.3 (0.76)		184.8 (0.79)	228.4 (0.81)	29.9 (0.85)
	smh16	76.1 (0.61)		196.2 (0.90)	216.8 (0.66)	21.2 (0.61)
	smh17	71.4 (0.82)		191.5 (0.94)	216.2 (0.87)	19.9 (0.95)
IF2	mln11	40.9 (0.69)	792.0 (0.76)		123.2 (0.83)	25.5 (0.93)
	smh10	44.2 (0.75)	983.9 (0.69)		180.0 (0.83)	28.9 (0.74)
	smh11	43.1 (0.90)	715.3 (0.88)		148.7 (0.88)	26.1 (0.84)

1117

1118

Table 2. Variance decomposition and BIC values (smaller is better, the lowest value in bold) for each trait and population considering mean phenotypic values across environments $\overline{Y_{h..}}$. Traits: FLO, flowering time; HT, plant height; GM, grain moisture; GY, grain yield. % hom is the percentage of homozygosity. Variance components: A, additive; D, dominant; AA, additive by additive; AD, additive by dominant; DD, dominant by dominant. Submodels A, additive effect only; A_Inb, additive and inbreeding effects; AD, additive and dominant effects; AD_Inb, additive, dominant and inbreeding effects; ADI_Inb, additive, dominant, epistasis and inbreeding effects.

Population	Trait	*M2* and submodels	BIC	Variance components						Broad sense heritability	Inb reg coef (%hom)	Wald test (p-value)	Mean pred. ability (SD)
				A	D	AA	AD	DD	Error				
Het2	FLO	A	1284.7	6.12					0.93	0.89	-	-	0.742 (0.017)
		A_Inb	**1282.9**	6.21	-	-	-	-	0.84	0.90	16.71	5.65E-03	0.748 (0.017)
		AD	1288.9	6.14	0.43	-	-	-	0.49	0.95	-	-	0.741 (0.017)
		AD_Inb	1288.6	6.2	0.02	-	-	-	0.81	0.91	16.66	6.02E-03	0.747 (0.017)
		ADI_Inb	1305.7	6.15	0.03	0.1	0.21	0.24	0.31	0.97	16.52	7.29E-03	0.747 (0.017)
	HT	A	2492.7	421.7	-	-	-	-	48.42	0.92	-	-	0.653 (0.023)
		A_Inb	2495.2	421.78	-	-	-	-	47.09	0.92	-83.56	7.86E-02	0.655 (0.023)
		AD	**2483.6**	423.12	41.51	-	-	-	5.68	0.99	-	-	0.653 (0.023)
		AD_Inb	2487.3	422.41	40.91	-	-	-	5.29	0.99	-89.32	8.41E-02	0.655 (0.023)
		ADI_Inb	2517.7	364.16	1.34	87.8	0	0	0	1.00	-83.32	9.08E-02	0.656 (0.023)
	GM	A	**1265.5**	5.21	-	-	-	-	1.06	0.86	-	-	0.734 (0.017)
		A_Inb	1270.8	5.24	-	-	-	-	1.06	0.86	3.78	5.54E-01	0.731 (0.017)
		AD	1271.2	5.21	0.09	-	-	-	0.98	0.87	-	-	0.733 (0.017)
		AD_Inb	1276.5	5.25	0.23	-	-	-	0.83	0.89	3.61	5.57E-01	0.730 (0.017)
		ADI_Inb	1293.6	5.17	0.1	0.15	0.25	0.26	0.37	0.95	3.82	5.55E-01	0.730 (0.017)
	GY	A	2235.7	120.19	-	-	-	-	43.13	0.78	-	-	0.589 (0.027)
		A_Inb	**2229.6**	117.44	-	-	-	-	40.8	0.79	-117.32	5.49E-04	0.608 (0.026)
		AD	2234.7	115.38	36.24	-	-	-	8.43	0.96	-	-	0.592 (0.026)

