

基于分子标记技术的玉米品种杂种优势模式鉴定方法

杨扬[†], 田红丽[†], 易红梅[†], 王璐, 范亚明, 常诗惠, 王风格^{*}, 宋伟^{*}, 赵久然^{*}

北京市农林科学院玉米研究所, 农业农村部农作物DNA指纹创新利用重点实验室(部省共建), 玉米DNA指纹及分子育种北京市重点实验室, 北京 100097

[†] 同等贡献

^{*} 联系人, E-mail: maizezhao@126.com; songwei@baafs.net.cn; wangfengge@baafs.net.cn

2025-01-01 收稿, 2025-01-27 修回, 2025-04-07 接受, 2025-04-10 网络版发表

农业生物育种国家科技重大专项(2022ZD04019)、北京市农林科学院科技创新能力建设专项(KJCX20230303)和北京学者项目(BSP041)资助

摘要 单交种的杂种优势模式代表了玉米品种的遗传背景, 一直以来只能通过其系谱资料或亲本自交系的杂种优势群划分结果来确定. 随着现代农业进入商业化育种时代, 传统的杂种优势模式鉴定方法遇到系谱未知和亲本保密等因素而无法有效实施. 本文提出了一种完全基于分子标记技术的杂种优势模式鉴定方法(简称完全分子标记法), 在品种系谱信息未知的前提下, 通过分析单交种的DNA指纹就能鉴定出其两个亲本的杂种优势群. 结果表明: (1) 230个常用自交系组成的对照自交系库通过三种分类方法互相验证, 确认该库包含了9个杂种优势群, 分别是塘四平头群(黄早四改良群), X群、旅大红骨群、P群、兰卡斯特群、瑞德群、改良瑞德群、混合群1和混合群2; (2) 在对照自交系指纹库的基础上构建包含36种模式的杂种优势模式权重库, 并通过虚拟组合确定其置信阈值和判定规则; (3) 335个已通过国家审定品种作为待测杂交种, 通过完全分子标记法鉴定出其主要的杂种优势模式(不区分正反交)分别是改良瑞德群×塘四平头群、改良瑞德群×P群和P群×旅大红骨群. 本研究可为获取玉米单交种遗传背景信息提供一种客观、简便、有效的分析手段, 不仅有助于精准推广栽培品种和评估遗传多样性, 还对保障国家粮食安全具有重要意义.

关键词 玉米, 单交种, SSR标记, 杂种优势群, 杂种优势模式

20世纪以来, 杂种优势利用极大地促进了农业生产, 是植物遗传学对农业最大的贡献之一^[1]. 基于孟德尔遗传定律等经典遗传学, 科学界建立了自交系和自交系之间的有性杂交等概念. 国内外作物育种实践表明, 选择合适的杂种优势利用模式能够有效提升新品种培育的成功概率. 近40年来, 育种工作者在逐渐理解杂种优势遗传学三种经典假说的前提下, 通过合理分类亲本材料的杂种优势群并建立有效的杂种优势模式, 是进一步提升育种成效的重要策略^[2]. 在玉米育种中, 育种家利用其异花授粉的特点, 培育出了众多优异自交系并成功组配了多种杂交组合, 从而使玉米杂交种的商业化利用取得了显著进展. 了解玉米品种的遗传

背景不仅能帮助育种家对其组配材料分门别类和精准推广, 还能有效评估、保护玉米品种的遗传多样性.

杂种优势群的传统概念是指遗传基础广阔、遗传变异丰富, 具有较多的有利基因、较高的一般配合力, 种性优良的育种基础群体^[3]. 将此概念与现阶段的育种技术融合, 玉米杂种优势群特指从杂种优势利用的角度对人工培育的玉米自交系遗传背景进行分类的统称. 它具有以下三个特征: (1) 每个群包含的自交系有一定数量规模; (2) 不同自交系的亲缘关系, 均表现为群内遗传距离小、群间遗传距离比较大; (3) 不同自交系通过组配, 均表现出群内杂种优势小、群间杂种优势大. 早期的玉米杂种优势群鉴定主要通过搜集系谱

引用格式: 杨扬, 田红丽, 易红梅, 等. 基于分子标记技术的玉米品种杂种优势模式鉴定方法. 科学通报, 2025, 70: 3191–3202

Yang Y, Tian H, Yi H, et al. Heterosis pattern detection method of maize single-cross hybrids via molecular marker (in Chinese). Chin Sci Bull, 2025, 70: 3191–3202, doi: [10.1360/TB-2025-0001](https://doi.org/10.1360/TB-2025-0001)

来源进行确定,例如孙琦等人^[4]通过对2001~2012年我国玉米骨干自交系相关资料的统计与分析得知当时玉米种质的遗传基础主要是塘四平头群、旅大红骨群、Reid群、PB群、Lancaster群。现阶段的杂种优势群鉴定则依靠不同分子标记技术进行分类,例如刘志斋等人^[5]利用SSR(simple sequence repeat)标记将820份中国重要玉米自交系划分为兰卡斯特群、旅大红骨群、塘四平头群、瑞德群与P群;赵久然等人^[6]基于SNP(single nucleotide polymorphism)芯片把有重要育种价值的344份玉米自交系划分为旅大红骨群、黄改群(塘四平头群)、Iodent群、兰卡斯特群、P群、改良瑞德群、瑞德群和X群;Zhang等人^[7]使用SNP芯片对代表中国温带气候玉米种质的269个自交系进行了聚类分析,结果划分为Imported Reid群、Domestic Reid群、兰卡斯特群、Zi330群、塘四平头群、P群和其他群。综上所述,目前国内主要的玉米杂种优势群主要包括塘四平头群(黄改群)、旅大红骨群(Zi330群)、瑞德群(Reid群或Domestic Reid群)、改良瑞德群(Imported Reid群)、兰卡斯特群(Lancaster群)、P群(PB群)以及X群这七大类。

杂种优势模式的传统概念是指两个不同的杂种优势群之间具有较高的基因互作效应,具有较高的特殊配合力,相互配对成为产生强杂种优势的模式^[3]。本文中玉米杂种优势模式特指玉米单交种的两个亲本对应的杂种优势群组合,它代表了玉米单交种的遗传背景,具有以下三个特征:(1) 每种模式均由两个杂种优势群组配出超亲优势;(2) 每种模式都能组配出一定数量规模的品种;(3) 每种模式都有一个或多个代表性品种且被广泛应用和认可。在育种理论研究中,玉米杂种优势模式鉴定分为两类方法,一类是正向推导,基于数量遗传学分析法对已知杂种优势群的自交系通过双列杂交模式设计组配出单交组合,该组合的杂种优势模式是已知信息^[8,9];另一类是反向推导,在亲本信息获取有限情况下鉴定单交种双亲的杂种优势群,该情况下单交种的杂种优势模式是未知信息。目前反向推导类的玉米杂种优势模式鉴定方法有三种:第一种是完全系谱法,它基于表型组遗传关系,根据单交种的品种来源(即父母本名称)及亲本的系谱来源来推断其杂种优势模式,例如王懿波等人^[10]依据审定品种的品种来源得出中国玉米有10种主要杂交优势利用模式;第二种是系谱结合分子标记法,它基于基因组遗传多态性,通过单交种给定的双亲名称搜集对应自交系样本,再借助分子标记对亲本自交系进行杂种优势群划分,进而推

断单交种的杂种优势模式,例如滕文涛等人^[11]利用SSR标记对71个杂交种的84份亲本自交系进行杂种优势类群的划分和利用模式的分析;第三种是种皮分子标记法,它基于单交种的种皮母本发育特征,在无亲本自交系样本的情况下,推测单交种双亲分子标记等遗传特征^[12],再借助双亲的类群划分推测其杂种优势模式,例如王春喜^[13]采用SSR标记验证玉米籽粒种皮DNA与其母本一致。前两种方法本质上均依赖单交种的品种来源信息进行杂种优势模式鉴定,当遇到亲本的系谱来源信息或者亲本自交系难以获得时,这两种鉴定方法均会出现失效。第三种方法虽然理论上可行,但由于其试验操作要求高且工作量大,并未见采用该方案进行杂种优势模式鉴定的报道。

