植物学报 *Chinese Bulletin of Botany* 2024, **59** (6): 1041–1053, www.chinbullbotany.com
doi: 10.11983/CBB24087 cstr: 32102.14.CBB24087

·研究论文·

玉米自然群体自交系农艺性状的多环境全基因组预测初探

李园[1], 范开建[1], 安泰[2], 李聪[3], 蒋俊霞[1], 牛皓[1], 曾伟伟[2]
衡燕芳[1], 李虎[1], 付俊杰[1], 李慧慧[1], 黎亮[1*]

[1]中国农业科学院作物科学研究所, 作物分子育种国家工程研究中心, 作物基因资源与育种全国重点实验室, 北京 100081
[2]中国农业大学农学与生物技术学院, 北京 100193; [3]黑龙江八一农垦大学农学院, 大庆 163000

摘要 多环境田间测试是选育高产稳产品种的重要途径, 但因其成本高逐渐成为植物育种的瓶颈问题。将稀疏测试与全基因组预测方法相结合可实现对未测表型的预测, 进而减少田间测试的工作量和成本。利用244份玉米(*Zea mays*)自然群体自交系在两年(2022年和2023年)两点(北京顺义和黑龙江密山) 4个环境下, 针对散粉期、株高、穗位高、穗长、穗行数和行粒数6个代表性农艺性状开展研究, 比较了4种模型(Single、Across、M×E和R-norm)、2种训练群体组成方案(CV1和CV2)以及3种训练集抽样比例(0.5、0.7和0.9)对预测精度的影响。结果表明, 上述6个农艺性状的平均预测精度分别为0.67、0.58、0.50、0.33、0.33和0.48; Single模型、Across模型、M×E模型和R-norm模型的平均预测精度分别为0.36、0.52、0.53和0.53; 其中CV1各模型在不同性状中的预测精度范围在0.19–0.65之间, CV2预测精度范围在0.47–0.89之间; 不同抽样比例比较显示, 不同模型中训练集比例的提升对6个性状的预测精度提升有限, 最大提升幅度仅为0.05。综上表明, 在进行多环境预测时, 利用CV2训练群体组成方案并在预测模型中纳入多个环境下的表型数据可提升预测精度。

关键词 玉米, 全基因组预测, 多环境预测, 训练集优化

李园, 范开建, 安泰, 李聪, 蒋俊霞, 牛皓, 曾伟伟, 衡燕芳, 李虎, 付俊杰, 李慧慧, 黎亮 (2024). 玉米自然群体自交系农艺性状的多环境全基因组预测初探. 植物学报 **59**, 1041–1053.

玉米(*Zea mays*)是世界上重要的粮食作物之一, 也是重要的饲料和生物能源的主要原料。目前我国玉米种植面积高达4×10^7 hm^2, 已成为第一大粮食作物。但随着对其需求日益增长, 如何进一步提高玉米单产是育种家面临的重大难题。产量性状是综合性状, 受其它相关性状(如生育期、株高、穗长和穗行数)的影响。产量及其相关性状是典型的受微效多基因控制的数量性状, 其表型受基因型与环境互作(genotype-by-environment, G×E)效应的影响较大, 增加了遗传研究和育种选择的复杂性(Carena et al., 2010)。因此, 在育种中需要在多个环境下进行评价并综合考虑不同的产量相关性状, 协同改良以达到较为理想的产量水平。

在玉米种质资源评价及品种培育过程中, 不同的材料往往需要在几十甚至几百个环境下进行测试, 以评价其对不同环境的适应性和产量表现。然而, 表型评价的成本越来越高, 田间多环境测试面临的困难和挑战随之加大。此外, 近几年双单倍体(doubled haploid, DH)技术快速发展, 每年生产的DH系数量急剧增长, 可组配的杂交组合数量更是呈指数增长, 因此传统育种中依赖于多个测验种杂交后在多个环境下进行田间测试的DH系评价策略难以持续, 这已成为玉米DH系育种的瓶颈(Fu et al., 2022)。因此, 对不同环境下的表型预测成为育种中的新需求。全基因组预测(genomic prediction, GP)亦称全基因组选择(genomic selection, GS), 可利用覆盖全基因组的高密度分子标记和已获得的表型数据构建模型, 预测只有基因型个体的育种值(Meuwissen et al., 2001)。利用该方法可进行个体的早期预测和选择, 有效缩短育种周期, 提高遗传增益。近年来, 随着基因组测序和

收稿日期: 2024-06-06; 接受日期: 2024-10-14
基金项目: 国家自然科学基金(No.32272190)和中国农业科学院创新工程(2024)
* 通讯作者。E-mail: liliang05@caas.cn

芯片检测技术逐渐成熟，高密度分子标记的成本显著降低，加之预测模型的不断优化，GP已成为高效且精准的育种技术之一，越来越多应用于动植物育种(Hayes et al., 2009; Jannink et al., 2010; Crossa et al., 2010)。

GP方法将群体分为训练集(training populations, TRN)和测试集(testing populations, TST)，其中TRN的构成是影响预测准确性的重要因素，包括样本量大小和亲缘关系等(Xu et al., 2019)。在多环境试验中，TRN的构成除需要考虑亲缘关系，还要考虑不同环境的抽样。在GP中，常通过交叉验证(cross-validation, CV)来评估预测精度。Burgueño等(2012)针对不同基因型在多个环境下测试如何进行分配的问题，利用2种不同的CV方式模拟育种者可能面临的2种真实场景，即评估模型用于预测完全未经测试的基因型在不同环境下的表现(CV1)和评估在某些环境中已测试但未在其它环境中测试的预测表现(CV2)。Burgueño等(2012)利用小麦(*Triticum aestivum*)多环境试验数据比较CV1和CV2两种不同CV方案的预测精度，发现CV2的预测精度高于CV1。Jarquin等(2020)在控制测试总量不变的前提下，通过设计不同环境间基因型完全不重叠和基因型完全重叠2种情况，结果表明TRN设计为CV2时预测结果较好。此外，在设计GP试验时，TRN与TST的比例也影响预测精度。Luo等(2023)通过对285份玉米杂交种设置不同的TRN与TST比例，发现比例为1:1时可实现最优预测。TRN与TST的比例也受群体亲缘关系影响，当对全同胞家系进行预测时，较少比例的训练群体就能对大多数性状和家系实现合理的预测精度(Zhu et al., 2021)。

统计模型是影响预测精度的重要因素之一。统计模型主要包括参数、半参数和非参数三大类(Alemu et al., 2024)。其中参数方法中基因组最佳线性无偏预测(genomic best linear unbiased prediction, GBLUP)用途最广。GBLUP法假设所有标记均遵循相同的遗传变异，通过构建基因组关系矩阵G来预测表型，进而直接估计个体的育种值(VanRaden, 2008)。该方法具有高效稳健的优点，是目前应用最广泛的方法。与动物育种不同，植物育种包括广泛的多点多年田间试验数据，基因组预测模型应考虑不同环境下的G×E效应。2012年，Burgueño等(2012)首次提出在GP模型中加入G×E效应，该方法可以帮助育种者预测某个环境下完全未经测试的杂交组合的表现或者完全未知环境下杂交组合的表现。Lopez-Cruz等(2015)通过对所有标记和环境(M×E)之间的相互作用进行建模，以区分跨环境共同影响的标记和环境特定影响的标记，显著提升了预测精度。近年来，研究者通过在GP模型中结合G×E效应来提高作物表型的预测能力并取得了重要进展(Cuevas et al., 2016; Ferrão et al., 2017; Sousa et al., 2017; Roorkiwal et al., 2018)，这些研究结果对于育种中资源的合理分配至关重要。

