

董春林, 张利, 李昊洋, 等. 基于全基因组选择预测玉米籽粒含水率[J]. 山西农业科学, 2025, 53(5): 1-7.

DONG C L, ZHANG L, LI H Y, et al. Prediction of maize grain moisture content using genomic selection[J]. Journal of Shanxi Agricultural Sciences, 2025, 53(5): 1-7.

doi:10.26942/j.cnki.issn.1002-2481.2025.05.01

基于全基因组选择预测玉米籽粒含水率

董春林¹, 张利¹, 李昊洋², 宋莹璐², 张鹏艳², 张正², 卜华虎², 常建忠²

(1. 太原生态工程学校, 山西 太原 030025; 2. 山西农业大学 山西有机旱作农业研究院, 山西 太原 030031)

摘 要: 玉米作为我国重要的粮食作物, 在保障我国粮食安全等方面发挥着重要作用。玉米成熟期籽粒含水率是玉米机械化粒收的重要衡量指标之一。培育成熟期低籽粒含水率的玉米品种是当前玉米育种的重要方向之一。籽粒含水率受复杂数量性状控制, 单靠表型选择效率较低, 全基因组选择技术能够对复杂数量性状进行快速筛选改良。在忻州和榆次 2 个地点对组配的 250 份玉米杂交种的籽粒含水率进行鉴定, 结合基因型数据, 利用 9 种 GS 模型开展全基因组选择分析。结果表明, 玉米杂交种平均籽粒含水率为 21.51%, 变异系数为 5.31%, 广义遗传力为 0.41, 特殊和一般配合力方差比值 $\sigma^2_{SCA}/\sigma^2_{GCA}$ 为 0.20, 表明玉米杂交种的籽粒含水率主要受加性效应控制。全基因组选择分析结果表明, 9 个 GS 模型的预测准确度平均为 0.592, 其中, rrBLUP、RKHS、BayesC、SVM 等 4 个模型预测准确性较高, 准确度均为 0.600, LASSO 模型的预测准确性最低, 为 0.572; 当标记密度为 2 000 个、训练群体大小为 70% 时, 预测准确性就能达到较高水平。在此基础上, 利用 rrBLUP 模型对 4 700 份玉米杂交种 GMC 进行预测, 发现 Top 100 杂交种的籽粒含水率平均为 22.63%, Bottom 100 杂交种的籽粒含水率平均为 18.56%。选择籽粒含水率预测结果的 Bottom 100 杂交种进行育种, 新玉米品种籽粒含水率相对于 Top 100 降低 4.08%, 相当于 17.98% 的增益。

关键词: 玉米; 复杂数量性状; 籽粒含水率; 全基因组选择分析; 标记密度; 预测准确度

中图分类号: S513

文献标识码: A

文章编号: 1002-2481(2025)05-0001-07

Prediction of Maize Grain Moisture Content Using Genomic Selection

DONG Chunlin¹, ZHANG Li¹, LI Haoyang², SONG Yinglu², ZHANG Pengyan²,
ZHANG Zheng², BU Huahu², CHANG Jianzhong²

(1. Taiyuan Ecological Engineering School, Taiyuan 030025, China;

2. Shanxi Institute of Organic Dryland Farming, Shanxi Agricultural University, Taiyuan 030031, China)

Abstract: Maize, as an important food crop in China, plays a vital role in ensuring food security. The grain moisture content(GMC) at maturity is a key indicator for mechanical grain harvesting of maize. Breeding maize varieties with low GMC at maturity has become a central focus of current breeding programs. GMC is controlled by complex quantitative traits, and it is inefficient to rely solely on phenotypic selection. Genomic selection(GS) enables rapid screening and improvement of complex quantitative traits. In this study, the GMC of 250 maize hybrid varieties was identified at two locations, Xinzhou and Yuci, genomic selection analysis was conducted in combination with genotype data using 9 GS models. The results showed that an average GMC of maize hybrid varieties tested was 21.51%, with a coefficient of variation of 5.31%, broad-sense heritability of 0.41, and a $\sigma^2_{SCA}/\sigma^2_{GCA}$ ratio of 0.20, indicating that the trait was mainly controlled by additive effects. The GS analysis revealed that the average prediction accuracy of the 9 GS models was 0.592. Among them, four models such as rrBLUP, RKHS, BayesC, and SVM models had higher prediction accuracy, with values of 0.600, while the LASSO model had the lowest prediction accuracy of 0.572. When the marker density was 2 000 and the training population size was 70%, the prediction accuracy reached a higher level. Based on this, using the rrBLUP model, the GMC of 4 700 maize hybrids was predicted. The

收稿日期: 2025-07-11

基金项目: 山西农业大学生物育种工程(YZGC146); 中央引导地方科技发展资金项目(YDZJSX20231C018); 山西省重点研发计划(202302140601018-2); 山西省现代农业产业技术体系建设专项(2025CYJSTX01-09)

作者简介: 董春林, 研究员, 主要从事玉米遗传育种研究, E-mail: 13513612843@163.com

通信作者: 常建忠, 研究员, 主要从事玉米分子育种研究, E-mail: sxyjycjz@sxau.edu.cn

results showed that the average GMC of the top 100 hybrids was 22.63%, while the bottom 100 hybrids had an average GMC of 18.56%. If breeding were conducted on the bottom 100 hybrids based on GMC predictions, the GMC would decrease by 4.08 compared to the top 100, resulting in a 17.98% gain.