随着“十三五”以来国家和省级玉米品种审定的试验通道不断拓宽,在2014和2016年先后开通了绿色通道试验和联合体试验,通过两级审定的玉米杂交品种数量急剧增加^[14],给品种管理工作带来了新的挑战:如何在品种审定过程中采取有效的监督手段来保障玉米品种遗传背景的多样性,从而促进玉米种子产业的健康发展。由于国内两级审定的玉米品种以单交种为主,鉴定玉米单交种的杂种优势模式成为解析市场上玉米品种遗传背景的最直接、有效的技术手段。然而现有的杂种优势模式鉴定方法受到诸多因素的制约而存在局限性,在实际应用中亟需一种不依赖品种来源信息、无需取得亲本材料可精准鉴定玉米单交种杂种优势模式的方法。

本文设计了一种完全基于分子标记的玉米单交种杂种优势模式鉴定方法,无需提供任何品种来源信息,只需检测待测品种DNA指纹就能有效推断其亲本所属的杂种优势群。本研究选取230个常用玉米自交系作为对照自交系,利用SSR指纹数据,采用类群划分和虚拟杂交的方法确定杂种优势模式权重库和阈值;通过对335个玉米杂交种杂种优势模式的鉴定,进而验证该方法的有效性和准确性。在实际应用中,利用本文提出的方法分析农作物品种DNA指纹库,可深入解析已知或待测品种的遗传背景信息,从而为品种选育、管理和推广提供科学依据。

1 材料与方法

1.1 材料与试验设计

本文分别采用《常用玉米自交系SSR指纹图谱》

收集的230个常用玉米自交系和《国家审定玉米品种SSR指纹图谱》收集的335个玉米杂交种分别作为杂种优势模式鉴定的对照自交系和待测杂交种^[15,16],前者覆盖国内外常用的玉米自交系、后者主要包含2000~2013年通过国家审定的玉米单交品种。上述样品的SSR指纹库构建均设计了两组以上的平行试验,样品制备、DNA提取、PCR扩增、PCR产物检测和数据采集等实验流程及核心引物信息具体参考行业标准《玉米品种鉴定技术规程 SSR 标记法》(NY/T1432-2014)和已发表的论文^[17]。两组以上的平行试验数据通过指纹合并算法,最终获得所有样品的40个SSR核心引物位点的指纹图谱。

本文使用的对照自交系包含了遗传背景明确且代表性得到育种家公认的16个参照系,包括塘四平头群的代表系黄早四和京92, X群的代表系京724和京MC01, P群的代表系齐319和X178, 旅大红骨群的代表系自330和丹340, 兰卡斯特群的代表系Mo17, 瑞德群的代表系B73和U8112, 改良瑞德群的代表系掖478和郑58, 糯玉米群的代表系京糯6、白糯6和紫糯5B。

1.2 研究方法

结合玉米单交种的培育特点, 本文设计了一种基于分子标记数据的杂种优势模式鉴定分析方法(以下简称完全分子标记法), 如图1所示该方法的分析流程包括三个环节。

一是筛选对照自交系并分类, 具体流程分为四步: (1) 选择种质来源广泛、在常规育种领域有代表性且和待测杂交种有一定亲缘关系的材料作为对照自交系; (2) 选用一种准确率高的分子标记(如SSR, SNP和InDel(insertion/deletion)等)进行DNA指纹库构建; (3) 从位点和样品两个维度控制对照自交系的指纹数据质量, 位点维度可通过最大等位基因频率、最小等位基因频率、多态信息量、位点缺失率等评估参数筛选高质量位点作为本方法的鉴定依据, 样品维度可通过数据缺失率、位点纯合率等评估参数确定是否补充试验数据; (4) 以对照自交系的指纹数据为分析依据, 采用系统发生分析、群体结构分析等常规类群划分方法确定每个自交系所属的杂种优势群。

二是确定杂种优势模式权重库和判定阈值, 具体流程分为四步: (1) 采用杂种优势群之间两两组合、同群不杂交的方式设计杂种优势模式, 由于对照自交系作为母本或父本进行虚拟杂交时, 其DNA指纹数据无

法识别其亲本身份, 因此本方法在设计模式时不区分杂种优势群之间的正、反交; (2) 按照设定的杂种优势模式对分属于不同群下自交系的指纹数据进行全部遍历且不重复的模拟杂交, 得到所有杂种优势模式第一代虚拟杂交种(first generation virtual hybrids, FGVH)的指纹数据; (3) 依据FGVH的指纹数据及其所属杂种优势模式信息, 统计每个模式对应的FGVH在所有位点的基因型频率作为评估待测杂交种的杂种优势模式的打分依据(下文统称为杂优模式权重库); (4) 以FGVH为训练样品, 累加求和每个模式对应的基因型频率, 作为FGVH在对应杂种优势模式的得分, 统计第*i*种模式分值的平均值 $\bar{X}_i$ 和标准偏差 s_i , 并将 $(\bar{X}_i - 2s_i)$ 作为该模式得分置信区间的下限。

杂优模式权重库根据每个模式的FGVH样品指纹, 统计每个位点在每一种基因型频率, 通过求和方式归并为每个模式的打分依据, 公式如下:

$$P_k = \sum_{i=1}^m \frac{g_i}{G_i}, \quad (1)$$

式中, P_k 为第*k*种杂种优势模式的评分值, g_i 为该模式第*i*个位点出现和待测杂交种相同基因型的FGVH数量, G_i 为该模式第*i*个位点有基因型数据的FGVH数量, m 为评估杂种优势模式使用的位点数量。

三是分析待测杂交种的杂种优势模式, 具体流程分为四步: (1) 待测杂交种采用和对照自交系相同的分子标记检测平台进行DNA指纹库构建, 建库位点和第一个环节保持一致; (2) 从样品维度控制待测杂交种的指纹数据质量, 数据缺失率是评估待测样品是否补充试验数据的主要参数; (3) 结合第二个环节得到的杂优模式权重库, 通过基因型匹配计算待测杂交种在每个杂种优势模式的得分值, 符合判定规则的杂优模式即为待测杂交种的杂种优势模式; (4) 查询该模式对应的得分阈值, 并与待测杂交种的最高得分进行比较, 如果在其阈值范围内, 说明该分析结果可靠, 否则需要扩大对照自交系的样品范围并重新分析。

1.3 数据处理

由于部分软件不支持中文字符, 本文将玉米自交系的样品名称中文字符替换为其对应的拼音字母。在SSR指纹数据质量控制方面, 采用PowerMarker V3.0软件^[18]提供的Summary Statistics功能对所有样品位点维度的数据质量进行评估, 样品维度的数据质量则采用

