

张敖,关媛,张学才,等.玉米全基因组选择育种研究进展[J].上海农业学报,2023,39(5):13-18.

玉米全基因组选择育种研究进展

张敖¹,关媛^{2*},张学才²,杨双³,计舒文¹,刘翼宁¹,阮燕晔^{1**},郑洪建^{2**}

(¹ 沈阳农业大学生物科学技术学院,沈阳市玉米全基因组选择育种重点实验室,沈阳 110866;

² CIMMYT-中国特用玉米研发中心,上海 201403;³ 沈阳市农业科学院,沈阳 110121)

摘要:介绍了全基因组选择技术的由来和原理,阐释了不同的基因分型平台(GBS、DArT和rAmpSeq)和技术对实施全基因组选择育种技术的影响。此外,还介绍了开展全基因组选择育种所采用统计模型中的四大类算法,以及全基因组选择在玉米育种中的应用、面临的挑战以及发展前景。

关键词:全基因组选择;玉米育种;分子育种;分子标记辅助选择

中图分类号:S513 **文献标志码:**A **文章编号:**1000-3924(2023)05-013-06

Research progress in genomic selection breeding of maize

ZHANG Ao¹,GUAN Yuan^{2*},ZHANG Xuecai²,YANG Shuang³,JI Shuwen¹,

LIU Yining¹,RUAN Yanye^{1**},ZHENG Hongjian^{2**}

(¹ College of Bioscience and Biotechnology,Shenyang Agricultural University; Shenyang Key Laboratory of Maize Genomic Selection Breeding,Shenyang 110866,China;² CIMMYT-China Specialty Maize Research Center, Shanghai 201403,China;³ Shenyang Academy of Agricultural Sciences,Shenyang 110121,China)

Abstract: This article introduces the origin and principles of genome selection (GS) technology, and explains the impact of different genotyping platforms (GBS, DArT, and rAmpSeq) and technologies on the implementation of genome selection breeding technology. In addition, it also introduces the 4 major types of algorithms used in statistical models for conducting genome selection breeding, as well as the application, challenges, and development prospects of genome selection in maize breeding.

Key words: Genomic selection; Maize breeding; Molecular breeding; Molecular marker-assisted selection

玉米(*Zea mays* L.)是世界上最重要的谷物作物之一,其产量高、适应性广,在全球农业系统中起着越来越重要的作用^[1]。玉米的高产源自杂交种的广泛应用。杂交种是遗传上具有不同特点的两个亲本自交系的杂种一代,具有杂种优势,即优于亲本的田间表现^[1]。杂交种选育有两个关键环节:1)育种家需要选育出优良的亲本自交系;2)育种家需要筛选出优良的自交系杂交组合。优良自交系主要从不同遗传特性玉米种质杂交后代所形成的自交系群体中选育,通常利用测验种与大量待选自交系进行杂交,然后根据杂种一代的田间生长表现进行选择;而优良杂交种的选择则根据自交系杂交组合的生长表现。无论是优良自交系选择,还是优良杂交组合的选择,都需要做大量的杂交和田间比较试验。随着杂交育种水平的提高,需要更多的自交系和杂交组合才有可能选择到超过现有优良杂交种的杂交组合。然而,种植的自交系和杂交组合过多将增大育种成本,甚至降低田间比较试验的质量。事实上,育种单位很难将所有自交系都进行杂交,并将杂交后代都种植到田间进行比较试验,而是根据血统和经验抽取部分自交系进行杂交,这无疑会漏掉一些有潜力的杂交组合。所以,完全基于田间表现的表型育种方法已经不能满足高水平玉米育种的需要。为减少进行田间种植的数量,对育种材料包括自交系和杂交组合进行初步筛

收稿日期:2023-06-15

基金项目:辽宁省教育厅面上(青年)项目(LJKZ0657)