Keywords: maize; complex quantitative traits; grain moisture content; genomic selection analysis; marker density; prediction accuracy

玉米(*Zea mays* L.)是我国重要的粮食作物之一,在保障我国粮食安全、经济稳定等方面发挥着重要作用。当前,我国玉米生产正在由追求高产的单一目标向高产高效、节本增效和提质增效协调统一的多重目标转变。机械化粒收是我国未来玉米机械收获的必然趋势和发展方向^[1],而玉米成熟期籽粒含水率是玉米机械化粒收的重要衡量指标之一。籽粒含水率的高低对玉米的产量和品质均有较大影响^[2],玉米成熟期籽粒含水率过高,会导致机械化粒收过程中籽粒破碎率升高,进而增加了霉变风险,增加了后期烘干、储藏等环节的生产成本,严重制约了玉米机械化粒收技术的大规模推广应用^[3]。因此,培育成熟期籽粒含水率低的玉米品种,是农业现代化、机械化发展的迫切要求^[4-5],也是当前我国玉米育种的重要方向之一。

全基因组选择(Genomic selection, GS)作为一种现代育种方法,最早由 MEUWISSEN 等^[6]于 2001 年提出,该技术利用覆盖全基因组的分子标记对训练群体进行基因型鉴定,结合训练群体的数据构建预测模型,进而对预测群体的表型进行精准预测,提高优良后代选择效率。传统分子标记辅助选择(MAS)仅针对主效基因,相比之下,GS 利用全基因组范围内的标记信息,能够显著提升对微效多基因控制的复杂数量性状的预测精度,使育种者能在育种早期高效筛选优良基因型,加速育种进程。GS 技术最先在动物育种中的研究和应用较为广泛。阳文攀等^[7]基于芯片数据对长白猪繁殖性状进行基因组选择研究,发现支持向量机(SVM)与随机森林(Random forest)等非参数机器学习模型在猪繁殖性状预测中具有一定的优势。章峰等^[8]利用 5 种杂交肉牛品种进行 GS,使用 GBLUP、BayesB 和 BayesR 等 3 种模型对肉牛干物质摄入和日增量等性状进行预测,发现选用高密度 SNP 标记、BayesR 模型和多品种训练群体的 GS 准确性最高。近年来,随着基因分型成本的下降,GS 技术已经逐渐在作物分子育种领域推广应用,如在水稻中主要用于预测杂交后代的育种潜力,涉及表型有产量^[9]、株高^[10]、抽穗期^[11]、抗病性^[12]等;在大豆中针对品质

性状的 GS 研究也较多,如蛋白质和油分^[13]、氨基酸含量^[14]、脂肪酸含量^[15]等;此外,在一些经济作物,如花生和棉花中也有少量应用^[16-17]。在玉米中,GS 技术研究主要针对自交系选育和杂交种选育 2 个方面,在自交系选育中,RIEDELSEIMER 等^[18]利用 635 份 DH 系对玉米产量三要素进行 GS 分析,发现全同胞家系群体预测准确性最高,若采用半同胞群体,其预测准确性下降 42%;YU 等^[19]基于 369 个玉米自交系组成的训练群体,对 2 687 个不同玉米自交系的 8 个玉米茎尖分生组织(SAM)相关性状进行预测,预测准确度为 0.37~0.57。在杂交种选育中,LUO 等^[20]采用部分双列杂交设计,组配了 285 份杂交种,在春夏播区 5 个地点进行表型鉴定,利用 GBLUP 模型进行 GS,结果发现,夏播区和春播区的玉米杂交种产量预测准确度分别达 0.51 和 0.46;ZHANG 等^[21]对 3 个玉米测交群体的表型、一般和特殊配合力进行了 GS 分析,结果表明,当将加性效应整合到模型中时,杂交种产量的预测准确度为 0.59~0.81;当将加性、非加性效应均整合到模型中时,预测精度为 0.64~0.86。籽粒含水率受复杂数量性状控制,单纯靠表型选择效率较低,利用 GS 有助于对玉米品种籽粒含水率进行快速改良。

本研究对 250 份玉米杂交种的玉米籽粒含水率进行了全基因组选择分析,评估了标记密度、群体大小和不同模型对预测准确度的影响,并以 250 份玉米杂交种的籽粒含水率建模,对 4 700 个玉米杂交种的籽粒含水率进行了全基因组预测,旨在为 GS 在选育低籽粒含水率玉米育种中的应用提供参考。