目前，利用多环境表型数据进行预测已成为全基因组预测的热点和重要研究方向。但尚未见利用国内不同环境下表型数据开展多环境全基因组预测的研究报道。因此，本研究利用自然群体自交系在4个环境下的6个农艺性状，即散粉期、株高、穗位高、穗长、穗行数和行粒数的表型数据，探讨环境间抽样方式、预测模型及抽样比例对多环境测试中表型预测精度的影响。

1 材料与方法

1.1 试验材料与设计

试验材料主要包括244份来源广泛的玉米(*Zea mays* L.)自然群体自交系，该群体为结合已报道的自然群体自交系(Yang et al., 2014; Wang et al., 2020)，在其基础上进行收集整理的适宜温带地区种植的自交系，主要包含9个亚群，各亚群中的材料数较为均衡，具有很好的代表性。2022年和2023年，试验材料分别种植于中国农业科学院作物科学研究所北京顺义基地(116.28°E, 40.00°N)和黑龙江省八一农垦大学密山基地(45.54°E, 131.87°N)，每个年份和地点的组合为一个环境，2022年北京、2022年密山、2023年北京和2023年密山分别简写为22BJ、22MS、23BJ和23MS。田间试验设计采用扩增α-设计，包含2次重复，每20个材料为1个区组，每个区组包含2个对照(B73和Mo17)，行长为1.5 m，株距为0.2 m，行距为0.5 m。田间管理同大田常规管理。

1.2 表型调查

在4个环境下(22BJ、22MS、23BJ和23MS)分别调查6个农艺性状，包括花期性状：散粉期(days to an-

thesis, DTA); 株型性状: 株高(plant height, PH)和穗位高(ear height, EH); 产量性状: 穗长(ear length, EL)、行粒数(kernel number per row, KNR)和穗行数(ear row number, ERN)。调查标准见表1。

1.3 表型数据处理与统计分析

4个环境下6个性状的表型数据采用R包lme4进行线性混合模型拟合, 得到最佳线性无偏预测估计(best linear unbiased estimation, BLUE)值, 公式如下:

$$y_{im}=\mu+G_i+R_m+B_nR_m+\varepsilon_{im}$$

其中, y_{im}是i^{th}个基因型在m^{th}个重复下的表型, μ是群体均值, G_i是i^{th}个个体的遗传效应, R_m为m^{th}个重复效应, B_n是m^{th}下重复内n^{th}个区组效应, ε_{im}为误差, 服从正态分布ε_{im}–$N(0, \sigma_m^2)$, σ_m^2为m^{th}下重复内的误差方差。

1.4 方差分析和遗传力计算

方差分析公式如下:

$$y_{ijb}=\mu_{ij}+B_{b/j}+G_i+E_j+GE_{ij}+\varepsilon_{ijk}$$

y_{ijb}表示第i^{th}个基因型在第j^{th}个环境下的第b^{th}个区组下的表型值, 服从正态分布y_{ijb}–$N(\mu_{ij}, \sigma_\varepsilon^2)$, μ_{ij}表示群体内所有个体的表型平均值, $E_{b/j}$表示第j^{th}个环境下第b^{th}个区组的效应, G_i为第i^{th}个基因型的遗传效应, E_j为j^{th}个环境的效应, GE_{ij}表示第i^{th}个基因型与第j^{th}个环境之间的互作效应, ε_{ijk}为随机误差, 服从正态分布ε_{ijk}–$N(0, \sigma_\varepsilon^2)$。

遗传方差: $\sigma_g^2 \triangleq \frac{1}{g-1}\sum_i g_i^2$

互作方差: $\sigma_{ge}^2 \triangleq \frac{1}{(g-1)(e-1)}\sum_{i,j} ge_{ij}^2$

环境和重复平均数广义遗传力(H^2)计算公式如下:

$$H^2=\frac{\sigma_g^2}{\sigma_g^2+\frac{1}{e}\sigma_{ge}^2+\frac{1}{re}\sigma_\varepsilon^2}$$

遗传力估值的近似标准误(SE(H^2))计算公式如下:

$$\mathrm{SE}(H^2)=\frac{\mathrm{SE}(\sigma_g^2)}{\sigma_g^2+\frac{1}{e}\sigma_{ge}^2+\frac{1}{re}\sigma_\varepsilon^2}$$

1.5 基因型数据

试验所用基因型数据为全基因组重测序数据, 利用华大T7测序平台进行15×深度的测序, 以B73 V5为参考基因组, 经过严格的筛选共获得3×10^7个SNP位点。利用plink软件对基因型数据进行质控, 质控条件为最小等位基因频率(MAF)<0.5, 基因型缺失率>0.2, 质控后使用Beagle软件进行填充, 最终获得301 838个高质量SNP用于预测分析。

1.6 统计模型

本研究主要采用4种统计模型进行多环境预测, 4种模型均采用DIC准则进行选择(Pérez and De Los Campos, 2014):

(1) 单环境模型(Single) (Vanraden, 2008)

该模型表示为:

$$y_k = \mu_k + X_k\beta_k + \varepsilon_k$$

$y_k=[y_{1k}, \cdots, y_{nk}]'$, 该模型是在第$k$个环境($k$=1, 2, …, s)中$n$个个体的表型向量使用线性模型对$p$个标记进行回归得到。$X_k=\{x_{ijk}\}$是在第$k$个环境中可用的中心化和标准化的$n \times p$标记矩阵, $\beta_k=[\beta_{1k}, \cdots, \beta_{pk}]$是一个p维的标记效应向量, $\varepsilon_k=[\varepsilon_{1k}, \cdots, \varepsilon_{nk}]'$是一个$n$维的残差向量, 1=[1, …, 1]′是一个$n$维的全1向量。其中

表1 244份玉米自交系6个性状的一般信息

Table 1 General information of six maize agronomic traits for 244 inbred lines

Trait	Abbreviate	Unit	Description
Days to anthesis	DTA	Days	Recorded the number of days from the planting day to anthesis data when 50% of the plant anthers in the plot were extruded to 1/2 length of the main tassel spindle
Plant height	PH	cm	Measured the height of the stem from the ground to the top of the tassel of 3–5 plants
Ear height	EH	cm	Measured the height of the stem from the ground to the base of the ear of 3–5 plants
Ear length	EL	cm	Measured the length of 3–5 ears
Ear row number	ERN	Count	Counted the number of ear row of 3–5 ears
Kernel number per row	KNR	Count	Counted the number of kernels per row of 3–5 ears

i=1, 2,…, n个个体, μ_k为截距项。

令$g_k=X_k\beta_k$，其中$g_k=[g_{1k}, \cdots, g_{nk}]'$是随机效应，假设$g_k$–N$(0, \sigma^2_{gk}G_k)$，$G_k=\frac{X_kX_k'}{p}$，其GBLUP形式可表示为:

$$y_k=\mu_k1+g_k+\varepsilon_k$$

(2) Across-Environment模型(Lopez-Cruz et al., 2015)