		AD_Inb	2235.3	117.44	0.04	-	-	-	40.77	0.79	-117.31	5.51E-04	0.605 (0.026)
		ADI_Inb	2252.3	117.44	0.09	0.01	0.03	0.01	40.67	0.79	-117.29	5.54E-04	0.605 (0.026)
IF2	FLO	A	2356.5	706.42	-	-	-	-	163.7	0.85	-	-	0.744 (0.017)
		A_Inb	**2348.5**	716.35	-	-	-	-	149.66	0.86	62.17	1.85E-04	0.761 (0.016)
		AD	2361.8	698.66	7.05	-	-	-	157.73	0.85	-	-	0.741 (0.017)
		AD_Inb	2354	716.36	0	-	-	-	149.66	0.86	62.17	1.85E-04	0.760 (0.016)
		ADI_Inb	2370.2	606.95	0	53.29	0	6.91	122.32	0.88	61.64	4.19E-04	0.761 (0.016)
	HT	A	2028.4	228	-	-	-	-	44.05	0.87	-	-	0.789 (0.015)
		A_Inb	**2021.3**	232.27	-	-	-	-	40.24	0.88	-31.74	2.98E-04	0.800 (0.014)
		AD	2032.3	229.56	6.48	-	-	-	37.39	0.89	-	-	0.787 (0.016)
		AD_Inb	2026.8	232.09	0.93	-	-	-	39.35	0.88	-32.17	4.04E-04	0.797 (0.015)
		ADI_Inb	2037.8	184.86	0	47.81	0.01	0	14.51	0.95	-30.9	1.28E-03	0.804 (0.015)
	GM	A	**1089.1**	7.14	-	-	-	-	1.2	0.88	-	-	0.756 (0.015)
		A_Inb	1094.7	7.13	-	-	-	-	1.21	0.88	-0.24	8.73E-01	0.752 (0.016)
		AD	1092.5	7.22	0.19	-	-	-	1	0.91	-	-	0.756 (0.015)
		AD_Inb	1098.1	7.23	0.22	-	-	-	0.98	0.91	0.02	9.90E-01	0.753 (0.016)
		ADI_Inb	1109.3	4.84	0.09	1.37	0	0.09	0.41	0.95	0.01	9.94E-01	0.760 (0.015)
	GY	A	1705.2	30.72	-	-	-	-	20.11	0.66	-	-	0.557 (0.023)
		A_Inb	**1672.9**	29.45	-	-	-	-	16.79	0.69	-32.46	1.95E-10	0.635 (0.021)
		AD	1700.1	30.25	6.77	-	-	-	13.5	0.78	-	-	0.592 (0.024)
		AD_Inb	1677.9	29.03	1.92	-	-	-	15.06	0.73	-35.18	7.12E-09	0.633 (0.021)
		ADI_Inb	1689.4	20.02	0.51	10.13	0	0	10.74	0.79	-34.16	1.49E-09	0.648 (0.021)

1132 FIGURE LEGENDS

1133

1134 **Figure 1**: Experimental design used for cross-validations in each scenario. 'h', 'e' and 'r'
1135 describe the levels of factors 'hybrid', 'environment' and 'plot' (see Methods). For scenario
1136 « GxE_new_env », $\widetilde{Y_{her}}$ values are used for training and $\overline{Y_{he.}}$ values are used for validation.
1137 $\overline{Y_{he.}}$ are mean values per environment and $\overline{Y_{h..}}$ are mean values across environments. For the
1138 « Within-env. » scenario, environment A serves as an example.

1139 **Figure 2**. Variance decomposition within each environment using model M3 and
1140 corresponding submodels in A, population Het2 and B, population iF2. FLO, flowering time;
1141 HT, plant height; GM, grain moisture; GY, grain yield. Each column represents a single
1142 environment.