1 材料和方法

1.1 试验材料

100 份玉米自交系来自山西农业大学山西有机旱作农业研究院分子育种研究室,其中,父本群自交系 50 份,母本群自交系 50 份。利用这 100 份自交系组配出 250 份杂交种。

1.2 试验设计

试验于 2023 年在山西农业大学榆次东阳试验

示范基地(37.69°N,112.73°E)和山西农业大学玉米研究所忻州试验基地开展(38.42°N,112.73°E)。试验采用随机区组设计,每个玉米杂交种植2行,行长5 m,行距0.5 m,2次重复,种植密度6.75万株/hm²,常规大田水肥管理。

1.3 试验方法

1.3.1 玉米籽粒含水率测定及分析 于玉米收获期在各小区中间采集大小均匀一致的穗子,脱粒后,利用PM-8188A水分测定仪测定各杂交种的籽粒含水率(Grain moisture content, GMC),重复3次。使用ASReml 3.0软件对250个杂交种的籽粒含水率进行统计分析并计算最佳线性无偏预测值(Best Linear Unbiased Prediction, BLUP)和广义遗传力(H^2),利用Origin 2024进行绘图。

1.3.2 玉米基因型鉴定及分析 玉米出苗30 d后采集250份玉米杂交种植株的新鲜叶片,利用植物DNA提取试剂盒(DP350,北京天根生物科技有限公司)提取DNA,并通过1%琼脂糖凝胶检测DNA的质量和完整性。然后,采用45 K液相芯片(石家庄博瑞迪生物技术有限公司)进行基因分型,利用Plink 2.0软件基于最小等位基因频率(MAF)≥0.01,缺失率≤50%的标准对基因分型数据进行过滤^[22],进一步利用Beagle 5.4填充缺失位点^[23],获得高质量SNP信息。利用TASSEL 5的Create tree功能进行聚类分析^[24]。根据杂交种组配信息,将亲本的基因型数据合并即为杂交种基因型数据。

1.3.3 玉米籽粒含水率全基因组选择分析 试验

选择9种GS模型,利用组配的250份杂交种的籽粒含水率表型开展GS研究,模型包括岭回归最佳线性无偏预测(Ridge regression best linear unbiased prediction, rrBLUP)、BayesB、BayesC、Bayes-LASSO、岭回归(Ridge regression, RR)、弹性网络(Elastic net, EN)、再生核希尔伯特空间(Reproducing kernel hilbert space, RKHS)、随机森林(Random forest, RF)和支持向量机(Support vector machine, SVM),采用随机取样500次的5-fold交叉验证,其中,5-fold交叉验证具体为:将训练群体的个体,随机分成5等份,其中1份作为测试集,另外4份作为训练集,每一次建模都统计预测表型值与真实表型值间的皮尔逊(Pearson)相关系数来对模型的预测能力进行评价。所有模型均采用R4.3.0 (<https://www.r-project.org/>)程序实现。

为研究训练群体大小对GS的影响,本研究以10%为梯度分别提取群体的10%~90%的自交系作为训练群体,剩下的为预测群体,比较不同大小训练群体的预测准确性;为了探究不同标记密度对GS准确性的影响,在玉米10条染色体上均匀提取SNP标记,从每条染色体1个标记(标记密度为10)递增到每条染色体4 000个标记(标记密度40 000),共24个标记密度,每个标记密度重复500次,研究标记密度对预测准确性的影响。

2 结果与分析

2.1 玉米杂交种籽粒含水率的统计分析

玉米杂交种GMC的统计分析如表1所示。

表1 玉米杂交种GMC的统计分析

Tab.1 Statistical analysis of maize hybrid GMC

指标 Index	平均值±标准差 Mean±SD	最大值 Max.	最小值 Min.	范围 Range	变异系数/% CV	偏度 Skew	峰度 Kurt	中位数 Median	四分位数 Quartile	Herita- bility	$\sigma^2_{SCA}/\sigma^2_{GCA}$
籽粒含水率/% Grain moisture content	21.51±1.53	24.00	17.92	6.08	5.31	0.097	-0.069 1	20.60	Q1:15.2, Q2: 20.6, Q3:22.5	0.41	0.20
变异来源 Source of variance	σ^2_G	$\sigma^2_{GCAfemale}$	$\sigma^2_{GCAmale}$	σ^2_{SCA}	$\sigma^2_{G \times E}$	$\sigma^2_{GCAfemale \times E}$	$\sigma^2_{GCAmale \times E}$	$\sigma^2_{SCA \times E}$	σ^2_{Error}		
P	0.47***	0.08	0.31**	0.08	0.18***	0.10	0.08**	0.00	2.39		

注:**和***分别表示在0.01水平上显著和0.001水平上极显著。Heritability表示广义遗传力;GCA表示一般配合力;SCA表示特殊配合力;G表示基因型;E表示环境。Q1~Q3分别代表第一、第二和第三四分位数。

Note: ** and *** indicated significance at the 0.01 level and extremely significant at the 0.001 level, respectively. Heritability indicated Broad-sense heritability; GCA indicated general combining ability; SCA indicated specific combining ability; G indicated genotype; E indicated environment. Q1-Q3 respectively represented the first, second, and third quartiles, respectively.