该模型假设标记在不同环境中的影响一致，模型描述见Lopez-Cruz等(2015)，该模型GBLUP形式为:

$$\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}=\begin{bmatrix}\mu_1 1\\\mu_2 1\\\mu_3 1\end{bmatrix}+\begin{bmatrix}g_1\\g_2\\g_3\end{bmatrix}+\begin{bmatrix}\varepsilon_1\\\varepsilon_2\\\varepsilon_3\end{bmatrix}$$

(3) M×E模型(Lopez-Cruz et al., 2015)

该模型使用M×E方法对G×E互作进行建模，模型公式见Lopez-Cruz等(2015)的描述，每个标记对每个环境的影响被分解为所有环境共有的效应和每个环境下特有的效应，该模型GBLUP形式为:

$$\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}=\begin{bmatrix}\mu_1 1\\\mu_2 1\\\mu_3 1\end{bmatrix}+\begin{bmatrix}g_{0,1}\\g_{0,2}\\g_{0,3}\end{bmatrix}+\begin{bmatrix}g_{1,1}\\g_{1,2}\\g_{1,3}\end{bmatrix}+\begin{bmatrix}\varepsilon_1\\\varepsilon_2\\\varepsilon_3\end{bmatrix}$$

(4) R-norm (reaction norm)模型(Jarquín et al., 2014)

该模型假设环境(e_k)为随机效应，模型描述见Jarquín等(2014)，R-norm模型通过引入交互项$g_{ik}e_{ik}$加入GBLUP模型中，GBLUP模型形式为:

$$\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}=\begin{bmatrix}\mu_1 1\\\mu_2 1\\\mu_3 1\end{bmatrix}+\begin{bmatrix}g_1\\g_2\\g_3\end{bmatrix}+\begin{bmatrix}g_1e_1\\g_2e_2\\g_3e_3\end{bmatrix}+\begin{bmatrix}\varepsilon_1\\\varepsilon_2\\\varepsilon_3\end{bmatrix}$$

1.7 多环境训练集组成试验方案设计

将数据集随机划分为TRN和TST，研究不同训练群体组成对预测结果的影响，评估不同模型的预测能力。(1) 设置2种不同的CV方案，CV1和CV2 (图1)；使用TRN占总样本量比例为0.5的数据集进行预测；对于CV1，50%的自交系在所有环境下都有表型，用于构建训练模型预测剩余50%未经测试自交系在不同环境下的表型；对于CV2，总体缺失50%的表型数据，但是每个材料至少在一个环境下有表型数据，利用此数据集构建训练模型预测未测环境下的表型；(2) 不同模型比较：将M×E模型和R-norm模型与忽略G×E互作建模的Across模型以及在每个环境中拟合的单环境模型(Single)的预测精度进行比较；(3) 不同TRN比例比较：比较不同梯度TRN占总样本量比例(0.5、0.7和0.9)对预测精度的影响。

使用TRN数据集训练模型，通过计算不同环境中TST中预测值与真实值之间的Pearson相关系数评估预测准确性。每个模型进行50次交叉验证，参数nIter为30 000; burnIn为2 000。

	CV1			
	E1	E2	E3	E4
Line1				
Line2	NA	NA	NA	NA
Line3				
Line4				
Line5	NA	NA	NA	NA
Line6				
Line7				
Line8	NA	NA	NA	NA
Line9				
Line10	NA	NA	NA	NA

	CV2			
	E1	E2	E3	E4
Line1	NA			
Line2	NA	NA		
Line3			NA	
Line4		NA		
Line5				NA
Line6	NA			
Line7		NA		
Line8			NA	
Line9				NA
Line10	NA			NA

图1 两种交叉验证方案(CV1和CV2)

表格中E1–E4代表不同的环境，NA表示品种表型在该环境下未测试，黄色空格表示品种表型在该环境测试。

Figure 1 Two distinct cross-validation schemes (CV1 and CV2)

In the table, E1–E4 indicate different environments, NA indicate that the variety phenotype was untested in this environment, and the yellow space indicate that the variety phenotype tested in this environment.

2 结果与分析

2.1 表型数据分析

在22BJ、22MS、23BJ和23MS四个环境下，分别对散粉期、株高、穗位高、穗长、行粒数和穗行数6个性状进行描述性统计分析(表2)。结果表明，各性状在不同环境间的平均值和变异幅度均存在差异。2022年和2023年DTA，MS比BJ均晚20天左右；但是2022年PH和EH，BJ均高于MS，2023年MS均高于BJ；22BJ高于22MS，但是23BJ低于23MS；EL、KNR和ERN并无明显的年份或者地点间的显著差异。通过偏度、峰度和变异系数分析，发现不同性状的离散程度不同，大部分性状呈偏态分布(图2)。6个性状在基因

表2 六个农艺性状在4个环境下表型数据描述性统计分析

Table 2 Descriptive statistical analysis of six agronomic traits across four environments

Trait	Environment	Range	Means±SD	Skew	Kurt	Coefficient of variation (%)
DTA	22BJ	48.00–72.00	61.41±15.29	–0.56	0.11	0.25
	22MS	54.00–99.00	83.05±31.11	–1.32	2.55	0.37
	23BJ	53.00–82.00	67.88±17.28	–0.13	0.00	0.25
	23MS	69.00–104.00	88.23±12.59	–0.04	0.90	0.14
PH	22BJ	131.50–315.67	230.39±52.01	–0.33	0.17	0.23
	22MS	113.00–302.00	227.27±45.40	–0.28	0.13	0.20
	23BJ	90.33–257.33	183.66±39.61	–0.44	0.68	0.22
	23MS	134.00–323.67	233.58±54.45	–0.29	0.06	0.23
EH	22BJ	34.67–142.33	87.70±23.85	–0.07	–0.46	0.27
	22MS	17.33–168.00	74.66±24.73	0.28	0.36	0.33
	23BJ	20.00–125.33	69.68±19.79	–0.16	–0.28	0.28
	23MS	23.67–145.33	86.81±27.89	–0.07	–0.41	0.32
EL	22BJ	5.00–22.00	13.99±3.99	0.11	0.62	0.29
	22MS	6.70–20.80	14.54±4.03	0.16	0.15	0.28
	23BJ	5.00–22.33	13.23±4.22	0.25	0.16	0.32
	23MS	8.50–20.00	13.60±4.53	0.46	0.17	0.33
KNR	22BJ	2.00–44.67	21.93±8.32	–0.45	0.58	0.38
	22MS	10.00–41.00	24.36±7.90	–0.16	0.05	0.32
	23BJ	3.00–39.00	19.47±7.56	0.13	0.15	0.39
	23MS	10.00–39.33	24.48±8.68	–0.06	–0.22	0.35
ERN	22BJ	7.50–19.67	13.49±3.78	0.04	–0.41	0.28
	22MS	6.00–20.00	13.62±3.89	0.05	–0.10	0.29
	23BJ	8.00–20.67	12.88±4.01	0.43	–0.01	0.31
	23MS	7.33–20.00	13.86±4.48	–0.09	–0.50	0.32