作者简介:张敖(1987—),男,博士,讲师,主要从事玉米遗传育种研究。E-mail:zhangao7@syau.edu.cn

* 共同第一作者

** 通信作者,E-mail:yanyeruan@syau.edu.cn;E-mail:hjzh6188@163.com

选,选出有良好遗传潜力的材料进行田间种植,将会提高育种的效率和质量。全基因组选择技术可以解决玉米杂交育种所面临的挑战,该技术根据育种材料的基因型数据预测其田间产量表现,可以对大批量育种材料进行初步筛选,随后将真正有开发潜力的育种材料进行田间试,节省育种资源。

1 全基因组选择在玉米育种上的优势

全基因组选择(Genomic selection 或 Genome-wide selection, GS)始于 1994 年,Bernardo^[2]提出了利用遗传协方差矩阵和表型方差/协方差矩阵计算 BLUP(最佳线性无偏估计)的方法,并证明该方法能够对玉米产量表现进行预测。随后,Meuwissen 等^[3]提出了全基因组选择的概念。全基因组选择虽然属于分子标记辅助选择育种技术的一种,但其在方法、程序和应用目标上都不同于传统的分子标记辅助选择技术(Molecular assisted selection, MAS),全基因组选择可应用于尚未进行分子标记鉴定的性状,它可以通过提高单位时间的遗传收益来加快育种进程。该技术利用遍布整个基因组的全部或显著分子标记预测个体的基因组估计育种值(Genomic Estimated Breeding Value, GEBV)作为筛选依据。在进行全基因组选择时,首先,构建一个建模群体(Training population 或 modeling population),测定其基因型和表现型,根据表型和基因型数据建立预测模型;其次,测定被预测群体(Testing population)的基因型;最后,利用已经建立的预测模型和被预测群体的基因型数据计算育种值(GEBV)并进行选择。

全基因组选择技术与传统的表型选择育种相比,具有诸多明显的优势。(1)玉米在全球分布广泛,遗传多样性丰富,不同来源的种质遗传构成差异大。全基因组选择通过分析整个基因组的分子标记和基因构成信息,可以直接评估个体的遗传变异,降低环境变异的影响,能更准确地预测个体的遗传潜力。(2)玉米植株高大,在田间试验时,玉米与小麦和水稻相比需要更大的土地面积,当育种规模巨大时,表型选择将无法企及。全基因组选择利用表型和基因型之间的联系预测群体中个体的性状表现,而不需要将所有的材料全部进行田间种植,从而增大选择的规模和效率。(3)玉米的生育期长,一年内繁殖的代数有限。当建立了可用的预测模型后,全基因组选择可以不经田间表型测定对个体进行预测,可以大大加快育种进程,增大遗传收益。(4)玉米的一些性状如花粉粒数等难以大规模测定,如果用全基因组模型进行预测选择,则可大大降低表型鉴定所需的人力、物力和时间成本。(5)由于极端天气难以遇到或难以适时遇到,对于极端逆境抗性玉米材料的选择就极为困难,利用长期积累的恶劣天气下玉米的表型数据与基因型建立预测模型,可在非极端逆境下对抗性材料进行选择。(6)全基因组选择可以迭代进步。通过积累大量的基因型和表型数据,结合多组学数据、多环境数据、基因功能及调控网络数据等,汇集到高级的统计模型,从而更全面、准确地评估玉米的遗传潜力,而传统育种的表型选择由于缺乏表型数据和基因型数据关系的信息积累,难以通过时间延续提高选择的质量。

2 基因组选择的基因型检测技术

若要将全基因组选择技术整合到实际的玉米育种流程中,需要高通量及经济的基因分型平台。目前用于玉米的主要基因分型平台有 GBS^[4]、DArT(<https://www.diversityarrays.com/>)和 rAmpSeq^[5-6]。

GBS(Genotyping-by-Sequencing)是一种简化的基因组测序平台,它是一种基于单/多酶切的高效基因组分析技术,可用于研究和鉴定多种生物体的基因组信息。GBS 主要通过将 DNA 样品进行切割和标记,然后使用高通量测序技术对其进行测序,从而获取目标生物体的基因组数据。相较于传统的全基因组测序方法,GBS 更为经济高效,并且适用于大规模的基因组分析研究。GBS 具有高通量、高覆盖度和高分辨率的特点,能够提供丰富的基因组数据,在玉米使用的 2.7 版中,可以将单核苷酸多态性(SNP)数据统一到 995 690 个标准碱基座上,方便多群体的合并和历史数据的使用^[7-9]。GBS 每个样品的成本为 \$ 30,约合人民币 200 元。