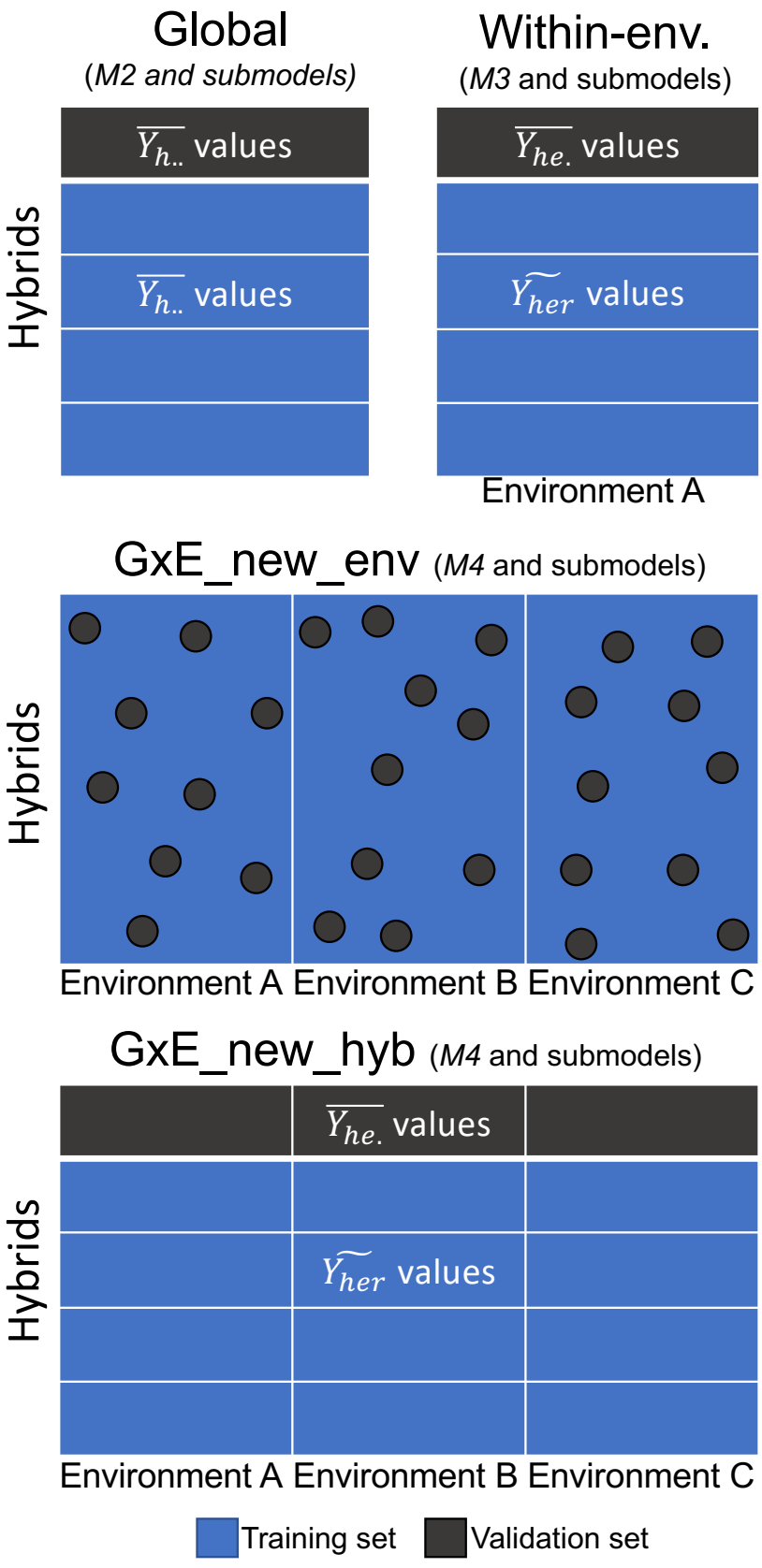
1143 **Figure 3**. Variance decomposition across environments for grain yield using ADI×Ecom_Inb
1144 and ADI×Espec_Inb models and submodels derived from model *M4* in populations Het2
1145 (A) and iF2 (B). The error variance terms $\sigma^2_{\varepsilon(e)}$ are environment specific and represented by
1146 their average value. For ADI×Espec_Inb model and submodels G×E variance terms
1147 $\sigma^2_{AE(e)} \ldots \sigma^2_{DDE(e)}$ are environment specific and represented by their average value. In B, and
1148 represented by their average value.