DTA: 散粉期; PH: 株高; EH: 穗位高; EL: 穗长; KNR: 行粒数; ERN: 穗行数

DTA: Days to anthesis; PH: Plant height; EH: Ear height; EL: Ear length; KNR: Kernel number per row; ERN: Ear row number

型间和基因型与环境间的互作效应均存在极显著差异(表3)。不同性状的遗传力存在差异，其中株高的广义遗传力最高为0.76，KNR的广义遗传力最低为0.38。

2.2 CV1和CV2两种交叉验证方案对预测结果的影响

对于不同模型，在CV1和CV2两种TRN组成方案下，Single模型的预测精度最低，其它3种模型的预测精度相似(图3)。在CV1中，各模型在DTA、EH、PH、EL、KNR和ERN 6个性状中平均预测精度范围在0.19–0.65之间；在CV2中，除Single模型外，各模型在6个性状中的平均预测精度范围在0.47–0.89之间。除Single模型之外，其它3种模型都表现为CV2下的预测精度显著高于CV1。在所分析的6个性状中，CV2的预测精度均高于CV1，且不同性状CV2较CV1平均预测精度提升的幅度不一致。在CV1中，DTA的平均预测精度最高(0.68)，EL的平均预测精度最低(0.12)；在CV2中，PH的平均预测精度最高(0.89)，较CV1提升了0.57，除Single模型外，KNR的平均预测精度最低(0.44)，比CV1提升了0.23。对于4个环境，同一模型在分析DTA和KNR时，22BJ环境下的预测精度最高；在分析PH、EH、EL和ERN时，22MS环境下的预测精度均最高。

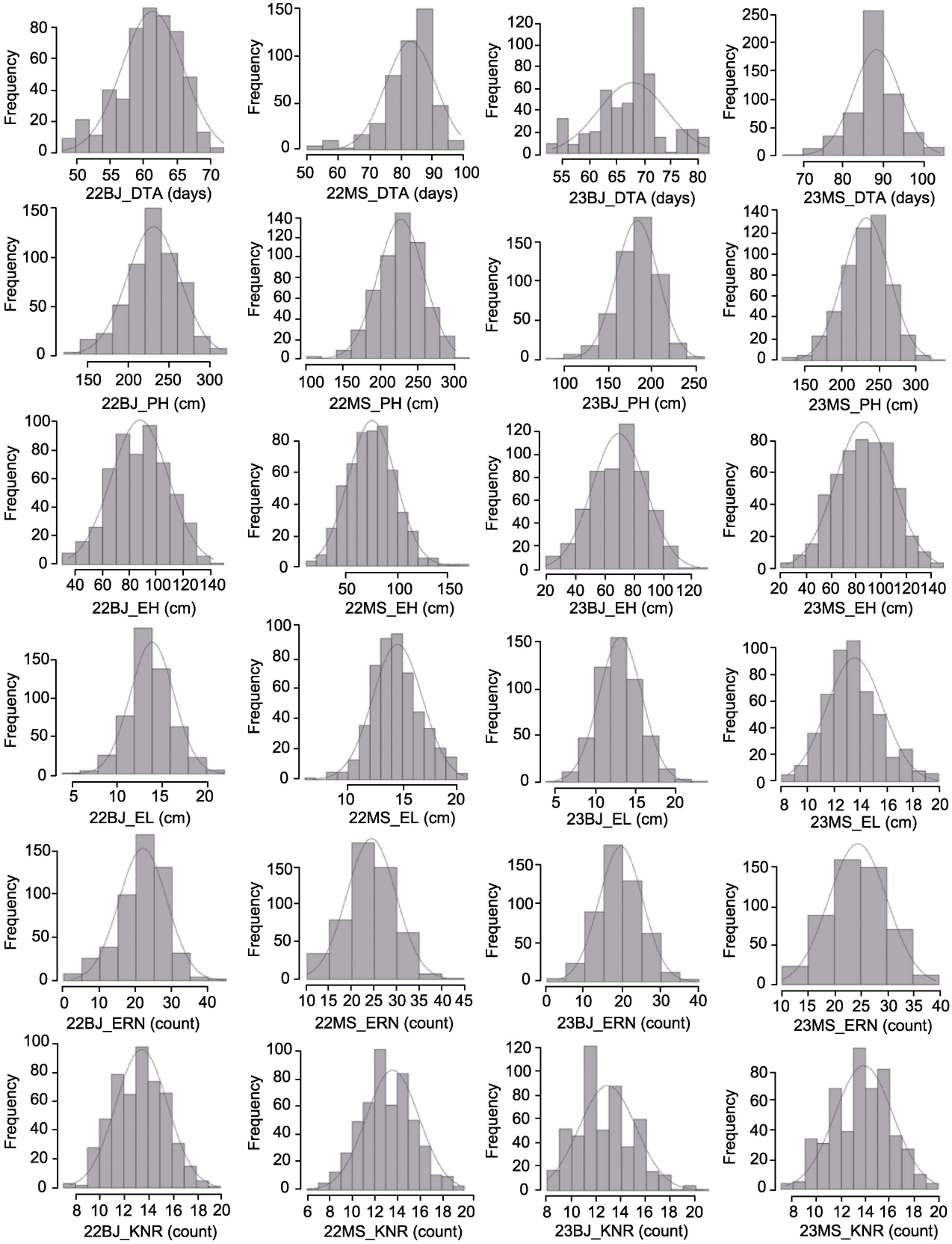

图2

Figure 2

图2 6个农艺性状4个环境下表型的正态分布
DTA、PH、EH、EL、KNR和ERN同表2。

Figure 2 Normal distribution of phenotype performance of six traits in four environments
DTA, PH, EH, EL, KNR, and ERN are the same as shown in Table 2.

表3 6个农艺性状4个环境联合方差分析
Table 3 ANOVA analysis of variance of phenotypes of six agronomic traits in four environments

Trait	σ_g^2	σ_{ge}^2	H^2	SE (H^2)
DTA	27.65***	8.79***	0.67	0.08
PH	736.31***	104.94***	0.76	0.01
EH	370.33***	67.62***	0.74	0.02
EL	2.94***	1.09***	0.50	0.26
KNR	13.74***	7.69***	0.38	0.04
ERN	3.25***	0.49***	0.60	0.46

σ_g^2: 遗传方差; σ_{ge}^2: 基因型与环境互作方差; H^2: 广义遗传力; SE(H^2): 遗传力估计的近似标准误; *** P<0.001; DTA、PH、EH、EL、KNR和ERN同表2。

σ_g^2: Genotypic variance; σ_{ge}^2: Genotype-by-environment interaction variance; H^2: Heritability in the broad sense; SE(H^2): Approximate standard error of the heritability estimate; *** P<0.001; DTA, PH, EH, EL, KNR, and ERN are the same as shown in Table 2.