250份玉米杂交种GMC在榆次、忻州的平均值分别为20.65%和20.75%,两地的BLUP均值为

21.51%, BLUP为17.92%~24.00%,变异系数为5.31%。由表1可知,杂交种群体的偏度和峰度值

的绝对值均接近于 0,表明杂交种的 GMC 数据基本接近正态分布,且均匀性较强。方差分析结果表明,GMC 基因型方差和基因型 $\times$ 环境方差均达到极显著水平($P<0.001$)。进一步将基因型方差分解为亲本效应方差,发现父本 GCA 方差及其与环境交互方差均达到极显著水平($P<0.001$)。广义遗传力分析发现,玉米杂交种 GMC 的 H^2 为 0.41, GMC 的 $\sigma^2_{SCA}/\sigma^2_{GCA}$ 值为 0.20,表明 GMC 具有中等大小的广义遗传力,适合早代筛选,且一般配合力效应占主导地位,说明对于 GMC,亲本的遗传优势对后代表型起决定作用。

2.2 基因型分析及聚类分析

玉米 45 K 液相芯片有效 SNP 标记数为 44 935,将原始标记按照 $MAF \geq 0.01$, $MR \leq 50\%$ 进行过滤,最终筛选到 43 034 个高质量 SNP 标记(图 1-A)。

过滤后的 SNP 标记,其平均缺失率为 0.043,MAF 为 0.182,多态性信息含量为 0.198,等位基因多态性为 0.287。将 100 份玉米自交系进行聚类分析发现,这些玉米自交系可分为 6 个类群,第 I 类为旅大红骨群,包括丹 340、E28、F141、KH128、DH218-9、LYA52 等 22 份玉米自交系;第 II 类为 P 群,包括 Qi319、X178、P138、JH49、H161、LYB92 等 16 份玉米自交系;第 III 类为塘四平头群,包括 Chang7-2、Huang-Zao4、K12、HuangC、789-1 等 10 份自交系;第 IV 类为 Reid 群,包含 Zheng58、Ye478、DH775、DH814-3、PH6 等 11 份玉米自交系;第 V 类为兰卡斯特群,包括 Mo17、35S、PHR32、PHK52、F118、H151 等自交系 23 份;第 VI 类为其他类群,包含 XZ3113、DH509、PHW5、YCZ1862、G17-3F 等 15 份自交系(图 1-B)。