DArT(Diversity Arrays Technology),即多态性芯片技术,2001 年该技术被应用于基因型测定。2010 年,DArT 公司将 DArT 技术与简化基因组测序技术相结合,开发了 DArT-seq 测序平台,目前该公司开发了不同简化级别和不同价格的测序产品(<https://www.diversityarrays.com>)。目前,DArT 公司与国际玉米小麦改良中心(CIMMYT)合作在墨西哥建有 SAGA 测序平台。DArT-Seq 技术和其他简化基因组测序 GBS 技术相似,使用双酶法(*Pst*I 和 *Hpa*II 酶)对 DNA 样品进行酶切,减低染色体组的复杂性。酶切后来自不

同样本的 DNA 链接不同碱基组合的条形码,并链接测序接头,构建 DNA 简化测序文库。结合 Illumina 二代短片段测序技术(150 bp),将混合样本简化测序 DNA 文库放在单个甬道上测序。DArT-seq 获得的标记有两种,Silico-DArT 和 SNP,前者完全基于酶切的结果,不依赖于参考基因组,可使用七千多个标记信息用于全基因组预测;SNP 则将获得的 Silico 标记比对到参考基因组上。DArT 标记已经应用于玉米全基因组选择研究^[10-11]。DArT 平台每个样品的成本为 \$20,约合人民币 140 元。

rAmpSeq 是一种基于重复序列的遗传分型方法,可以更加高效地进行全基因组选择和物种保护的研究^[5]。在基因组 DNA 中使用 PCR 扩增一组选择性扩增的基因片段,这些片段包含可重复序列(如转座子)和周围的反向互补引物结合位点。PCR 反应产生的扩增产物使用 Illumina 平台上的序列技术进行测序,使基因组的 SNP 分析成为现实,并且仅需在重复序列的黏附区域记录变异。该平台在玉米上的标记数目为 7 595 个,已经足够满足全基因组选择的标记数量要求,需要注意的是该平台只适合于纯合材料的测序。rAmpSeq 平台标记在全基因组选择籽粒锌含量和第一阶段育种上都有应用,rAmpSeq 标记的全基因组预测精度略低于 GBS^[8,12]。该平台每个样品的成本仅为 \$5,是非常有潜力的基因组分型平台。

靶向测序基因型检测技术(GBTS)也是比较新的基因分型技术,基于多重 PCR 的 GenoPlexs 和基于液相探针捕获的 GenoBaits。多聚单核苷酸多态性(mSNP)技术极大地提高了目标位点(扩增子)内变异的检测效率。与 GBS 和固相芯片相比,GBTS 技术具有平台广适性、标记灵活性、检测高效性、信息可加性、支撑便捷性和应用广谱性等特点。目前已经在 20 余种主要农作物、蔬菜以及部分动物和微生物中开发了 GBTS 标记 50 余套。GenoPlexs 和 GenoBaits 2 种技术相结合,可广泛应用于生物进化、遗传图谱构建、基因定位克隆、标记性状关联检测、后裔鉴定、基因渐渗、基因累加、品种权保护、品种质量监测、转基因成分/基因编辑/伴生生物检测等领域。未来 GBTS 技术发展方向为:便携式、自动化、高通量、智能化检测平台;根据用户需求定制的可变密度、多功能分子检测;GBTS 与其他技术的整合;基于资源共享的开源育种等。应用该分型技术理论上可以完成玉米全基因组预测,但相关研究还未见报道。

3 统计模型与预测算法

全基因组选择所利用的数学模型主要有四大类算法,其各有特点,无论是那一种方法都需要大量的复杂计算。

3.1 最小二乘法(Least-squares estimation,LS)