2.3 不同模型对预测结果的影响

单环境模型在所有性状中的预测精度最低(图3; 表4)。对于2种训练集构成方式, 在CV1中, Single模型与Across模型、M×E模型和R-norm模型的预测精度相似, 4个模型之间平均预测精度的差异在0–0.01之间; 在CV2中, Single模型预测效果最差, Across模型与M×E模型和R-norm模型表现相似, 3种模型的预测精度比Single模型提升幅度范围分别为0.18–0.57、0.20–0.56和0.18–0.57。对于性状而言, 在CV1中DTA、EH、PH、EL、KNR和ERN 6个性状不同模型与Single模型相比提升幅度均在0–0.01之间; 在CV2中6个性状平均预测精度提升幅度在0.19–0.55之间, 其中PH的预测提升幅度最大(为0.55), DTA的提升幅度最小(为0.19)。

在CV1中, 不同性状不同模型均方误差(mean squared error, MSE)差异较小, 但在CV2中MSE值差异较大, 且Across模型、M×E模型和R-norm模型的MSE值均小于Single模型。不同性状不同模型在CV1和CV2下差异较大, Single模型在CV1和CV2中的MSE值差异不大, 但在CV2中, Across模型、M×E模型和R-norm模型的MSE值均小于CV1。

2.4 不同梯度TRN占总样本量比例对预测结果的影响

在CV1和CV2两种方案下, 不同模型不同TRN比例的平均预测精度变化趋势相似(表5)。当训练集占比为0.7时, 与占比为0.5时相比, 在CV1下, 不同模型DTA、EH、PH、EL、KNR和ERN 6个性状的平均预测精度分别提升了0.03、0.03、0.04、0.02、0.02和0.03; 在CV2下平均预测精度分别提升了0.03、0.01、0.00、0.06、0.05和0.04; 当训练集占比为0.9时, 与占比为0.5时相比, 在CV1下不同模型6个性状的平均预测精度分别提升了0.03、0.07、0.07、0.04、0.02和0.04; 在CV2下平均预测精度分别提升了0.01、0.02、0.00、0.06、0.07和0.05。对于不同模型而言, Single模型的平均提升幅度介于0.02–0.07之间, Across模型的平均提升幅度介于0.02–0.08之间, M×E模型的平均提升幅度为–0.01–0.09, R-norm模型的平均提升幅度为0.02–0.09。

3 讨论

玉米是我国主要的粮食作物, 种植环境纵跨寒温带、暖温带、亚热带和热带生态区, 分布在低地、平原、丘陵和高原山区等不同地形条件。玉米产量对不同环境的响应差异较大, 因此, 进行多环境测试是品种培育的重要过程, 研究玉米对不同环境条件的响应规律具有重要意义。本研究选取2个具有明显环境差异的地点, 黑龙江密山(属北方早熟玉米区)和北京顺义(属东华北中晚熟玉米区), 2个地点的维度差异非常大, 因此DTA在地点间差异较大, 不同地点年际间的差异较小, 而其它性状并未表现出类似的规律。因此, 在多环境预测中考虑不同地理环境信息, 尤其是光温因子对不同性状的影响, 从而实现精准预测非常重要。本研究中, 我们针对散粉期、株高、穗位高、穗长、

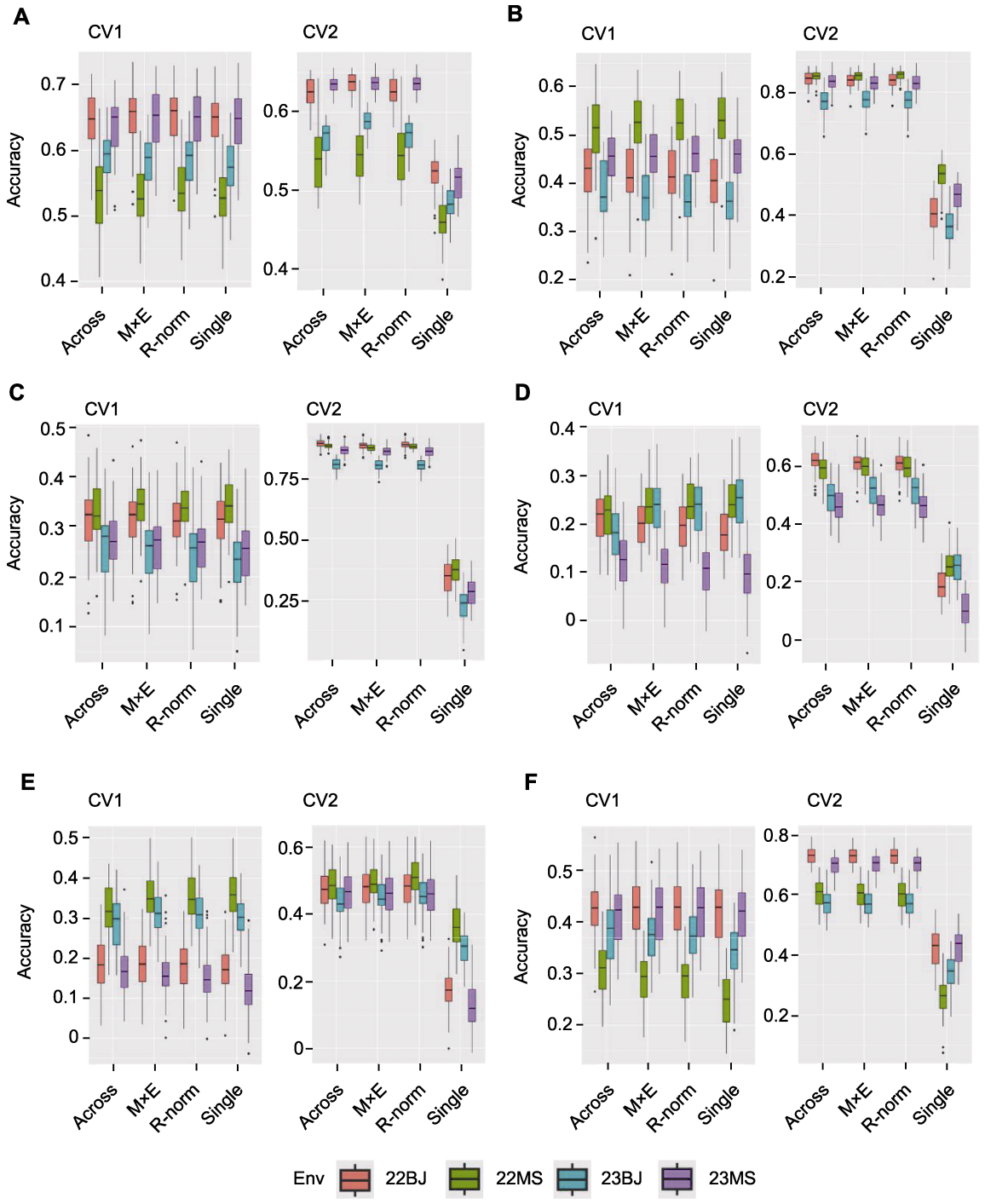

图3 不同性状CV1和CV2不同模型预测值与观测值的相关性

Single模型为利用单环境数据进行预测，Across模型、M×E模型和R-norm模型利用3个环境数据预测剩余1个环境下的表型数据。**(A)** 散粉期(DTA); **(B)** 穗位高(EH); **(C)** 株高(PH); **(D)** 穗长(EL); **(E)** 行粒数(KNR); **(F)** 穗行数(ERN)