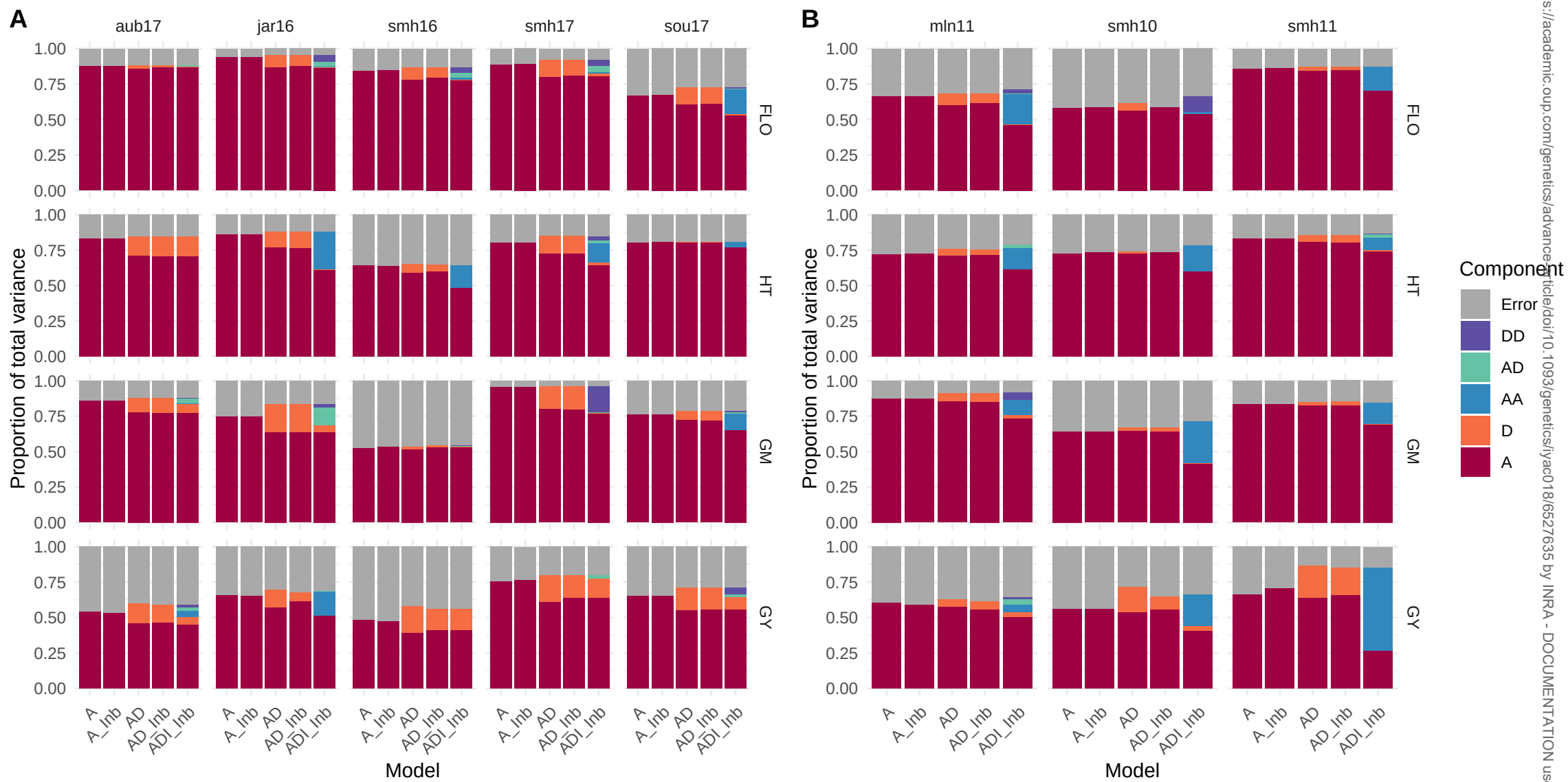
1149 **Figure 4.** Boxplots representing the distribution of predictive ability in scenario « Global »
1150 considering model *M2* and submodels and 100 cross-validations in Het2 (A) and in iF2 (B)
1151 populations. FLO, flowering time; HT, plant height; GM, grain moisture; GY, grain yield.