该方法又称最小平方方法,在建立预测模型时,用观测数据的最小化误差的平方和来寻找数据的最佳匹配参数。在利用该模型进行基因组(单倍型效应)预测时,由于受自由度不足的限制,该方法不能同时对所有的效应进行估计,而是需要采取逐步循环程序来纳入效应。这种方法会高估被纳入的效应,而且预测精度也显著低于后面介绍的 BLUP 和贝叶斯法。但是,随着 DNA 分子上标记之间的距离增大,LS 估计的育种值(EBV)与真实育种值(TBV)之间的相关性增加,这或许是因为当标记数量减少时需要估计的效应也相应减少所致,抑或是因为需要估计的效应减少时,弱化了 LS 法自由度不足所产生的不利影响^[12]。最小二乘法直接应用于玉米全基因组选择的案例很少,更多的是与机器学习连用,计算模型的均方误差(MSE)和平均绝对百分比误差(MAAPE),或用于模型拟合的早停,防止过拟合^[13]。

3.2 BLUP 法

即最优线性无偏估计法(Best linear unbiased prediction)。在建立预测模型时,将表型数据(标记的效应)看作是各个标记的线性函数,然后利用最小二乘原理估计出各个标记的效应值,再进行方差检验,包括 F 测验和多重比较,据此确定各个标记与表型效应的关系参数。目前常用的 BLUP 法包括基因组最佳线性无偏预测(Genomic best linear unbiased prediction,GBLUP)^[14]和岭回归最佳线性无偏预测(Ridge regression best linear unbiased prediction,rrBLUP)^[15]。BLUP 法是目前最稳健的预测方法。

3.3 贝叶斯法(Bayesian Analysis)

在建立预测模型时,先假定未知参数的先验信息,然后利用贝叶斯公式根据表型数据(标记的效应)获得后验信息,再去推测未知参数。与前两种方法相比,贝叶斯法含有迭代过程,计算量更大,耗时更长。Bayes 全基因组回归需要大规模的建模群体去实现其优势,研究表明,采用 1 254 个材料建模,Bayes B 需

要巨大的计算量,相对而言,GBLUP 的计算量较少^[16]。CIMMYT 开发了包含贝叶斯模型的预测工具包 BGLR,包含了基因型与环境交互工具和多性状选择工具^[17-19]。贝叶斯方法已经应用在一些研究中^[20]。

3.4 机器学习法

机器学习是计算机科学的一个领域,其利用算法和已有范例捕捉目标模式特征。该方法又分为监管学习和无监管学习两种方式。监管学习的目的是依据输入数据推理预测出目标性状的数值;而无监管学习是发现输入变量之间的分组和关联。在机器学习中算法能够从已有样本中学习基因型与表现型的关系,然后根据基因型数据推测未进行观测的表型数据。用于全基因组预测的机器学习算法有支持向量机(Support vector machine, SVM)、随机森林(Random forests, RF)、人工神经网络(Artificial neural networks, ANN)和再生核希尔伯特空间(Reproducing kernel Hilbert space, RKHS)等,用于相应统计和数据挖掘的开源软件有 Weka^[21]和 R^[22]等,都可在 Windows、macOS 和 Linux 下运行。目前已经有两个比较知名的机器学习/深度学习建模方法^[23-24]。

除上述 4 类方法外,还有最小绝对值收敛和选择算子(Least absolute shrinkage and selection operator, LASSO)、非参数方法(Non-parametric methods)等。