Figure 3 Correlation between observed and predicted values of different models of CV1 and CV2 for different traits

The Single model uses single-environment data to predict, and the Across model, M×E model and R-norm model use three environmental data to predict phenotypic data in the remaining environment. **(A)** Days to anthesis (DTA); **(B)** Ear height (EH); **(C)** Plant height (PH); **(D)** Ear length (EL); **(E)** Kernel number per row (KNR); **(F)** Ear row number (ERN)

表4 不同性状CV1和CV2不同模型预测精度均方误差(MSE)分析

Table 4 Mean squared error (MSE) analysis of different models of CV1 and CV2 for different traits

		CV1				CV2			
		Single	Across	M×E	R-norm	Single	Across	M×E	R-norm
DTA	22BJ	17.05	15.96	16.19	16.19	17.03	8.07	6.28	7.47
	22MS	49.26	49.31	48.03	48.22	47.63	34.46	33.25	33.94
	23BJ	29.44	28.85	28.53	28.56	29.49	19.29	16.63	18.75
	23MS	23.67	23.52	22.91	22.94	24.07	9.41	9.43	9.62
EH	22BJ	351.07	339.97	343.11	343.84	351.35	125.72	129.73	128.55
	22MS	388.25	400.24	384.77	384.73	377.25	164.68	162.07	158.30
	23BJ	237.92	232.14	233.07	233.30	237.86	144.16	133.37	134.57
	23MS	405.39	401.68	397.68	398.11	406.28	174.36	176.11	175.36
PH	22BJ	819.45	816.72	816.21	816.97	833.93	218.41	232.62	225.98
	22MS	786.00	787.15	780.27	780.59	784.60	214.57	224.97	216.90
	23BJ	468.62	470.77	469.01	470.20	468.00	223.87	216.31	217.63
	23MS	895.11	885.29	893.44	893.03	919.63	292.51	302.67	299.63
EL	22BJ	4.61	4.55	4.58	4.60	4.61	3.00	3.05	3.05
	22MS	3.89	3.95	3.92	3.91	3.77	2.68	2.63	2.64
	23BJ	4.60	4.74	4.63	4.63	4.59	3.71	3.60	3.59
	23MS	3.95	3.90	3.96	3.98	3.88	3.23	3.20	3.23
KNR	22BJ	2.60	2.58	2.57	2.57	2.60	1.51	1.51	1.52
	22MS	3.67	3.53	3.58	3.58	3.57	2.46	2.47	2.47
	23BJ	3.09	3.00	3.01	3.01	3.08	2.52	2.49	2.49
	23MS	3.62	3.62	3.58	3.58	3.47	2.18	2.16	2.16
ERN	22BJ	29.68	29.41	29.60	29.59	29.65	23.72	23.80	23.64
	22MS	22.05	22.62	22.10	22.00	21.78	19.06	18.84	18.47
	23BJ	21.78	21.98	21.59	21.57	21.77	19.90	19.60	19.32
	23MS	26.27	25.70	25.90	26.09	25.12	20.06	20.02	20.11

DTA、PH、EH、EL、KNR和ERN同表2。DTA, PH, EH, EL, KNR, and ERN are the same as shown in Table 2.

穗行数和行粒数6个代表性农艺性状进行预测，以初步探究在国内不同经纬度和不同气候条件下玉米自交系响应不同环境的表型变化。

GP的应用可帮助育种家以更快速的方式选择目标基因型，从而增加遗传增益(Meuwissen et al., 2001; VanRaden, 2008; Crossa et al., 2010)。研究表明，使用多环境模型即在模型中纳入多个环境的数据进行建模的预测精度优于在单一环境下的预测精度(Burgueño et al., 2012; Jarquín et al., 2014)。多环境模型可以使用协方差函数(Burgueño et al., 2012)、标记和环境协变量(Jarquín et al., 2014)对G×E相互作用进行建模，或者通过对M×E相互作用建模(Lopez-Cruz et al., 2015)。在本研究中，当预测在某个环境下缺失的表型时(CV2)，多环境模型预测精度显著优于Single模型，这与Burgueño等(2012)的研究结果一致。本研究结果表明在多环境预测时模型中纳入多个环境数据的重要性。此外，多环境模型还有另一个优势，即能够以较少的样本量提供较高的预测精度，节省了表型评价资源，这对于减少育种中多环境测试压力具有重要意义。

在先前的报道中，研究人员利用模拟的数据(Lorenz, 2013)、小麦(*Triticum aestivum*)多环境试验数据(Burgueño et al., 2012)和玉米多环境试验数据(Jarquin et al., 2020)探究了在固定训练集大小的情况下，不同TRN组成方式对GP预测精度的影响。在本研究中，我们利用DTA、EH、PH、EL、KNR和ERN 6个性状分别对2种CV方案进行预测，同一种模型下，CV2的平均预测精度均高于CV1，这与前人的研究结

表5 CV1和CV2中不同梯度训练集(TRN)占总样本量比例4种模型预测结果

Table 5 Prediction results of four models with different percentage of training populations (TRN) in CV1 and CV2