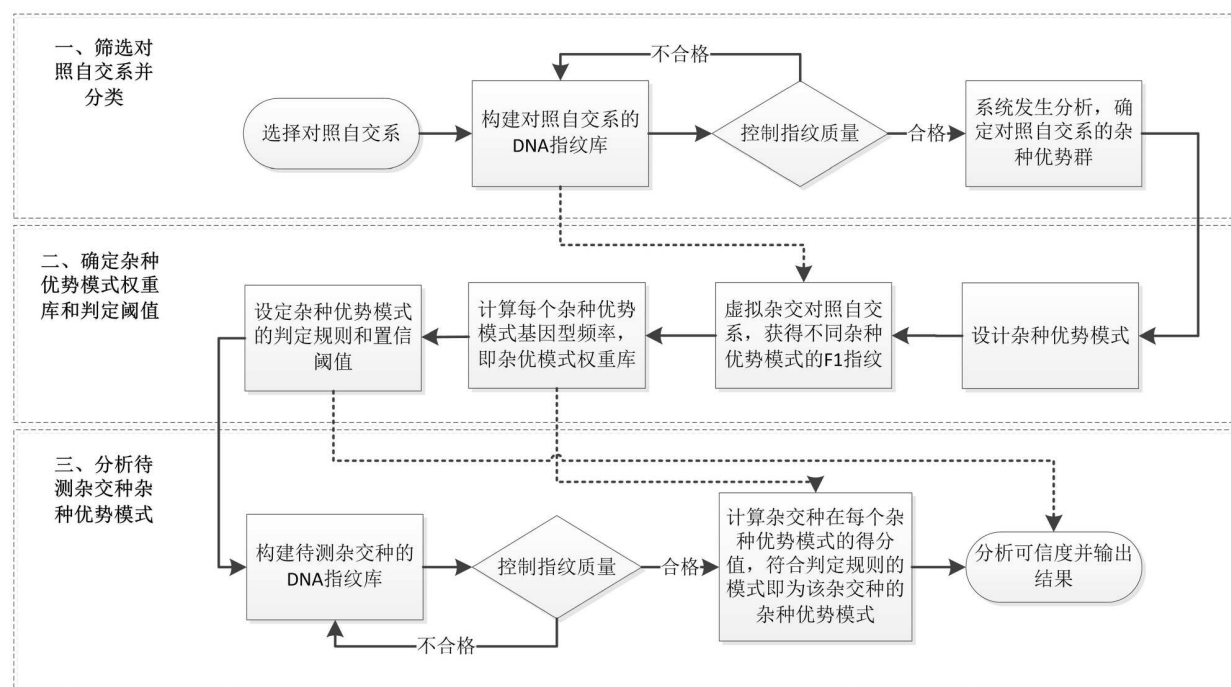


图1 基于完全分子标记法的农作物杂种优势模式鉴定分析流程图

Figure 1 Analysis flow chart of crop heterosis pattern detection based on the pure molecular marker method

“SSR比对统计工具V1.0”的样品统计功能进行评估。所有数据质量评估参数采用Graphpad Prism 8软件提供的Frequency distribution功能进行频率分布和描述性统计。

230个玉米自交系作为杂优模式库的对照自交系需要借助常规类群划分方法确定其杂种优势群，具体分析步骤如下。首先，利用PowerMarker V3.0软件^[18]提供的Rogers’s 1972方法对230个自交系之间的遗传距离矩阵进行计算，选择邻接法构建系统发生树，并以参照系作为分类依据进行类群划分，系统发生树则采用Evolview V3.0在线软件^[19]进行渲染。其次，通过MVSP V3.13软件对上述的遗传距离矩阵进行主坐标分析，同时用上述类群划分结果对所有样品进行分组标注。最后，采用Structure 2.3.4软件^[20]对230个自交系进行群体结构分析，其中退火周期长度和马尔科夫链蒙特卡罗重复次数均设置为10000， K 设置为2-14，每个 K 重复计算10次，输出结果后导入Structure Harvester工具^[21]计算 ΔK 。

杂优模式权重库通过将对照自交系指纹数据和分群结果输入“农作物杂优模式鉴定工具 V1.0”计算获得，335个玉米杂交种的杂种优势模式结果则通过将其

指纹数据和杂优模式权重库输入相同软件计算获得。使用和对照自交系相同参数的系统发生树方法对待测杂交种进行分组验证，并将该结果和杂种优势模式推测结果同时用Evolview V3.0在线软件^[19]进行比较展示。杂交种品种来源对应的亲本自交系的系谱关系追溯采用“SSR比对统计工具V1.0”进行分析。通过基于遗传距离矩阵的主坐标分析对待测杂交种进行杂优模式特征验证，即所有样品分组标注均依据完全分子标记法分析的杂种优势模式结果。

2 结果

2.1 SSR指纹数据质量

230个玉米自交系SSR指纹数据缺失率均为0，位点纯合率平均值为0.954，其中位点纯合率 ≥ 0.960 的样品占比76.1%。40个SSR标记在对照样品中的等位基因总数和为503，最大等位基因频率平均值为0.386，遗传多样性指数平均值为0.742，多态信息量平均值为0.711。上述数据质量评估结果一方面说明本文使用的对照样品具备自交系高纯合度、种质来源广泛和数据完整的特点，另一方面也证明本文选用的SSR标记作为建库位

点在特异性、多态性和稳定性等方面均表现出色。

335个玉米杂交种SSR指纹数据缺失率均为0, 位点杂合率平均值为0.639, 其中位点杂合率 ≥ 0.600 的样品占比81.8%。40个SSR标记在待测样品中的等位基因数总和为533, 最大等位基因频率平均值为0.403, 遗传多样性指数平均值为0.724, 多态信息量平均值为0.713。上述数据质量评估结果不仅说明待测样品符合玉米杂交种特征和指纹数据完备, 也证明本套SSR标记在玉米杂交种鉴定效果同样具备特异性、多态性和稳定性。

2.2 对照自交系的类群划分

以参照系的遗传背景信息作为参照, 基于邻接法的系统发生树将230个玉米自交系划分为9个杂种优势群(图2(a))。除了自330被分到兰卡斯特群、紫糯5B被分到混合群1、京糯6被分到混合群2、白糯6被分到混合群2、其余13个参照系的杂种

良瑞德群, 其余12个参照系的分群结果与遗传背景信息一致。选取有系谱信息的135个玉米自交系对分群结果进行校验, 其中107个自交系的系谱信息和分群结果吻合, 准确率为79.3%。基于主坐标分析的散点图选取的两个向量解释率分别为28.82%和16.28%(图2(b)), 其中塘四平头群、P群和改良瑞德群对应的样品不存在重叠交叉区域, 其余6个杂种优势群对应的样品则存在重叠交叉区域, 说明在当前解释向量条件下前3个杂种优势群划分结果可以通过主成分分析得到验证, 后者则不能通过该方法验证。

群体结构分析的结果表明, ΔK 在 $K=4, 7, 9$ 表现出较高数值(图2(c)), 本文结合参照系的实际情况, 选择与系统发生树相同类群数 $K=9$ 的群体结构结果进行分析比较(图2(d)), 紫糯5B被分到混合群1、京糯6被分到混合群1、白糯6被分到混合群2, 其余13个参照系的杂种