4 全基因组选择在玉米育种中的应用实例

玉米育种的两大目标是选系和杂交种选育。目前,大多数玉米全基因组选择育种研究都集中在选系上。在淹水和干旱胁迫条件下,全基因组选择比传统的分子标记辅助选择更适用于玉米对于产量和生理性状的选择^[25]。基因型、基因型 + 环境、基因型 + 环境 + 基因型环境互作模型被用于玉米籽粒锌含量自交系选育的研究^[26]。在抗旱研究中,岭回归、贝叶斯 LASSO、贝叶斯 A、贝叶斯 B、随机森林、RKSH 等用于玉米选系研究,其中,贝叶斯 B 在研究所用的数据集中表现出最高的预测精度^[27]。曹士亮等^[9]对热带玉米抗复合焦油斑病抗性进行了全基因组选系研究,双亲群体获得了比关联群体更高的预测精度,最高平均 r_{MG} 为 0.74,能够满足育种需要。张敖等^[28]用高质量低密度 SNP 标记对 22 个热带玉米自交系群体进行了全基因组选择研究,评估了遗传力、建模群体规模、标记密度等对预测精度的影响,发现建模群体与预测群体的亲缘关系是最重要的影响因素,遗传力也有非常重要的影响。刘玉博等^[29]利用 CIMMYT 种质评估了玉米穗腐病的遗传解析和全基因组预测,发现整合 GWAS 的 SNP 信息用于全基因组预测能够得到较好的预测结果。崔震海等^[30]评估了苞叶松紧度的全基因组预测精度,并提出当预测精度有限时,可以选择去掉预测结果最差的部分,更加贴合生产实践。Li 等^[31]使用来自甘肃和陕西 2 个异源亲本系列的 119 个自交系之间的 490 个杂交组合进行了全基因组预测,并测试了 8 个不同的统计预测模型,得到了许多有潜力的单交组合的选择。王楠等^[12]进行了早期测交的跨年份预测,发现累加的多年数据可以有效地提升预测精度。张敖等^[32]评估了测交试验中杂交种表现和配合力的全基因组预测,发现育种早期的测交阶段,材料数比较少,将自交系与测验种的数据都添加到预测模型中,可以大幅提高杂交种表现和一般配合力的预测。

5 全基因组选择在玉米育种中面临的挑战

全基因组选择技术在玉米育种中虽然具有巨大的潜力,但也面临一些挑战。

(1) 技术成本:全基因组选择技术需要进行大规模的基因组测序和数据分析,使得其实施成本较高。庞大数据的存储和处理需求都需要投入大量的资金,这对于一些资源有限的育种项目或育种工作者而言可能是一个限制因素。然而,随着技术的不断进步和成本的逐渐下降,预计这个挑战会逐渐减轻。

(2) 数据管理和分析:全基因组选择技术生成的数据量庞大,需要高效的数据管理和分析方法。基因组数据的存储、处理和解读需要强大的计算和信息技术基础设施。育种者需要具备数据科学和生物信息学的知识和技能,以便有效地处理和解释这些大规模的基因组数据。此外,数据共享和标准化也是一个重要的问题,以便不同研究团队之间可以共享和比较数据,进一步推动全基因组选择技术的发展。

(3) 性状复杂性:玉米的许多重要性状是微效多基因控制的,并且这些基因还可能与环境因素相互作用。全基因组选择技术可以帮助鉴定与复杂性状相关的基因组区域,但解析这些区域的功能和相互关系仍然具有挑战性。理解复杂性状的遗传机制需要深入的功能研究和生物信息学分析。此外,复杂性状的

遗传改良可能需要对大量基因进行选择,增加了育种者在实践中的困难。

(4)基因型与环境互作:玉米的生长和发育过程受到环境条件的影响,而且这些环境条件可能与基因型之间存在复杂的相互作用。全基因组选择技术主要关注基因组水平的遗传信息,但忽略了环境对表型的影响。这可能导致在不同环境条件下,基于全基因组选择的育种策略效果不稳定。因此,为了更好地利用全基因组选择技术,需要进一步研究基因型与环境互作的机制,以开发更稳健和适应性强的品种。

(5)表型数据的准确性与稳定性:随着科技的进步,基因型鉴定的成本在大幅降低的同时,准确性却在不断提高,而表型鉴定误差较大的问题正在被放大。全基因组选择预测精度的挑战已经从基因型转到了表型,开发高通量的准确表型鉴定系统,是全基因组选择预测准确性进一步提高的前提。

(6)模型的优化和选择:全基因组预测模型已经从线性到非线性再到机器学习、深度学习。基于深度学习的预测模型展现出了很强的活跃度。深度学习模型可以自动学习和提取复杂的特征表示,从而能够更好地捕捉基因型与表型之间的非线性关系,提高预测精度。

(7)多组学数据整合:将基因组学数据与其他组学数据进行整合,如转录组、甲基化组和蛋白质组数据,来揭示基因型与表型之间的关联。通过综合多组学数据的信息,可以更准确地识别与目标性状相关的基因型,提高全基因组选择的预测精度。