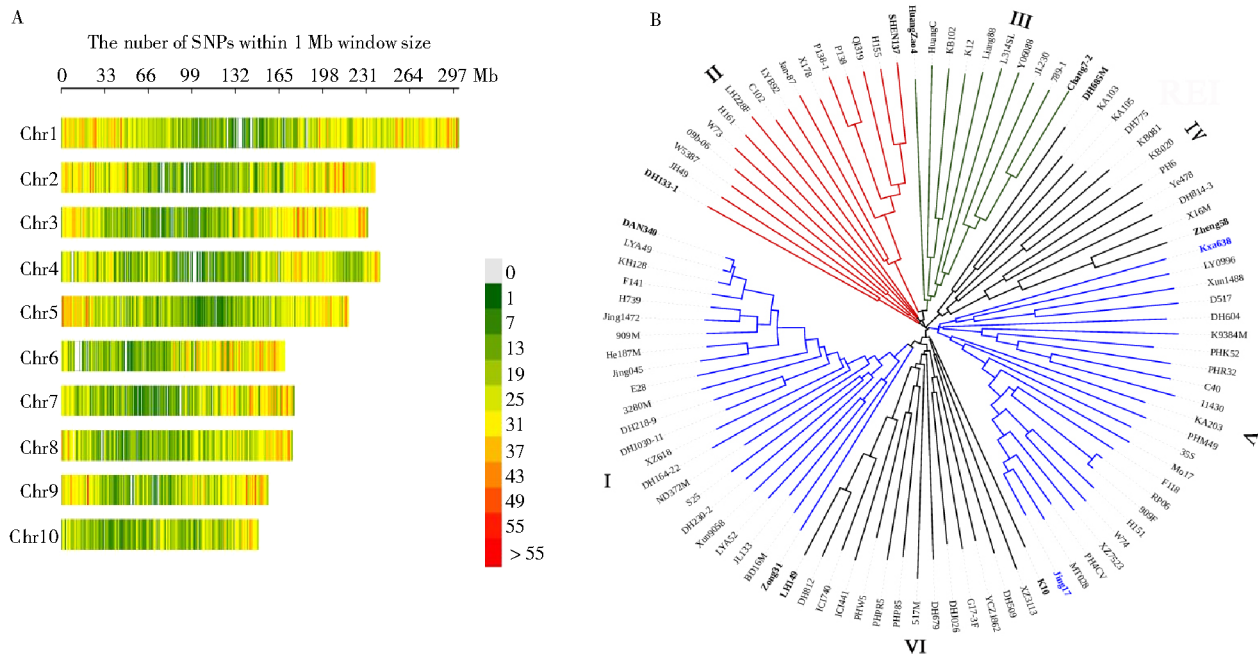


图 1 100 份玉米自交系的基因型信息和聚类分析

Fig.1 Genotype information and cluster analysis of 100 maize inbred lines

2.3 预测模型对 GMC 预测准确性的影响

为了评估现有预测模型对 GMC 预测准确性的影响,本研究选择了包括 rrBLUP、RKHS、BayesB、BayesC、RR、EN、LASSO、RF、SVM 等在内的 9 种模型,基于 43 034 个高质量 SNP 标记对榆次和忻州 2 个地点玉米杂交种的 GMC 的 BLUP 值进行全基因组选择分析,基于 10 倍交叉验证和皮尔逊相关系数对各模型的预测准确性进行评估,结果表明(表 2),9 个 GS 模型的预测准确度分别为 0.600、

0.600、0.599、0.600、0.591、0.576、0.572、0.585 和 0.600,平均值为 0.592,其中,rrBLUP、RKHS、BayesC、SVM 4 个模型的预测准确性较高,准确度均为 0.600,预测准确性最低的模型为 LASSO,准确度为 0.572。9 个预测模型预测准确性最高和最低之差仅为 0.028,可见,预测模型对 GMC 预测准确性的影响较小。考虑到不同预测模型运行时间成本,在同样条件下,由于 rrBLUP 运行时间最短,因此,在这 4 个模型中选择 rrBLUP 为最优模型。

表2 不同GS模型对玉米籽粒含水率的预测准确度

Tab.2 Prediction accuracy of different GS models for maize grain moisture content

指标	Index	rrBLUP	RKHS	BayesB	BayesC	RR	EN	LASSO	RF	SVM
预测准确度	Prediction accuracy	0.600	0.600	0.599	0.600	0.591	0.576	0.572	0.585	0.600

2.4 标记密度和群体大小对GMC预测准确性的影响

在筛选出最优预测模型的基础上,为探究不同标记密度对预测准确度的影响,在43 034个高质量SNP标记中选取标记,确保在玉米10条染色体上均匀分布,设置了24个标记密度,并基于rrBLUP模型进行GS分析(图2),发现随着标记密度逐渐增加,预测准确度总体呈上升趋势,当标记密度从10增加到60时,预测准确度从0.342迅速提升至

0.509,随后预测准确度波动较大,标记密度大于2 000时预测准确度趋于稳定,一直维持在0.6左右。随着训练群体大小从总群体的10%增加到90%,预测准确度呈逐渐上升趋势,当群体大小为总群体的10%时,其预测准确度为0.452,群体大小为总群体的70%时,预测准确度维持在0.6左右(图2)。综上,在rrBLUP模型下,标记密度为2 000和训练群体为群体大小的70%时,能够达到较好的预测效果。

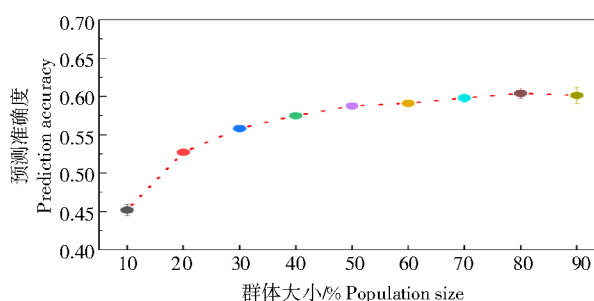
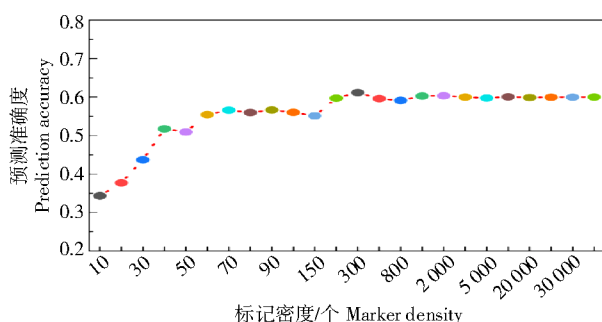
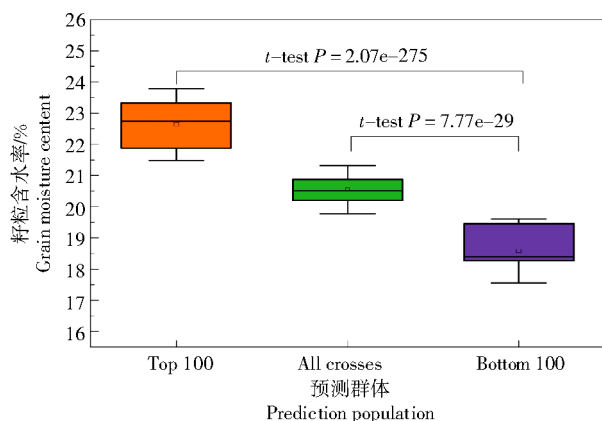