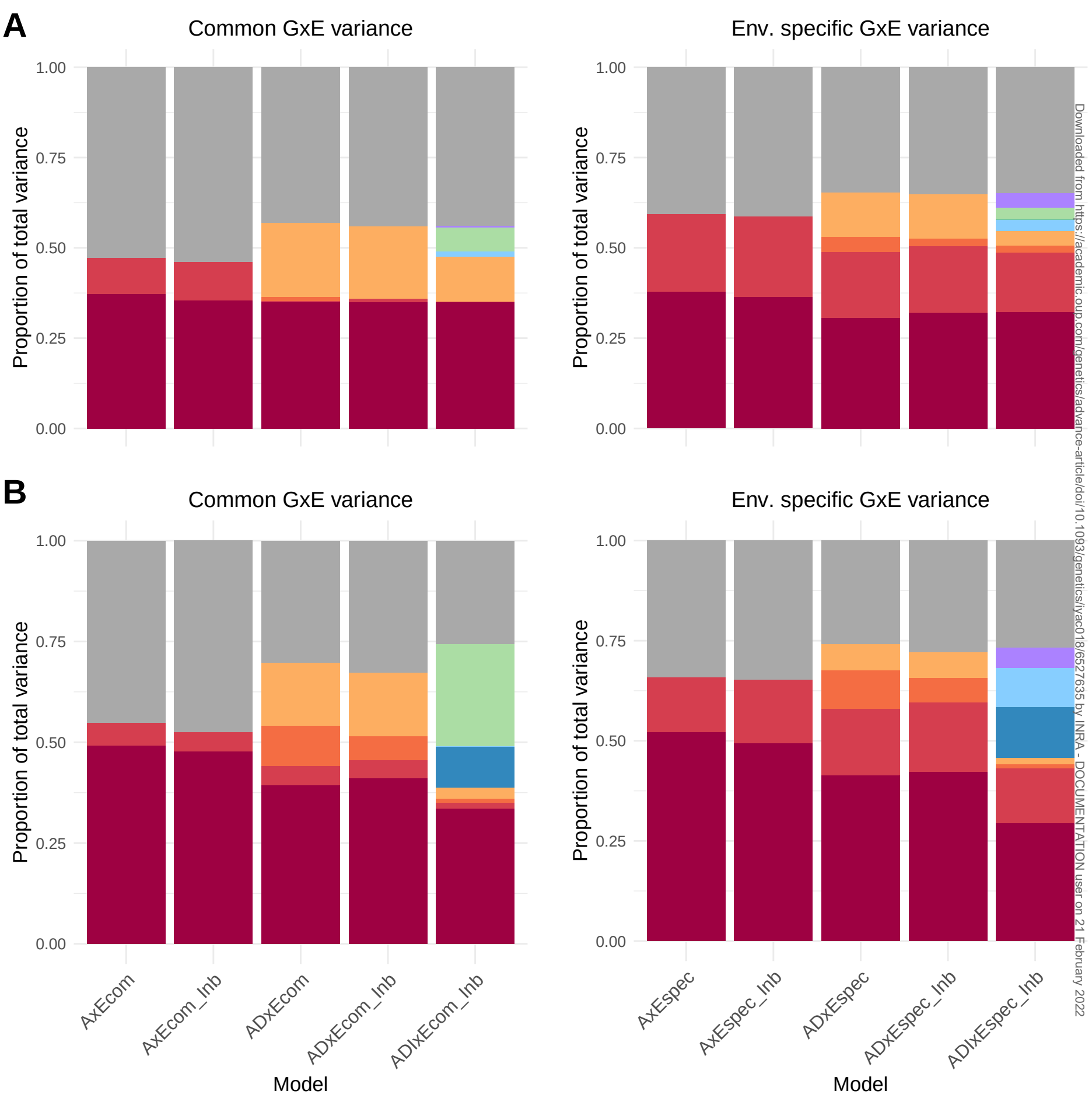
1152 **Figure 5.** Boxplots representing the distribution of predictive ability in scenario « Within
1153 environment » considering model *M3* and submodels and 100 cross-validations in Het2 (A)
1154 and in iF2 (B) populations. FLO, flowering time; HT, plant height; GM, grain moisture; GY,
1155 grain yield. Each column represents a single environment.