虽然全基因组选择技术在玉米育种中面临一些挑战,但随着技术的进步和研究的深入,这些挑战将逐渐得到克服。随着成本的降低、数据分析方法的改进以及对环境互作效应和复杂性状的更深入理解,全基因组选择技术有望成为玉米育种的重要工具,为培育出更高产、适应性更强和更优质的玉米品种作出重要贡献。

6 全基因组选择在玉米育种中的发展前景

随着生物技术的不断进步,全基因组选择技术已经成为现代育种的重要工具之一,它将基因组信息与遗传评估相结合,帮助育种者更高效地筛选和选择优良品种。

首先,全基因组选择技术极大地加速了玉米育种的进程。传统的育种方法通常需要多年甚至几十年的时间来选择和改良具有特定性状的品种。然而,全基因组选择技术利用高通量测序技术,可以快速而准确地识别出与目标性状相关的基因型。育种者可以更迅速地进行杂交和选择,从而大大缩短育种周期。

其次,全基因组选择技术提供了更高的遗传改良效率。通过对大量个体的基因组进行测序和分析,全基因组选择技术能够全面了解每个个体的遗传构成。育种者可以更好地预测目标性状在后代中的表现,并有针对性地选择具有优良基因型的个体进行繁殖。这种精准的遗传改良方法可以显著提高育种效率,加速优质玉米品种的培育。

综上所述,全基因组选择技术在玉米育种中具有广阔的发展前景。通过加速育种进程、提高遗传改良效率以及解决复杂性状问题,全基因组选择技术有望推动玉米育种的突破和进步,为农业生产提供更高产、更抗逆、更优质的玉米品种,从而为粮食安全和农业可持续发展作出贡献。

参 考 文 献

- [1] ERENSTEIN O, JALETA M, SONDER K, et al. Global maize production, consumption and trade: trends and R&D implications [J]. Food Security, 2022, 14 (5): 1295-1319.
- [2] BERNARDO R. Prediction of maize single-cross performance using RFLPs and information from related hybrids [J]. Crop Science, 1994, 34 (1): 20-25.
- [3] 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.
- [4] 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.
- [5] BUCKLER E, ILUT D C, WANG X Y, et al. rAmpSeq: Using repetitive sequences for robust genotyping [J]. bioRxiv, 2016. DOI: 10.1101/096628.
- [6] 徐云碧, 杨泉女, 郑洪建, 等. 靶向测序基因型检测 (GBTS) 技术及其应用 [J]. 中国农业科学, 2020, 53 (15): 2983-3004.
- [7] WANG N, YUAN Y B, WANG H, et al. Applications of genotyping-by-sequencing (GBS) in maize genetics and breeding [J]. Scientific Reports, 2020, 10: 16308.