Trait	TRN	Model	CV1	CV2	Trait	TRN	Model	CV1	CV2
DTA	0.5	Single	0.59	0.59	EH	0.5	Single	0.44	0.44
	0.5	Across	0.60	0.78		0.5	Across	0.45	0.82
	0.5	M×E	0.60	0.80		0.5	M×E	0.44	0.82
	0.5	R-norm	0.61	0.79		0.5	R-norm	0.44	0.82
	0.7	Single	0.62	0.63		0.7	Single	0.47	0.46
	0.7	Across	0.63	0.81		0.7	Across	0.47	0.82
	0.7	M×E	0.63	0.83		0.7	M×E	0.47	0.83
	0.7	R-norm	0.63	0.81		0.7	R-norm	0.47	0.83
	0.9	Single	0.62	0.63		0.9	Single	0.50	0.48
	0.9	Across	0.63	0.81		0.9	Across	0.51	0.84
	0.9	M×E	0.63	0.84		0.9	M×E	0.51	0.85
	0.9	R-norm	0.63	0.82		0.9	R-norm	0.51	0.84
PH	0.5	Single	0.28	0.31	EL	0.5	Single	0.20	0.20
	0.5	Across	0.29	0.86		0.5	Across	0.19	0.54
	0.5	M×E	0.29	0.85		0.5	M×E	0.20	0.55
	0.5	R-norm	0.29	0.86		0.5	R-norm	0.20	0.55
	0.7	Single	0.32	0.34		0.7	Single	0.23	0.22
	0.7	Across	0.33	0.85		0.7	Across	0.22	0.61
	0.7	M×E	0.33	0.85		0.7	M×E	0.23	0.62
	0.7	R-norm	0.33	0.85		0.7	R-norm	0.23	0.62
	0.9	Single	0.35	0.32		0.9	Single	0.23	0.22
	0.9	Across	0.37	0.85		0.9	Across	0.23	0.61
	0.9	M×E	0.36	0.85		0.9	M×E	0.24	0.62
	0.9	R-norm	0.36	0.85		0.9	R-norm	0.24	0.62
KNR	0.5	Single	0.24	0.24	ERN	0.5	Single	0.36	0.37
	0.5	Across	0.24	0.47		0.5	Across	0.39	0.65
	0.5	M×E	0.26	0.47		0.5	M×E	0.38	0.65
	0.5	R-norm	0.25	0.48		0.5	R-norm	0.38	0.65
	0.7	Single	0.27	0.27		0.7	Single	0.38	0.39
	0.7	Across	0.27	0.53		0.7	Across	0.40	0.69
	0.7	M×E	0.28	0.54		0.7	M×E	0.40	0.69
	0.7	R-norm	0.28	0.54		0.7	R-norm	0.40	0.69
	0.9	Single	0.26	0.28		0.9	Single	0.40	0.40
	0.9	Across	0.26	0.55		0.9	Across	0.42	0.70
	0.9	M×E	0.27	0.56		0.9	M×E	0.42	0.70
	0.9	R-norm	0.27	0.57		0.9	R-norm	0.42	0.70

DTA、PH、EH、EL、KNR和ERN同表2。DTA, PH, EH, EL, KNR, and ERN are the same as shown in Table 2.

果一致(Burgueño et al., 2012; Terraillon et al., 2023)。其可能的原因在于，在CV1中被预测的基因型在所有环境中均未进行过测试，无相关表型信息被纳入到模型中，因此预测精度只依赖于群体内不同个体间的亲缘关系。在CV2中，可利用同一基因型在其它环境下的表型数据进行模型拟合来预测，模型中考虑多个环境下的表型数据，提高了预测准确性。

在植物育种中，全基因组预测的准确性通常需要

训练集中包含良好且广泛的表型数据。在有限的测试资源下，为达到较好的预测效果，育种家需要平衡好测试的环境数量和每个环境下的基因型数量，因此抽样比例非常重要。如果在较小的抽样比例下可以达到较好的预测效果，训练群体中同等的样本量就可以包含更多的基因型或者环境数量，该结果不仅可为品种多环境测试节省大量成本，还可加快种质资源鉴定与评价。因此，我们比较了不同分配方案下不同抽样比例(0.5、0.7和0.9)对预测精度的影响，在进行比例设置时，CV1方案中TRN和TST之间无交叉材料，可随机进行选择预测，但在CV2方案中，模型的所有参数(包括方差组分)均是从每个TRN-TST分区的TRN数据中重新估计，结果表明减少训练集与总样本量的占比并未对预测结果产生较大影响。需要指出的是，本研究中采用的群体为个体之间亲缘关系较远的自然群体，适宜于种质资源的性状评价和筛选，但是与杂交育种的实际应用场景仍有一些区别，需进一步利用具有一定亲缘关系的群体(如半同胞群体、全同胞群体以及所组配的杂交种)，系统研究不同因素对预测精度的影响。

本研究表明，在对玉米自交系进行多环境预测时，不同训练群体构成方式对DTA、EH、PH、EL、KNR和ERN 6个性状的预测精度影响较大，在预测模型中纳入多环境测试数据可显著提升预测精度；当TRN占总样本量比例在0.5–0.9之间变化时，TRN比例的提升对于不同模型的预测精度影响有限。本研究通过对训练集进行设计，选择合适的预测模型实现大量个体在不同环境下的表型预测，有效减少了多环境测试的压力。利用该方法能够对种质资源开展多性状、多环境下表型精准鉴定，并结合基因型信息对资源进行全面综合评价，快速筛选具有高产、优质和抗逆等特性的材料，加速种质资源在育种中的有效利用，准确快速地培育出适应能力良好的品种。

作者贡献声明

李园: 收集表型数据、分析数据及撰写论文; 范开建: 论文润色; 安泰、李聪、蒋俊霞、牛皓、曾伟伟和衡燕芳: 协助收集表型数据; 李虎: 分析数据和修改论文; 李慧慧和付俊杰: 指导数据统计和计算及修改论文; 黎亮: 指导论文写作和修改。

参考文献

Alemu A, Åstrand J, Montesinos-López OA, Isidro Y, Sánchez J, Fernández-Gónzalez J, Tadesse W, Vetukuri RR, Carlsson AS, Ceplitis A, Crossa J, Ortiz R, Chawade A (2024). Genomic selection in plant breeding: key factors shaping two decades of progress. *Mol Plant* **17**, 552–578.

Burgueño J, De Los Campos G, Weigel K, Crossa J (2012). Genomic prediction of breeding values when modeling genotype × environment interaction using pedigree and dense molecular markers. *Crop Sci* **52**, 707–719.

Carena MJ, Hallauer AR, Filho JBM (2010). Quantitative Genetics in Maize Breeding. New York: Springer Science & Business Media. pp. 1–6.

Crossa J, De Los Campos G, Pérez P, Gianola D, Burgueño J, Araus JL, Makumbi D, Singh RP, Dreisigacker S, Yan JB, Arief V, Banziger M, Braun HJ (2010). Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. *Genetics* **186**, 713–724.

Cuevas J, Crossa J, Soberanis V, Pérez-Elizalde S, Pérez-Rodríguez P, De Los Campos G, Montesinos-López OA, Burgueño J (2016). Genomic prediction of genotype × environment interaction kernel regression models. *Plant Genome* **9**, plantgenome2016.03.0024.

Ferrão LFV, Ferrão RG, Ferrão MAG, Francisco A, Garcia AAF (2017). A mixed model to multiple harvest-location trials applied to genomic prediction in *Coffea canephora*. *Tree Genet Genomes* **13**, 95.

Fu JJ, Hao YF, Li HH, Reif JC, Chen SJ, Huang CL, Wang GY, Li XH, Xu YB, Li L (2022). Integration of genomic selection with doubled-haploid evaluation in hybrid breeding: from GS 1.0 to GS 4.0 and beyond. *Mol Plant* **15**, 577–580.

Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME (2009). Invited review: genomic selection in dairy cattle: progress and challenges. *J Dairy Sci* **92**, 433–443.

Jannink JL, Lorenz AJ, Iwata H (2010). Genomic selection in plant breeding: from theory to practice. *Brief Funct Genomics* **9**, 166–177.

Jarquín D, Crossa J, Lacaze X, Du Cheyron P, Daucourt J, Lorgeou J, Piraux F, Guerreiro L, Pérez P, Calus M, Burgueño J, De Los Campos G (2014). A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theor Appl Genet* **127**, 595–607.