图2 标记密度和群体大小对全基因组选择预测准确性的影响

Fig.2 Effect of marker density and population size on genomic selection prediction accuracy

2.5 基于rrBLUP模型预测玉米杂交种GMC结果

玉米籽粒含水率全基因组选择结果分析如图3所示。



利用 t -test 进行显著性分析, $P < 0.01$ 表示不同处理间差异极显著
Using t -test for significance analysis, $P < 0.01$ indicated a highly significant difference between different treatments

图3 玉米籽粒含水率全基因组选择结果分析

Fig.3 Analysis of genomic selection results for maize grain moisture content

利用100份自交系通过不完全双列杂交,获得250份杂交种(实际上可以组配4 950个组合),并对其进行表型鉴定,利用250份玉米杂交种的GMC建模,在此基础上利用rrBLUP模型、2 000标记密度、70%训练群体,对剩余的4 700份(All crosses)玉米杂交种的GMC进行预测(图3),预测的GMC平均值为20.55%,最小值为18.79%,最大值为22.55%,变异系数为6.24%;选择按GMC大小排序前100(Top 100)和后100(Bottom 100)的杂交种进行分析,以Top 100相对于Bottom 100均值增加比例为衡量标准,评估GS的选择效益。Top 100杂交种平均GMC为22.63%,Bottom 100杂交种平均GMC为18.56%,因此,选择GMC预测结果Bottom 100进行育种,新品种的GMC相对于Top 100将降低4.07%,相当于17.98%的增益。

3 结论与讨论

GS准确性受多种因素影响,包括预测模型、标记密度、群体结构、预测性状的遗传力、训练群

体大小等^[25]。目前,存在多种基因组预测模型,如用于估算标记效应的混合模型 rrBLUP 和 GBLUP 模型^[26-27],贝叶斯线性回归模型(BayesB、BayesC、BayesLASSO等),以上模型均为参数化的线性模型,它们假设标记效应是加性的,即每个位点对表型的贡献是独立的^[28]。此外,还有半参数模型(RKHS)和非参数模型(BF、SVM等),这些模型能够在无需显式建模的情况下考虑非加性效应^[29]。ZHOU等^[30]基于11个非线性回归模型进行全基因组选择分析,结果发现,支持向量回归模型(SVR)对大豆分枝数的预测准确度为0.637,优于其他模型。MERRICK等^[31]分析了小麦深播出苗率的全基因组选择结果发现,在17种预测模型中,SVM的预测准确性最高,达0.56。本研究评估了包括rrBLUP、BayesB、BayesC、BayesLASSO等在内的9种GS模型对GMC的预测准确性的影响,结果发现,9种模型的预测准确度相当,均在0.6左右。可见,对于籽粒含水率来说,预测模型对其预测准确度影响不大,但一些研究表明预测模型对不同性状的预测能力有较大差异^[32-33],这种差异可能由性状本身的遗传结构、群体特性等因素导致。

训练群体大小和标记密度同样对预测准确度有影响。已有研究表明,随着训练群体规模的增加,性状的预测准确性会逐渐增加^[34]。本研究发现,在rrBLUP模型下,随着群体大小的增加,GMC的预测准确性逐渐增加,预测准确度从0.452增加到0.598,增加了32.28%,表明可以通过适当增加训练群体大小来提高预测准确度。但是,也有研究表明,当个体间亲缘关系较远或无亲缘关系时,增加群体大小并不能提高预测准确度^[35]。因此,在做GS之前,需要充分评估训练群体的遗传结构。

标记密度也是影响GS预测准确性的重要因素,随着标记密度的增加,预测准确度会相应增加。已有研究表明,当标记密度增加到一定值时,预测准确度会达到一个平衡,不再增加。POLAND等^[36]研究发现,在小麦中,当GBS标记达到1 827个时的预测准确度和34 729个标记的预测准确度相当。孙强等^[37]研究发现,当标记密度为500时,其对株高和穗位高的预测准确度与标记为50 000时相当。这种情况的出现,可能是由于标记间存在较强的连锁不平衡,导致增加更多的标记并不会有效提高模型的预测准确性^[9]。本研究在玉米各染色体上均匀选取标记,设置24个标记密度来研究标记密度对GMC预测准确性的影响,结果表明,当标记

密度为60时,即每条染色体选取6个标记,利用rrBLUP对GMC的预测准确度可达0.5,当标记密度为2 000时,即每条染色体选取200个标记,预测准确度可达到标记密度为40 000时的水平。在育种实际中,较低的标记密度有利于GS技术在育种中推广应用,过高的标记密度不仅会显著增加基因型检测成本,而且建模时间成本也会随之增加^[38]。

参考文献:

- [1] 王荣焕,徐田军,赵久然,等.收获期对玉米籽粒含水率及籽粒机收质量的影响[J].中国农业科技导报,2020,22(11):35-41.
WANG R H, XU T J, ZHAO J R, et al. Effects of harvest date on maize grain moisture content and grain mechanical harvesting quality[J]. Journal of Agricultural Science and Technology, 2020, 22(11): 35-41.
- [2] 姜艳喜,王振华,金益,等.玉米收获期籽粒含水量相关性状的遗传及育种策略[J].玉米科学,2004,12(1):21-25.
JIANG Y X, WANG Z H, JIN Y, et al. Genetics on water content at harvesting and correlative traits and breeding strategy[J]. Journal of Maize Sciences, 2004, 12(1): 21-25.
- [3] XIANG K, REID L M, ZHANG Z M, et al. Characterization of correlation between grain moisture and ear rot resistance in maize by QTL meta-analysis[J]. Euphytica, 2012, 183(2): 185-195.
- [4] 赵久然,王荣焕.中国玉米生产发展历程、存在问题及对策[J].中国农业科技导报,2013,15(3):1-6.
ZHAO J R, WANG R H. Development process, problem and countermeasure of maize production in China[J]. Journal of Agricultural Science and Technology, 2013, 15(3): 1-6.
- [5] 王克如,李少昆.玉米机械粒收破碎率研究进展[J].中国农业科学,2017,50(11):2018-2026.
WANG K R, LI S K. Progresses in research on grain broken rate by mechanical grain harvesting[J]. Scientia Agricultura Sinica, 2017, 50(11): 2018-2026.
- [6] MEUWISSEN T H E, HAYES B J, GODDARD M E. Prediction of total genetic value using genome-wide dense marker maps[J]. Genetics, 2001, 157(4): 1819-1829.
- [7] 阳文攀,刘相杰,罗冬香,等.基于芯片数据的长白猪繁殖性状基因组选择研究[J].畜牧兽医学报,2025,56(1):213-221.
YANG W P, LIU X J, LUO D X, et al. Research on genomic selection of reproductive traits in Landrace pigs based on chip data[J]. Acta Veterinaria et Zootechnica Sinica, 2025, 56(1): 213-221.
- [8] 章峰.肉牛剩余采食量及其组成性状的全基因组关联分析和基因组选择研究[D].南昌:江西农业大学,2017.
ZHANG F. Genome-wide association studies and genomic prediction for residual feed intake and their component traits in beef cattle[D]. Nanchang: Jiangxi Agricultural University, 2017.
- [9] SPINDEL J, BEGUM H, AKDEMIR D, et al. Genomic selection and association mapping in rice (*Oryza sativa*): effect of trait genetic architecture, training population composition, marker number and statistical model on accuracy of rice genomic selection in elite, tropical rice breeding lines[J]. PLoS Genetics, 2015, 11(2): e1004982.
- [10] 李广伟.水稻生殖隔离遗传结构解析、S5位点演化起源与基