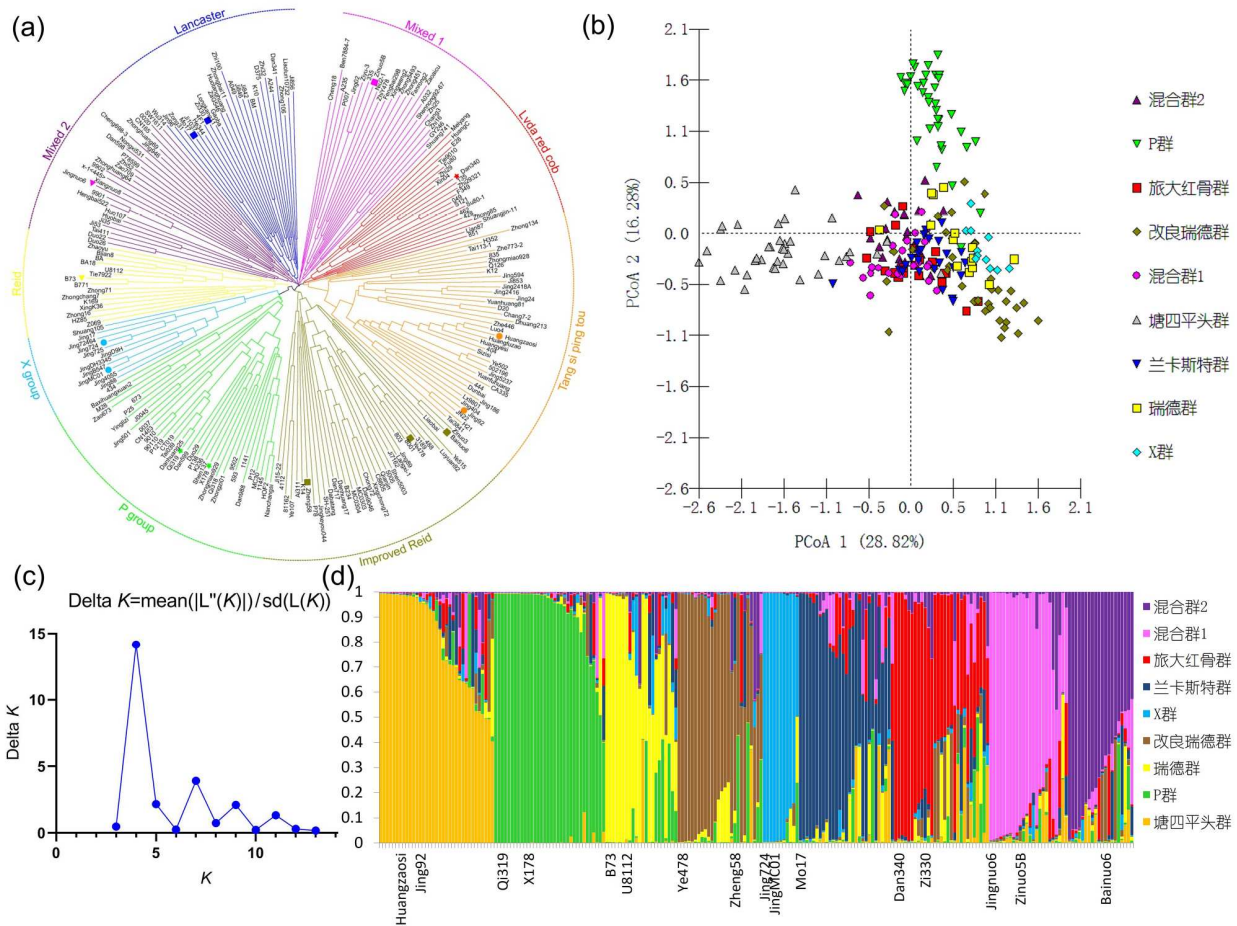


图2 230个玉米自交系的杂种优势群划分。(a) 基于邻接法的系统发生树; (b) 基于主坐标分析的散点图; (c) 基于Structure Harvester的 ΔK 折线图; (d) 基于Structure的群体结构图($K=9$)

Figure 2 Heterosis group classification of 230 maize inbred lines. (a) Phylogenetic tree base on neighbor-joining; (b) scatter plot based on principal coordinate analysis; (c) ΔK line chart based on Structure Harvester; (d) population structure bar chart based on Structure while K equals to 9

优势群与遗传背景信息一致. 与135个自交系的系谱信息比较, 有123个系的系谱信息和分群结果吻合, 准确率为91.1%. 将邻接树与群体结构的样品及结果进行一一对比, 两者分群结果一致率为73.9%, 以邻接树为参考, 结果有分歧的样品主要集中在改良瑞德群、混合群1和混合群2, 样品数量分别为12, 9, 20个, 而以群体结构为参考, 则集中在兰卡斯特群、旅大红骨群、混合群1和混合群2, 样品数量分别为11, 11, 10, 15个.

本文选择Structure的分群结果进入下一环节分析, 其中以黄早四和京92为代表的塘四平头群有35个系, 京724和京MC01为代表的X群有11个系, 齐319和X178为代表的P群有34个系, 自330和丹340为代表的旅大红骨群有29个系, Mo17为代表的兰卡斯特群有29个系,

B73和U8112为代表的瑞德群有22个系, 掖478和郑58为代表的改良瑞德群有26个系, 京糯6和紫糯5B为代表的混合群1有24个系, 白糯6为代表的混合群2有20个系. 分析不同杂种优势群的遗传多样性指数, 9个群的平均值为0.580, 其中混合群2群内遗传变异程度最高, 其值为0.700, X群群内遗传变异程度最低, 其值为0.360.

2.3 虚拟后代对杂优模式权重库的评估

9个杂种优势群进行两两之间不区分正反交且不重复的自由组合可以设计36个杂种优势模式, 按照设定的杂种优势模式模拟杂交获得23290个FGVH指纹数据. 对杂优模式权重库中不同模式的FGVH评分情况进行描述性统计(图3), 虽然所有模式评分的直方图都表

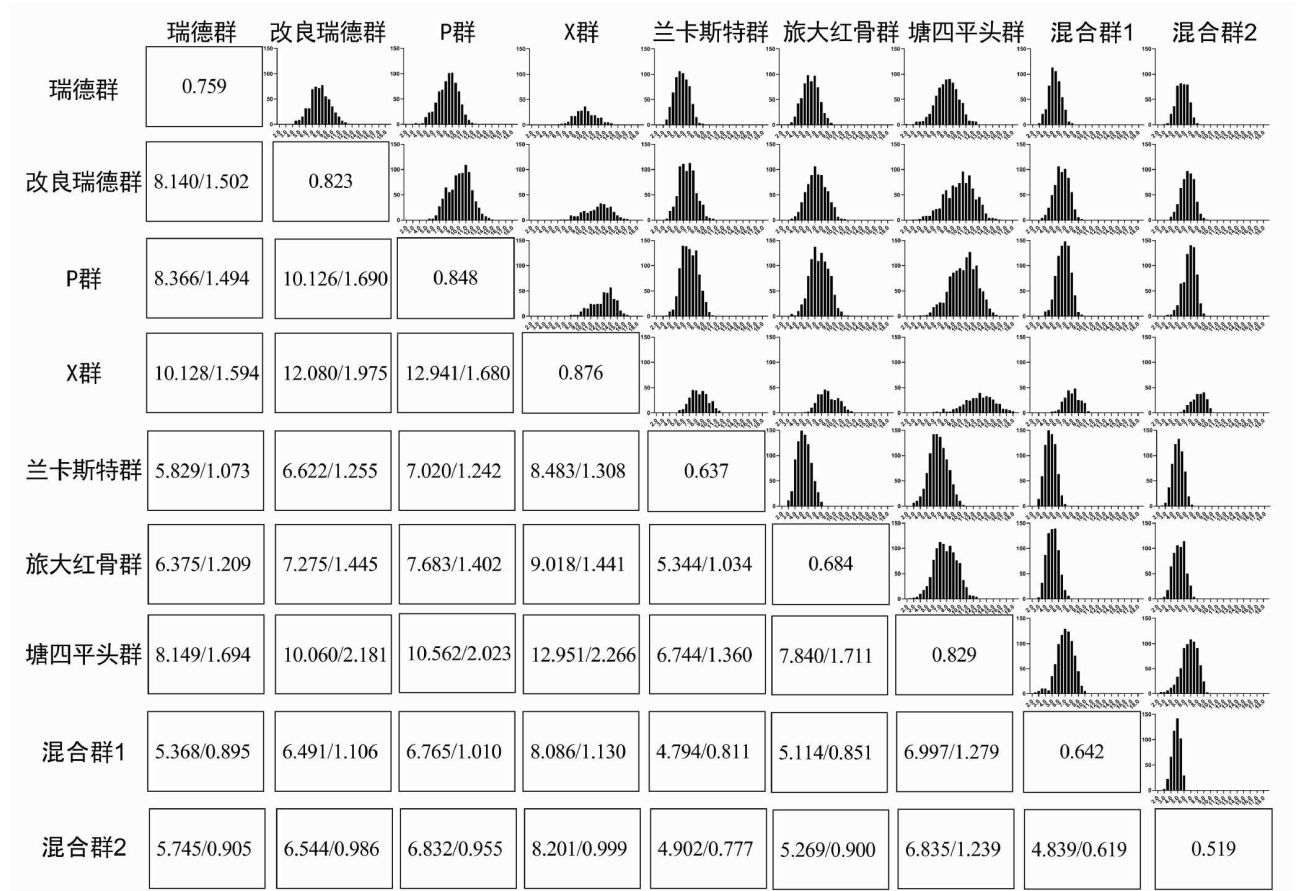


图3 杂种优势模式权重库对不同模式第一代虚拟杂交种评分的描述性统计. 右上角中的直方图为其对应杂种优势模式设计的虚拟后代评分的频数分布图, 横坐标为杂种优势模式评分的分组区间, 纵坐标为每个区间出现的样品数量; 对角线上的数值为杂种优势群在其设计杂种优势模式的准确率; 左下角的数值为其对应杂种优势模式设计的虚拟后代评分的平均值和标准偏差