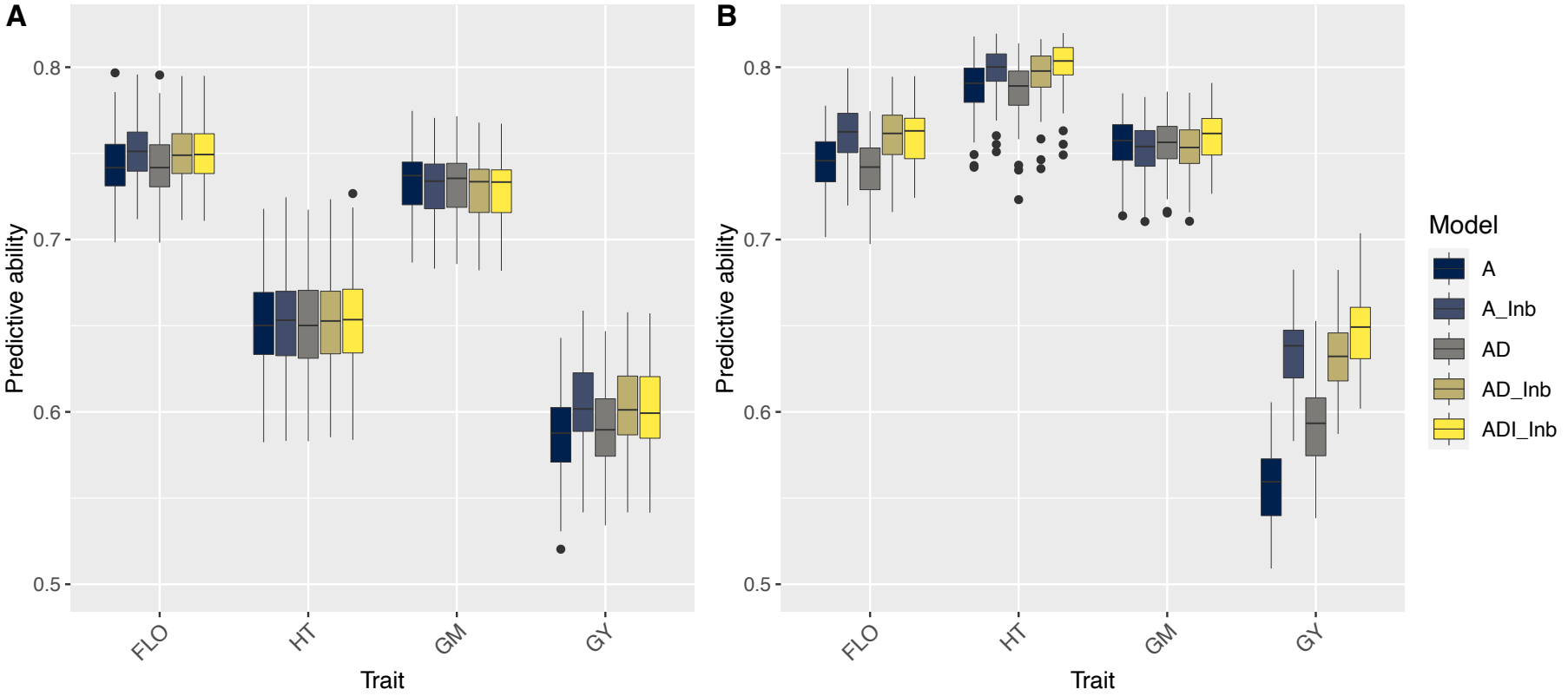
1156 **Figure 6.** Boxplot representing the distribution of predictive ability for grain yield in iF2
1157 population using G×E modelling with model *M4* and submodels. A, scenario
1158 « G×E_new_env », 100 cross-validations; B, scenario « G×E_new_hyb », 100 cross-