- 于全基因组预测的亚种间杂种优势利用研究[D]. 武汉:华中农业大学,2018.
- LI G W. Genetic architecture of reproductive isolation in rice, evolutionary origination of S5 and utilization of inter-subspecific heterosis *via* genomic prediction[D]. Wuhan: Huazhong Agricultural University, 2018.
- [11] KIM K W, NAWADE B, NAM J, et al. Development of an inclusive 580K SNP array and its application for genomic selection and genome-wide association studies in rice[J]. *Frontiers in Plant Science*, 2022, 13: 1036177.
- [12] MAHANTESH, GANESAMURTHY K, et al. Analysis of the efficiency of genomic selection models for predicting sheath blight resistance in rice(*Oryza sativa* L.)[J]. *International Journal of Bio-Resource and Stress Management*, 2022, 13(3): 268-275.
- [13] MILLER M J, SONG Q J, LI Z L. Genomic selection of soybean(*Glycine max*) for genetic improvement of yield and seed composition in a breeding context[J]. *The Plant Genome*, 2023, 16(4): e20384.
- [14] SINGER W M, SHEA Z, YU D J, et al. Genome-wide association study and genomic selection for proteinogenic methionine in soybean seeds[J]. *Frontiers in Plant Science*, 2022, 13: 859109.
- [15] WANG F M, ZHAO T T, FENG Y, et al. Identification of candidate genes and genomic prediction of soybean fatty acid components in two soybean populations[J]. *Theoretical and Applied Genetics*, 2024, 137(9): 211.
- [16] 鲁清, 刘浩, 李海芬, 等. 花生含油量全基因组选择及近红外光谱筛选的育种技术探究[J]. *作物学报*, 2024, 50(4): 969-980.
- LU Q, LIU H, LI H F, et al. Research on oil content screen with genomic selection and near infrared ray in peanut(*Arachis hypogaea* L.)[J]. *Acta Agronomica Sinica*, 2024, 50(4): 969-980.
- [17] 李晓宇. 陆地棉纤维品质性状的全基因组选择方法的评价与优化[D]. 保定:河北农业大学, 2023.
- LI X Y. Evaluation and optimization of genome-wide selection method for fiber quality traits of upland cotton[D]. Baoding: Hebei Agricultural University, 2023.
- [18] RIEDELSHEIMER C, ENDELMAN J B, STANGE M, et al. Genomic predictability of interconnected biparental maize populations[J]. *Genetics*, 2013, 194(2): 493-503.
- [19] YU X Q, LEIBOFF S, LI X R, et al. Genomic prediction of maize microphenotypes provides insights for optimizing selection and mining diversity[J]. *Plant Biotechnology Journal*, 2020, 18(12): 2456-2465.
- [20] LUO P, WANG H W, NI Z Y, et al. Genomic prediction of yield performance among single-cross maize hybrids using a partial diallel cross design[J]. *The Crop Journal*, 2023, 11(6): 1884-1892.
- [21] ZHANG A, PÉREZ-RODRÍGUEZ P, SAN VICENTE F, et al. Genomic prediction of the performance of hybrids and the combining abilities for line by tester trials in maize[J]. *The Crop Journal*, 2022, 10(1): 109-116.
- [22] PURCELL S, NEALE B, TODD-BROWN K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses[J]. *American Journal of Human Genetics*, 2007, 81(3): 559-575.
- [23] BROWNING S R, BROWNING B L. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering[J]. *The American Journal of Human Genetics*, 2007, 81(5): 1084-1097.
- [24] 黄婧, 张敏. 基于SNP分子标记的乌饭树种质遗传多样性研究[J]. *植物遗传资源学报*, 2025, 26(3): 530-538.
- HUANG J, ZHANG M. Genetic diversity analysis of *Vaccinium bracteatum* germplasm based on SNP markers[J]. *Journal of Plant Genetic Resources*, 2025, 26(3): 530-538.
- [25] SHIKHA M, KANIKA A, RAO A R, et al. Genomic selection for drought tolerance using genome-wide SNPs in maize[J]. *Frontiers in Plant Science*, 2017, 8: 550.
- [26] ROBERTSEN C D, HJORTSHØJ R L, JANSS L L. Genomic selection in cereal breeding[J]. *Agronomy*, 2019, 9(2): 95.
- [27] CROSSA J, PÉREZ-RODRÍGUEZ P, CUEVAS J, et al. Genomic selection in plant breeding: methods, models, and perspectives[J]. *Trends in Plant Science*, 2017, 22(11): 961-975.
- [28] HESLOT N, YANG H P, SORRELLS M E, et al. Genomic selection in plant breeding: a comparison of models[J]. *Crop Science*, 2012, 52(1): 146-160.
- [29] DANILEVICZ M F, GILL M, ANDERSON R, et al. Plant genotype to phenotype prediction using machine learning[J]. *Frontiers in Genetics*, 2022, 13: 822173.
- [30] ZHOU W, YAN Z X, ZHANG L T. A comparative study of 11 non-linear regression models highlighting autoencoder, DBN, and SVR, enhanced by SHAP importance analysis in soybean branching prediction[J]. *Scientific Reports*, 2024, 14: 5905.
- [31] MERRICK L F, CARTER A H. Comparison of genomic selection models for exploring predictive ability of complex traits in breeding programs[J]. *The Plant Genome*, 2021, 14(3): e20158.
- [32] DAETWYLER H D, PONG-WONG R, VILLANUEVA B, et al. The impact of genetic architecture on genome-wide evaluation methods[J]. *Genetics*, 2010, 185(3): 1021-1031.
- [33] JANNINK J L, LORENZ A J, IWATA H. Genomic selection in plant breeding: from theory to practice[J]. *Briefings in Functional Genomics*, 2010, 9(2): 166-177.
- [34] DE AZEVEDO PEIXOTO L, MOELLERS T C, ZHANG J P, et al. Leveraging genomic prediction to scan germplasm collection for crop improvement[J]. *PLoS One*, 2017, 12(6): e0179191.
- [35] CHARMET G, STORLIE E, OURY F X, et al. Genome-wide prediction of three important traits in bread wheat[J]. *Molecular Breeding*, 2014, 34(4): 1843-1852.
- [36] POLAND J, ENDELMAN J, DAWSON J, et al. Genomic selection in wheat breeding using genotyping-by-sequencing[J]. *The Plant Genome*, 2012, 5(3): 103-113.
- [37] 孙强, 任姣姣, 徐晓明, 等. 玉米株高和穗位高QTL定位和全基因组选择探究[J]. *玉米科学*, 2022, 30(4): 40-47.
- SUN Q, REN J J, XU X M, et al. QTL mapping and genomic selection for plant height and ear height in maize[J]. *Journal of Maize Sciences*, 2022, 30(4): 40-47.
- [38] ELSHIRE R J, GLAUBITZ J C, SUN Q, et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species[J]. *PLoS One*, 2011, 6(5): e19379.