Figure 3 Descriptive statistics of the first generation virtual hybrids score of different patterns in the weight library of heterosis patterns. The histogram in the upper right corner is the frequency distribution of the virtual progeny score designed to correspond to the heterosis patterns. The horizontal axis represents the score intervals of the heterosis pattern, and the vertical axis represents the number of samples that appear in each interval. The values on the diagonal are the accuracy of the heterosis groups in their designed heterosis patterns. The values in the lower left corner are the mean and standard deviations of the virtual progeny scores designed for their heterosis patterns

现为单峰且基本对称，但经过Kolmogorov-Smirnov检验服从正态分布的模式只有20个。不同模式的数据离散程度有显著差异，表现为不同模式的标准偏差有较大差异，包含塘四平头群模式的标准偏差普遍较大，其中X群×塘四平头群模式的标准偏差最大，包含混合群2模式则反之，其中混合群1×混合群2模式的标准偏差最小。另外，不同模式的评分平均值差异也很显著，不同模式评分平均值差值最大为8.112，其中评分平均值大于12的模式分别为X群×塘四平头群、X群×P群、改良瑞德群×X群，评分平均值小于5的模式分别为兰卡斯特群×混合群1、兰卡斯特群×混合群2、混合群1×混合群2。

虽然最简单的判定规则是直接输出模式评分最大值作为推测结果，但是不同杂种优势模式的评分值频数分布有显著差异，会对该规则带来误判影响。因此本文设计了三种判定规则并采用FGVH作为检验数据对杂优模式权重库的准确率进行评估：一是得分最高且不做阈值判定；二是得分最高且在对应模式阈值范围内；三是得分减去两种模式均值之差的最大值。三种规则总体的准确率分别为67.8%，73.5%和71.3%，对比每个模式的三种规则的判定结果差异(表1)，规则一虽然对平均分偏高和居中的模式的鉴定效果极佳，但对平均分偏低模式出现很大程度的误判；规则二通

过对不同模式的分布区间进行限定，对绝大部分模式的判定结果进行一定程度的有效修正；规则三通过对不同模式的平均分差值进行标准化，虽然能对平均分偏低模式进行很大程度的修正，但同时也对平均分居中的模式进行很大程度的过度矫正，因此后续分析均采用第二种判定规则。依据该规则的判定结果，准确率最高的杂优模式是改良瑞德群×P群，最低是瑞德×混合群2，不同杂种优势群的准确率情况具体如图3对角线所示，其中X群的准确率最高、混合群2最低。

2.4 待测杂交种的杂种优势模式

借助230个玉米自交系构建的杂种优势模式权重库，基于完全分子标记法推测335个玉米杂交种的杂种优势模式情况(图4(a))，共发现32种模式，其中有2个品种不符合判定规则未输出其结果。以杂交种的杂种优势模式为分析单位，改良瑞德群×塘四平头群是品种数量最多的模式，共有56个玉米品种，其代表性品种是郑单958、浚单20、鲁单9002等；以农大108为代表的32个玉米品种属于改良瑞德群×P群，以农大3138为代表的32个玉米品种属于P群×旅大红骨群，这两种模式的品种数量并列第二。以亲本的杂种优势群为分析单位，改良瑞德群覆盖的品种数量最多，共有 160个玉米品

表 1 三种判定规则对第一代虚拟杂交种杂种优势模式鉴定的准确率

Table 1 The accuracy rates of three judgment rules in indentifying the heterosis patterns of first generation virtual hybrids

杂种优势模式	规则一	规则二	规则三	杂种优势模式	规则一	规则二	规则三
P群×X群	1.000	0.968	0.717	混合群1×混合群2	0.319	0.398	0.817
P群×混合群1	0.702	0.784	0.714	兰卡斯特群×混合群1	0.455	0.519	0.901
P群×混合群2	0.547	0.612	0.721	兰卡斯特群×混合群2	0.272	0.367	0.850
P群×兰卡斯特群	0.726	0.777	0.702	兰卡斯特群×旅大红骨群	0.408	0.532	0.812
P群×旅大红骨群	0.729	0.816	0.670	兰卡斯特群×塘四平头群	0.668	0.710	0.757
P群×塘四平头群	0.990	0.958	0.664	旅大红骨群×混合群1	0.443	0.536	0.851
X群×混合群1	0.826	0.864	0.773	旅大红骨群×混合群2	0.336	0.450	0.778
X群×混合群2	0.605	0.732	0.727	旅大红骨群×塘四平头群	0.801	0.843	0.648
X群×兰卡斯特群	0.796	0.824	0.677	瑞德群×P群	0.844	0.888	0.671
X群×旅大红骨群	0.727	0.818	0.705	瑞德群×X群	0.855	0.888	0.727
X群×塘四平头群	1.000	0.961	0.694	瑞德群×改良瑞德群	0.794	0.843	0.619
改良瑞德群×P群	0.993	0.983	0.554	瑞德群×混合群1	0.407	0.513	0.795
改良瑞德群×X群	0.979	0.951	0.608	瑞德群×混合群2	0.284	0.366	0.741
改良瑞德群×混合群1	0.675	0.731	0.675	瑞德群×兰卡斯特群	0.567	0.660	0.803
改良瑞德群×混合群2	0.531	0.615	0.694	瑞德群×旅大红骨群	0.586	0.679	0.737
改良瑞德群×兰卡斯特群	0.643	0.710	0.670	瑞德群×塘四平头群	0.766	0.810	0.692
改良瑞德群×旅大红骨群	0.649	0.797	0.625	塘四平头群×混合群1	0.751	0.788	0.732
改良瑞德群×塘四平头群	0.960	0.951	0.615	塘四平头群×混合群2	0.533	0.613	0.710