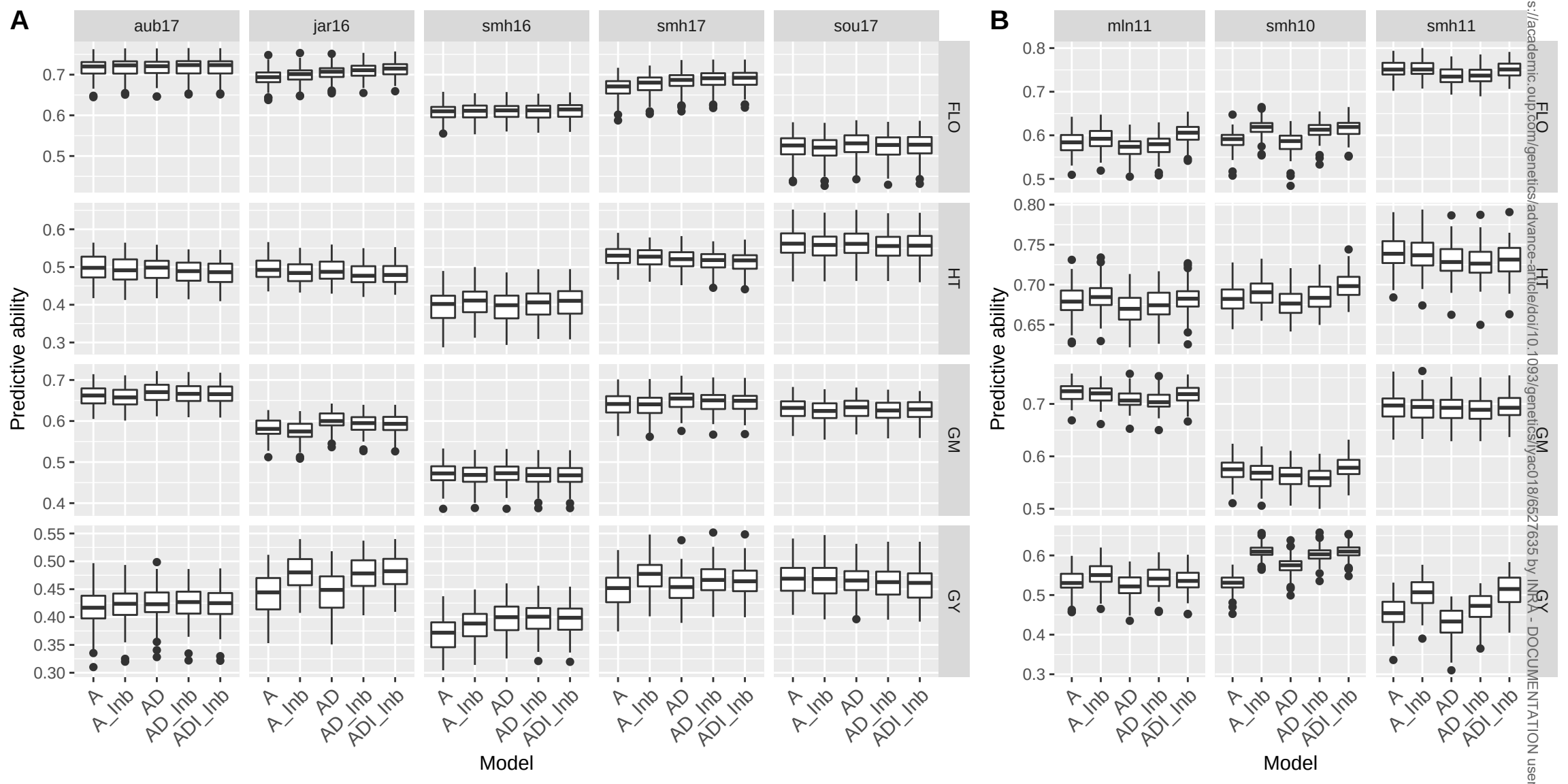
validations. Each column represents a single environment. The color legend indicates whether results are obtained with ADI×Ecom_Inb models and submodels (“Common”) or with ADI×Espe_Inb models and submodels (“Env. specific”).










A
B
aub17
jar16
smh16
smh17
sou17
mln11
smh10
smh11
FLO
HT
GM
GY
Predictive ability
Model
A
A_Inb
AD
AD_Inb
ADI_Inb