Jarquin D, Howard R, Crossa J, Beyene Y, Gowda M, Martini JWR, Pazaran GC, Burgueño J, Pacheco A,

Grondona M, Wimmer V, Prasanna BM (2020). Genomic prediction enhanced sparse testing for multi-environment trials. *G3* (*Bethesda*) **10**, 2725–2739.

Lopez-Cruz M, Crossa J, Bonnett D, Dreisigacker S, Poland J, Jannink JL, Singh RP, Autrique E, De Los Campos G (2015). Increased prediction accuracy in wheat breeding trials using a marker × environment interaction genomic selection model. *G3* (*Bethesda*) **5**, 569–582.

Lorenz AJ (2013). Resource allocation for maximizing prediction accuracy and genetic gain of genomic selection in plant breeding: a simulation experiment. *G3* (*Bethesda*) **3**, 481–491.

Luo P, Wang HW, Ni ZY, Yang RS, Wang F, Yong HY, Zhang L, Zhou ZQ, Song W, Li MS, Yang J, Weng JF, Meng ZD, Zhang DG, Han JN, Chen Y, Zhang RZ, Wang LW, Zhao M, Gao WW, Chen XY, Li WJ, Hao ZF, Fu JJ, Zhang XC, Li XH (2023). Genomic prediction of yield performance among single-cross maize hybrids using a partial diallel cross design. *Crop J* **11**, 1884–1892.

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

Pérez P, De Los Campos G (2014). Genome-wide regression and prediction with the BGLR statistical package. *Genetics* **198**, 483–495.

Roorkiwal M, Jarquin D, Singh MK, Gaur PM, Bharadwaj C, Rathore A, Howard R, Srinivasan S, Jain A, Garg V, Kale S, Chitikineni A, Tripathi S, Jones E, Robbins KR, Crossa J, Varshney RK (2018). Genomic-enabled prediction models using multi-environment trials to estimate the effect of genotype × environment interaction on prediction accuracy in chickpea. *Sci Rep* **8**, 11701.

Sousa MBE, Cuevas J, de Oliveira Couto EG, Pérez-Rodríguez P, Jarquín D, Fritsche-Neto R, Burgueño J, Crossa J (2017). Genomic-enabled prediction in maize using kernel models with genotype × environment interaction. *G3* (*Bethesda*) **7**, 1995–2014.

Terraillon J, Roeber FK, Flachenecker C, Frisch M (2023). Training set designs for prediction of yield and moisture of maize test cross hybrids with unreplicated trials. *Front Plant Sci* **14**, 1080087.

VanRaden PM (2008). Efficient methods to compute genomic predictions. *J Dairy Sci* **91**, 4414–4423.

Wang BB, Lin ZC, Li X, Zhao YP, Zhao BB, Wu GX, Ma XJ, Wang H, Xie YR, Li QQ, Song GS, Kong DX, Zheng ZG, Wei HB, Shen RX, Wu H, Chen CX, Meng ZD, Wang TY, Li Y, Li XH, Chen YH, Lai JS, Hufford MB, Ross-Ibarra J, He H, Wang HY (2020). Genome-wide selection and genetic improvement during modern maize breeding. *Nat Genet* **52**, 565–571.

Xu YB, Liu XG, Fu JJ, Wang HW, Wang JK, Huang CL, Prasanna BM, Olsen MS, Wang GY, Zhang AM (2019). Enhancing genetic gain through genomic selection: from livestock to plants. *Plant Commun* **16**, 100005.

Yang N, Lu YL, Yang XH, Huang J, Zhou Y, Ali F, Wen WW, Liu J, Li JS, Yan JB (2014). Genome wide association studies using a new nonparametric model reveal the genetic architecture of 17 agronomic traits in an enlarged maize association panel. *PLoS Genet* **10**, e1004573.

Zhu XT, Leiser WL, Hahn V, Würschum T (2021). Training set design in genomic prediction with multiple biparental families. *Plant Genome* **14**, e20124.

Study on Multi-environment Genome-wide Prediction of Inbred Agronomic Traits in Maize Natural Populations

Yuan Li[1], Kaijian Fan[1], Tai An[2], Cong Li[3], Junxia Jiang[1], Hao Niu[1], Weiwei Zeng[2]
Yanfang Heng[1], Hu Li[1], Junjie Fu[1], Huihui Li[1], Liang Li[1*]

[1]State Key Laboratory of Crop Gene Resources and Breeding, The National Engineering Center for Crop Molecular Breeding, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China; [2]College of Agronomy and Biotechnology, China Agricultural University, Beijing 100193, China; [3]College of Agronomy, Heilongjiang Bayi Agricultural University, Daqing 163000, China

Abstract Multi-environment field testing is an important way to select optimize maize yield and yield stability varieties. However, because of its high cost, it has gradually become a challenge in plant breeding. The combination of field sparse testing and genome-wide prediction method can be used to predict untested phenotypes, reduced the effort and cost on field testing. In this experiment, 244 inbred lines of natural populations were planted in Shunyi, Beijing and Mishan, Heilongjiang in 2022 and 2023. Six agronomic traits were studied, including days to anthesis, plant height, ear height, ear length, kernel number per row and ear row number. The effects of four different models (Single, Across, M×E and R-norm), two different cross-validation schemes (CV1 and CV2) and three different training sets sampling ratios (0.5, 0.7 and 0.9) on the prediction accuracy were compared. The results showed that the average prediction accuracy of the six agronomic traits was 0.67, 0.58, 0.50, 0.33, 0.33 and 0.48. The average prediction accuracy of the Single model, Across model, M×E model and R-norm model was 0.36, 0.52, 0.53 and 0.53 for each trait. In CV1, the average prediction accuracy of each model in six traits ranged from 0.19 to 0.65, and in CV2, the average prediction accuracy ranged from 0.47 to 0.89. The comparison of different training set sampling ratios shows that the improvement of the proportion of the training sets has limited improvement in the prediction accuracy of different traits in different models, and the maximum is only 0.05. The results show that the CV2 training set can be used to form a scheme and include phenotypic data from multiple environments in the prediction model to provide good prediction accuracy for multi-environment prediction.

Key words maize, genome-wide prediction, multi-environment prediction, optimal training sets

Li Y, Fan KJ, An T, Li C, Jiang JX, Niu H, Zeng WW, Heng YF, Li H, Fu JJ, Li HH, Li L (2024). Study on multi-environment genome-wide prediction of inbred agronomic traits in maize natural populations. *Chin Bull Bot* **59**, 1041–1053.

---

* Author for correspondence. E-mail: liliang05@caas.cn

(责任编辑: 朱亚娜)

通讯作者/团队简介

黎亮, 研究员, 博士生导师, 中国农业科学院杰出青年英才, 农业农村部神农英才“青年英才”。在玉米单倍体育种技术方面的工作(参与)获得2017年度教育部技术发明一等奖、2017年度大北农科技奖植物育种奖和2023年度国家技术发明二等奖。目前, 研究团队主要聚焦单倍体技术和全基因组选择技术的有机整合, 以构建工程化的育种流程, 实现不同环境下的田间表型预测, 为创制高产耐密宜机收种质提供技术支撑。