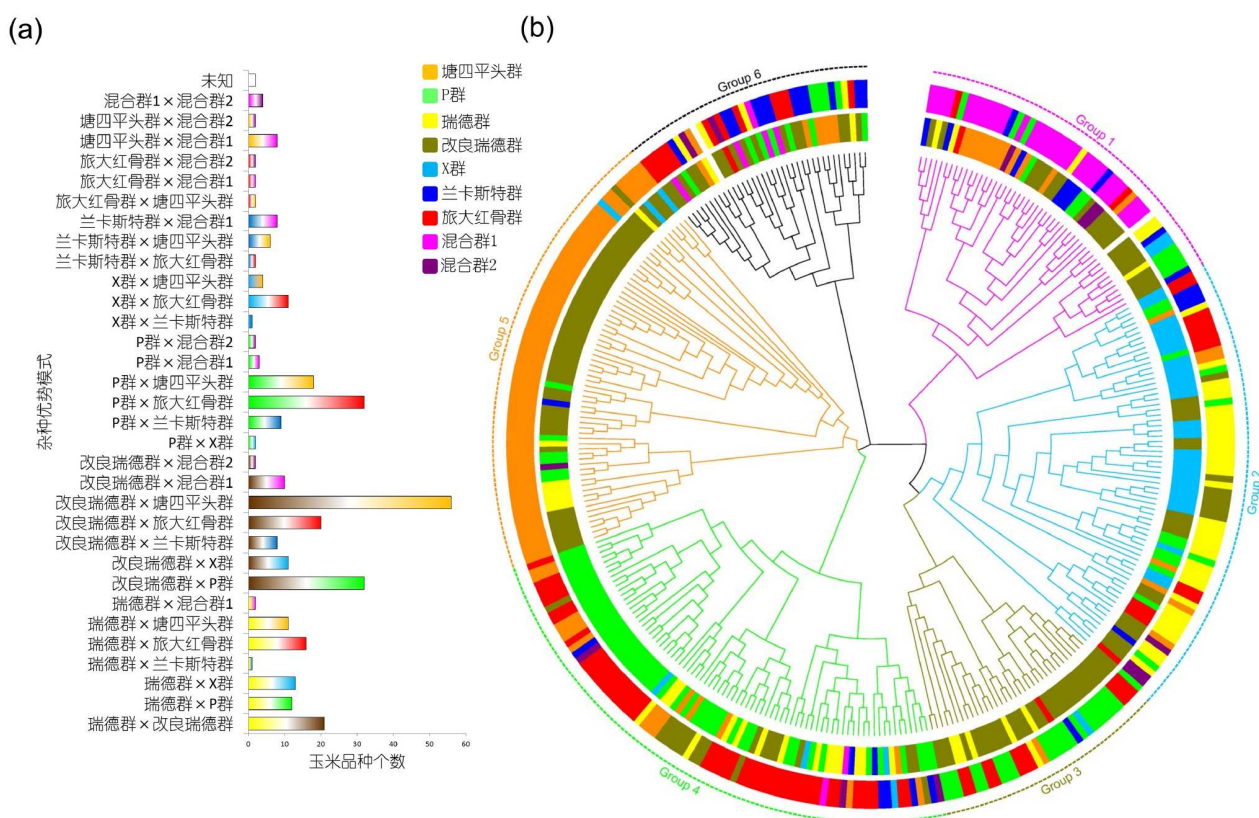


图 4 335个玉米杂交种基于完全分子标记法的杂种优势模式和基于邻接法的系统发生树分组情况。(a) 不同杂种优势模式的品种数量; (b) 系统发生树拓扑图及其对应的杂种优势模式, 其中外圈两条色带代表该树分支对应品种基于完全分子标记法获得的两个亲本杂种优势群, 内圈树枝颜色代表该树分支对应品种基于邻接法获得的组别

Figure 4 Heterosis patterns based on the pure molecular marker method and neighbor-joining phylogenetic tree grouping of 335 maize hybrids. (a) Number of varieties with different heterosis patterns; (b) phylogenetic tree topology diagram and corresponding heterosis patterns, where the two outer color circles represent the two heterosis groups of the parents of the varieties of the corresponding tree branches obtained by the pure molecular marker method, and the color of the inner circle branches represents the group of the corresponding tree branches obtained by the neighbor-joining method

种; P群排名第二, 有 110个玉米品种的亲本属于该群; 排名第三是塘四平头群, 涉及107个玉米品种。以2017年玉米主导品种的全国种植面积数据为参考^[22], 推广面积排名前三的杂种优势模式依次为改良瑞德群×塘四平头群、瑞德群×X群和X群×塘四平头群。

为验证结果准确性, 本文参考玉米品种审定公告中的品种来源信息, 选择其亲本与对照自交系中同名的99个样品进行二联体或三联体比较, 结果表明, 系谱类群与推测模式一致的样品有72个, 涵盖16种杂种优势模式。基于邻接法系统发生树分析, 所有杂交种SSR指纹可以划为6个分组, 该分组结果与其对应的杂种优势模式结果存在一定联系(图4(b)), 系统发生树外围的两圈彩色条柱代表每个树叶代表的杂交种的两个亲本对应的杂种优势群, 而相同树叶对应两个条柱的组合

则构成其杂种优势模式。第一分组的品种亲本以混合群1为主, 第二分组的品种亲本以X群或瑞德群为主, 第三分组以改良瑞德群×P群模式的品种为主, 第四分组以P群×旅大红骨群模式的品种为主且品种亲本以旅大红骨群或P群为主, 第五分组以改良瑞德群×塘四平头群模式的品种为主且品种亲本以塘四平头群为主, 第六分组的品种亲本以兰卡斯特群为主。从品种类型来看, 糯玉米品种集中在第一分组, 甜玉米和爆裂玉米品种集中在第六分组, 普通玉米品种分布在第二、三、四、五分组里。基于主坐标分析的杂交种散点图选取了排名前三的向量进行组合展示(图5), 其解释率分别为31.46%, 19.36%和7.81%。在第一和第二向量中(图5(a)), 只有改良瑞德群×塘四平头群模式的品种能显著区别于其他模式。在第一和第三向量中(图5(b)), 增加了对

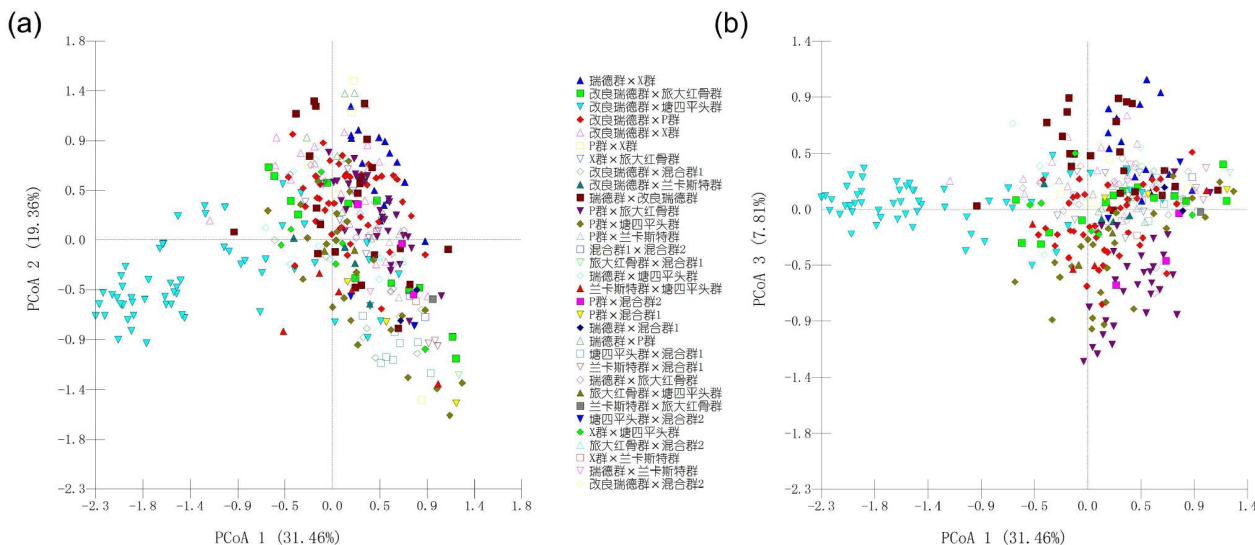


图 5 335个玉米杂交种基于主坐标分析的散点图及其对应品种基于完全分子标记法的杂种优势模式。(a) 第一坐标和第二坐标的散点图；(b) 第一坐标和第三坐标的散点图

Figure 5 Scatter plot of 335 maize hybrids based on principal coordinate analysis and their corresponding heterosis patterns based on pure molecular marker method. (a) Scatter plot for first and second coordinates; (b) scatter plot for first and third coordinates

P群×旅大红骨群、P群×塘四平头群、瑞德群×改良瑞德群、瑞德群×X群四种模式的区分度。

3 讨论

在已报道的研究中，无论是正向推导还是反向推导^[8-10,23]，玉米杂交种的杂种优势模式推测主要参考其品种来源信息，并且该信息只能依赖于育种单位提供的文字资料。本文开发的“完全分子标记法”杂种优势模式鉴定方法，只需提供玉米杂交种的DNA指纹便可推测该品种两个亲本对应的杂种优势群，打破了杂种优势模式信息来源单一的瓶颈，为追溯玉米杂交种遗传背景信息提供了一种客观、简便、有效的技术手段。国家玉米品种试验引入DNA指纹检测作为品种身份鉴定的判断依据已有近二十年^[24]，在很大程度上有效避免了新品种与老品种之间的趋同化选育。若引入本文提出的鉴定方法，则可以提升到杂种优势群层次引导育种家进行异质化育种，进而通过确保栽培品种的遗传背景多样性来保障国家粮食安全。