- [8] GUO R,DHLIWAYO T,MAGETO E K,et al. Genomic prediction of kernel zinc concentration in multiple maize populations using genotyping-by-sequencing and repeat amplification sequencing markers[J]. *Frontiers in Plant Science*,2020,11:534.
- [9] CAO S L,LOLADZE A,YUAN Y B,et al. Genome-wide analysis of tar spot complex resistance in maize using genotyping-by-sequencing SNPs and whole-genome prediction[J]. *The Plant Genome*,2017,10(2):plantgenome 2016. 10. 0099.
- [10] ZHANG A,CHEN S,CUI Z H,et al. Genomic prediction of drought tolerance during seedling stage in maize using low-cost molecular markers[J]. *Euphytica*,2022,218(11):154.
- [11] DOS SANTOS J P,PIRES L P,DE CASTRO VASCONCELLOS R C,et al. Genomic selection to resistance to *Stenocarpella maydis* in maize lines using DArTseq markers[J]. *BMC Genetics*,2016,17(1):86.
- [12] WANG N,WANG H,ZHANG A,et al. Genomic prediction across years in a maize doubled haploid breeding program to accelerate early-stage testcross testing[J]. *Theoretical and Applied Genetics*,2020,133(10):2869-2879.
- [13] MONTESINOS-LÓPEZ O A, MONTESINOS-LÓPEZ A, PÉREZ-RODRÍGUEZ P, et al. A review of deep learning applications for genomic selection[J]. *BMC Genomics*,2021,22(1):19.
- [14] HABIER D,FERNANDO R L,GARRICK D J. Genomic BLUP decoded;a look into the black box of genomic prediction[J]. *Genetics*,2013,194(3):597-607.
- [15] ENDELMAN J B. Ridge regression and other kernels for genomic selection with R package rrBLUP[J]. *The Plant Genome*,2011,4(3):250-255.
- [16] TECHNOW F,SCHRAG T A,SCHIPPRACK W,et al. Genome properties and prospects of genomic prediction of hybrid performance in a breeding program of maize[J]. *Genetics*,2014,197(4):1343-1355.
- [17] PÉREZ P,DE LOS CAMPOS G. Genome-wide regression and prediction with the BGLR statistical package[J]. *Genetics*,2014,198(2):483-495.
- [18] CUEVAS J,CROSSA J,MONTESINOS-LÓPEZ O A,et al. Bayesian genomic prediction with genotype $\times$ environment interaction kernel models[J]. *G3 Genes|Genomes|Genetics*,2017,7(1):41-53.
- [19] MONTESINOS-LÓPEZ O A, MONTESINOS-LÓPEZ A, LUNA-VÁZQUEZ F J, et al. An R package for Bayesian analysis of multi-environment and multi-trait multi-environment data for genome-based prediction[J]. *G3 Genes|Genomes|Genetics*,2019,9(5):1355-1369.
- [20] LIU X G,WANG H W,WANG H,et al. Factors affecting genomic selection revealed by empirical evidence in maize[J]. *The Crop Journal*,2018,6(4):341-352.
- [21] KOTTHOFF L,THORNTON C,HOOS H H,et al. Auto-WEKA; automatic model selection and hyperparameter optimization in WEKA[M]// *Automated Machine Learning*. Cham:Springer International Publishing,2019:81-95.
- [22] THE R DEVELOPMENT CORE TEAM. R: A language and environment for statistical computing[M]. Vienna: R Foundation for Statistical Computing,2014.
- [23] LIN X. DeepGS: deep representation learning of graphs and sequences for drug-target binding affinity prediction[EB/OL]. 2020, arXiv: 2003. 13902. [https://arxiv.org/abs/2003. 13902. pdf](https://arxiv.org/abs/2003.13902.pdf).
- [24] WANG K L,ALI ABID M,RASHEED A,et al. DNNP,a deep neural network-based method for genomic prediction using multi-omics data in plants[J]. *Molecular Plant*,2023,16(1):279-293.
- [25] CERRUDO D,CAO S L,YUAN Y B,et al. Genomic selection outperforms marker assisted selection for grain yield and physiological traits in a maize doubled haploid population across water treatments[J]. *Frontiers in Plant Science*,2018,9:366.
- [26] MAGETO E K,CROSSA J,PÉREZ-RODRÍGUEZ P,et al. Genomic prediction with genotype by environment interaction analysis for kernel zinc concentration in tropical maize germplasm[J]. *G3 Genes|Genomes|Genetics*,2020,10(8):2629-2639.
- [27] 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.
- [28] ZHANG A,WANG H W,BEYENE Y,et al. Effect of trait heritability, training population size and marker density on genomic prediction accuracy estimation in 22 bi-parental tropical maize populations[J]. *Frontiers in Plant Science*,2017,8:1916.
- [29] LIU Y B,HU G H,ZHANG A,et al. Genome-wide association study and genomic prediction of *Fusarium* ear rot resistance in tropical maize germplasm[J]. *The Crop Journal*,2021,9(2):325-341.
- [30] CUI Z H,DONG H X,ZHANG A,et al. Assessment of the potential for genomic selection to improve husk traits in maize[J]. *G3 Genes|Genomes|Genetics*,2020,10(10):3741-3749.
- [31] LI G L,DONG Y,ZHAO Y S,et al. Genome-wide prediction in a hybrid maize population adapted to Northwest China[J]. *The Crop Journal*,2020,8(5):830-842.
- [32] 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.

(责任编辑:闫其涛)