玉米自交系的杂种优势群是种质资源收集和品种选育等领域研究的重点，由于研究人员收集的材料存在差异性、自交系选用的育种方式不同、材料所处的育种阶段不同等因素，对中国玉米杂种优势群的命名、划分方法、类群数量和部分材料归属均难以达成统一共识。随着玉米杂种优势群划分研究的不断深入，塘四

平头群、旅大红骨群、P群、兰卡斯特群、瑞德群、改良瑞德群和X群已经成为国内公认存在的7个杂种优势群^[4-7,11,25-28]。本文筛选的对照自交系不仅包含了上述7个杂种优势群，而且对参照系的分类与前人研究一致^[7,29]。结合2.2节和2.3节的分析结果，包含塘四平头群模式的评分离散程度较大，和该群作为国内特有的杂种优势群并得到较多人工选择和改良有一定关系。另外本研究还包括了混合群1和混合群2，这两个群覆盖了糯质、早熟、苏湾等多种特定材料，使得这两个群的遗传背景较为复杂，进而造成对应的杂优模式权重库的分值偏低，影响包含这两个群的杂种优势模式评估。因此，对照自交系需要随着待测杂交种的种质来源不断扩充，增强其有待测杂交种的联系；鉴定方法中的判定规则也有待进一步完善，以便减少高分模式对低分模式的判定干扰。

为了验证杂种优势模式鉴定方法的有效性，本文采取三类方法进行验证。首先，选取对照自交系中有同名亲本的玉米品种进行“系谱结合分子标记法”分析验证，与“完全分子标记法”一致率为77.7%；其次，采用系统发生树和主坐标分析两种方法对待测杂交种的DNA指纹数据进行分类，发现邻接法的分组结果与“完全分子标记法”的鉴定结果在杂种优势模式和单亲杂种优势群两个层次上高度吻合，主要杂种优势模式的划分也在主坐标结果上得到验证；最后，选取有代表性的待

测杂交种与同类研究进行验证,与“完全分子标记法”结果基本一致,例如浚单20属于改良瑞德群×塘四平头群^[30]、京科968和京农科728均属于X群×塘四平头群^[31]、鲁单981属于P群×塘四平头群^[32]等。在“完全分子标记法”的通用性方面,虽然本文只验证了玉米的有效性,但理论上该方法没有限定研究对象所属的物种,即对存在杂种优势模式且采用单交模式的其他农作物均可采用本方法进行杂种优势模式鉴定。

4 结论

本文提出了一种玉米单交种杂种优势模式鉴定方法“完全分子标记法”。首先,筛选对照自交系进行杂种

优势群划分;其次,通过虚拟杂交对照自交系构建杂优模式权重库,并确定每个杂优模式的置信阈值;最后,依据待测玉米单交种的DNA指纹判定该品种的杂种优势模式。选取335个国家审定玉米品种对该方法进行验证,显示2000-2013年国家审定玉米品种利用最多且推广面积最大的杂种优势模式是改良瑞德群×塘四平头群。为了提高完全分子标记法的准确度,可通过补充代表性对照自交系来提高杂优模式权重库的代表性,并进一步研究本方法鉴定精度的影响因素。本方法目前只适用于杂交模式为单交的农作物,其他杂交模式的鉴定有待下一步研究。本研究可为追溯农作物单交种遗传背景提供客观且有效的分析方案。

参考文献

- 1 Zhan W, Yuan M Q, Xing Y Z. Progress in understanding molecular genetic basis of heterosis in rice (in Chinese). *Chin Sci Bull*, 2016, 6: 3842–3849 [湛蔚, 袁梦琦, 邢永忠. 水稻杂种优势遗传与分子基础. 科学通报, 2016, 61: 3842–3849]
- 2 Aslam M, Zafar S A. Heterotic group theory: a thriving vitality in hybrid maize breeding. *J Agric Basic Sci*, 2020, 05: 45–61
- 3 Liu J L. Maize Breeding Science (in Chinese). Beijing: China Agriculture Press, 2002 [刘纪麟. 玉米育种学. 北京: 中国农业出版社, 2002]
- 4 Sun Q, Li W C, Zhang F J, et al. Analysis on the pedigree of the parental lines of the maize hybrids approved by the nation from 2001 to 2012 (in Chinese). *J Maize Sci*, 2014, 22: 6–11 [孙琦, 李文才, 张发军, 等. 2001~2012年国审玉米品种亲本自交系系谱来源分析. 玉米科学, 2014, 22: 6–11]
- 5 Liu Z Z, Wu X, Liu H L, et al. genetic diversity and population structure of important chinese maize inbred lines revealed by 40 core simple sequence repeats (SSRs)(in Chinese). *Sci Agric Sin*, 2012, 45: 2107–2138 [刘志斋, 吴迅, 刘海利, 等. 基于40个核心SSR标记揭示的820份中国玉米重要自交系的遗传多样性与群体结构. 中国农业科学, 2012, 45: 2107–2138]
- 6 Zhao J R, Li C H, Song W, et al. Genetic diversity and population structure of important chinese maize breeding germplasm revealed by SNP-Chips (in Chinese). *Sci Agric Sin*, 2018, 51: 626–644 [赵久然, 李春辉, 宋伟, 等. 基于SNP芯片揭示中国玉米育种种质的遗传多样性与群体遗传结构. 中国农业科学, 2018, 51: 626–644]
- 7 Zhang R, Xu G, Li J, et al. Patterns of genomic variation in Chinese maize inbred lines and implications for genetic improvement. *Theor Appl Genet*, 2018, 131: 1207–1221
- 8 Peng Z B, Liu X Z, Fu J H, et al. Heterotic group and patterns of maize inbred lines established by RAPDs (in Chinese). *Acta Agron Sin*, 1998, 24: 711–717 [彭泽斌, 刘新芝, 傅骏骅, 等. 玉米自交系杂种优势类群与杂优模式构建的初步研究. 作物学报, 1998, 24: 711–717]
- 9 Liu Z H, Wang Q D, Zhao W Y, et al. Classification of heterotic group of elite maize inbred lines. *Acta Agric Boreali-Sin*, 2005, 20: 49–53 [刘宗华, 王庆东, 赵武英, 等. 优良玉米自交系的杂种优势类群划分研究. 华北农学报, 2005, 20: 49–53]
- 10 Wang Y B, Wang Z H, Wang Y P, et al. Studies on the heterosis utilizing models of main maize germplasms in China (in Chinese). *Sci Agric Sin*, 1997, 30: 17–25 [王懿波, 王振华, 王永善, 等. 中国玉米主要种质杂交优势利用模式研究. 中国农业科学, 1997, 30: 17–25]
- 11 Teng W T, Cao J S, Chen Y H, et al. Analysis of maize heterotic groups and patterns during past ecade in China (in Chinese). *Sci Agric Sin*, 2004, 37: 1804–1811 [滕文涛, 曹靖生, 陈彦惠, 等. 十年来中国玉米杂种优势群及其模式变化的分析. 中国农业科学, 2004, 37: 1804–1811]
- 12 Zhao J R, Liu L Z, Wang F G, et al. Identification authenticity of female parent of F_1 hybrids through SSR profiling of pericarp tissue in hybrid (in Chinese). *J Maize Sci*, 2004, 12: 6–8, 12 [赵久然, 刘龙洲, 王风格, 等. 利用杂交玉米 F_1 代种子果皮组织鉴定母本真实性的SSR研究. 玉米科学, 2004, 12: 6–8, 12]
- 13 Wang C X, Li X H, He Z F, et al. The rapid identification of maternal lines of maize hybrids with SSR based technology (in Chinese). *J Maize Sci*, 2007, 15: 9–12 [王春喜, 李新海, 郝转芳, 等. 玉米杂交种母本自交系的SSR快速鉴定技术. 玉米科学, 2007, 15: 9–12]
- 14 Wang S N, Wang S K, Zhao C P, et al. General situation of chinese variety certification system and comparison with domestic and foreign countries (in Chinese). *Sci Technol Cer Oils Foods*, 2024, 32: 211–218 [王赛楠, 王术坤, 赵翠萍, 等. 中国品种审定制度概况及国内外比较. 粮油食品科技, 2024, 32: 211–218]
- 15 Wang F G, Zhao J R, Yang Y. SSR Fingerprint Maps of Nationally Approved Maize Varieties (in Chinese). Beijing: China Agricultural Science and Technology Press, 2015 [王风格, 赵久然, 杨扬. 国家审定玉米品种SSR指纹图谱. 北京: 中国农业科学技术出版社, 2015]

- 16 Wang F G, Zhao J R, Yang Y, et al. SSR Fingerprint Maps of Commonly Used Maize Inbred Lines (in Chinese). Beijing: China Agricultural Science and Technology Press, 2020 [王凤格, 赵久然, 杨扬, 等. 常有玉米自交系SSR指纹图谱. 北京: 中国农业科学技术出版社, 2020]
- 17 Wang F G, Yang Y, Yi H M, et al. Construction of an SSR-based standard fingerprint database for corn variety authorized in China (in Chinese). *Sci Agric Sin*, 2017, 50: 1–14 [王凤格, 杨扬, 易红梅, 等. 中国玉米审定品种标准SSR指纹库的构建. 中国农业科学, 2017, 50: 1–14]
- 18 Liu K, Muse S V. PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics*, 2005, 21: 2128–2129
- 19 Subramanian B, Gao S, Lercher M J, et al. Evolview v3: a webserver for visualization, annotation, and management of phylogenetic trees. *Nucleic Acids Res*, 2019, 47: W270–W275
- 20 Pritchard J K, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*, 2000, 155: 945–959
- 21 Earl D A, vonHoldt B M. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genet Resour*, 2012, 4: 359–36
- 22 Committee N C V A. National Main Crop Variety Promotion and Application Report in 2017 (in Chinese). Beijing: China Agricultural Science and Technology Press, 2019 [国家农作物品种审定委员会. 2017年全国主要农作物品种推广应用报告. 北京: 中国农业科学技术出版社, 2019]
- 23 Wang Y B, Wang Z H, Lu L X, et al. Studies on maize germplasm base, division of heterosis groups and utilizing models of heterosis in China (in Chinese). *Maize Sci*, 1998, (1): 9–13 [王懿波, 王振华, 陆利行, 等. 中国玉米种质基础、杂种优势群划分与杂优模式研究. 玉米科学, 1998, (1): 9–13]
- 24 Yi H M, Ren J, Wang L, et al. DNA Fingerprinting and genetic diversity analysis of national maize regional trials in 2014–2019 (in Chinese). *Acta Agric Bor Sin*, 2020, 35: 87–93 [易红梅, 任洁, 王璐, 等. 2014~2019年国家玉米区域试验参试组合DNA指纹检测及遗传多样性分析. 华北农学报, 2020, 35: 87–93]
- 25 Hou M. Pan-genome assembly of maize founder inbred lines and analysis of heterosis mechanism (in Chinese). Doctor Dissertation. Hohhot: Inner Mongolia University, 2023 [侯美. 玉米骨干亲本自交系泛基因组组装及杂种优势机理解析. 博士学位论文. 呼和浩特: 内蒙古大学, 2023]
- 26 Li C, Guan H, Jing X, et al. Genomic insights into historical improvement of heterotic groups during modern hybrid maize breeding. *Nat Plants*, 2022, 8: 750–763
- 27 Shu G, Cao G, Li N, et al. Genetic variation and population structure in China summer maize germplasm. *Sci Rep*, 2021, 11: 8012
- 28 Li X H, Yuan L X, Li X H, et al. Heterotic grouping of 70 maize inbred lines by SSR markers (in Chinese). *Sci Agric Sin*, 2003, 36: 622–627 [李新海, 袁力行, 李晓辉, 等. 利用SSR标记划分70份我国玉米自交系的杂种优势群. 中国农业科学, 2003, 36: 622–627]
- 29 Wang B, Hou M, Shi J, et al. De novo genome assembly and analyses of 12 founder inbred lines provide insights into maize heterosis. *Nat Genet*, 2023, 55: 312–323
- 30 Song W, Zhao J R, Wang F G, et al. Analysis on heterosis pattern of xundan20 by SSR markers (in Chinese). *J Maize Sci*, 2011, 19: 43–47 [宋伟, 赵久然, 王凤格, 等. 利用SSR标记对浚单20进行杂优模式分析. 玉米科学, 2011, 19: 43–47]
- 31 Zhao J R, Li C H, Song W, et al. Elaboration of heterotic pattern in a series of maize varieties by SSR markers (in Chinese). *J Maize Sci*, 2017, 25: 1–8 [赵久然, 李春辉, 宋伟, 等. 利用SSR标记解析京科968等系列玉米品种的杂优模式. 玉米科学, 2017, 25: 1–8]
- 32 Li W C, Liu Z X, Sun Q, et al. Utilization and continuous improvement of “PN78599 Group×Tangsipingtou Group” maize hybrid model (in Chinese). *Shandong Agric Sci*, 2019, 51: 31–34 [李文才, 刘治先, 孙琦, 等. “PN78599群×塘四平头群”玉米杂优模式的利用及连续改良. 山东农业科学, 2019, 51: 31–34]

Summary for “基于分子标记技术的玉米品种杂种优势模式鉴定方法”

Heterosis pattern detection method of maize single-cross hybrids via molecular marker

Yang Yang[†], Hongli Tian[†], Hongmei Yi[†], Lu Wang, Yaming Fan, Shihui Chang, Fengge Wang^{*}, Wei Song^{*} & Jiuran Zhao^{*}

Maize Research Institute, Beijing Academy of Agriculture and Forestry Sciences; Key Laboratory of Crop DNA Fingerprinting Innovation and Utilization of the Ministry of Agriculture and Rural Affairs (Co-Construction by Ministry and Province); Beijing Key Laboratory of Maize DNA Fingerprinting and Molecular Breeding, Beijing 100097, China

[†] Equally contributed to this work

^{*} Corresponding authors, E-mail: maizezhao@126.com; songwei@baafs.net.cn; wangfengge@baafs.net.cn

Single-cross hybrids are the main maize variety types in national and provincial authorization. As the reflection of the genetic background of maize varieties, the heterosis pattern of single-cross hybrids can only be identified from either their pedigree or the heterosis group classification of their parental inbred lines for a long time. When the commercial breeding era comes with modern agriculture, the traditional heterosis pattern detection method has encountered with difficulties such as unknown pedigree information or confidential parental inbred lines. Therefore, this study presented a heterosis pattern detection method purely based on molecular marker technology (referred to as “the pure molecular marker method”), which is capable of detecting the heterosis groups of two parents of the single-cross hybrid by analyzing its DNA fingerprint, with no pedigree information ever needed. Firstly, we screened out the control inbred lines and classified the heterosis groups. Secondly, we used virtual hybridization of the control inbred lines to construct the heterosis pattern weight table and then determined the confidence threshold of each heterosis pattern. Finally, we decided the heterosis pattern of the variety according to the DNA fingerprints of the maize single-cross hybrids. The results showed that (1) through cross-verification by three cluster methods, a training inbred line library with 230 commonly used inbred lines was proved to contain 9 heterosis groups, i.e. Tangsipingtou (Huangzaosi improved lines), X, Lvda red cob, P, Lancaster, Reid, Improved Reid, Mixed 1 and Mixed 2; (2) a heterosis pattern weight table composed of 36 patterns was constructed, where the decision accuracy could be improved by using the confidence threshold of each pattern as the criterion rule, due to the fact that the rating of each pattern has shown quite different distribution; (3) 335 national authorized varieties were selected for method validation, and it was found that the results of the pure molecular marker method were not only consistent with 77.7% results of the genealogy-combined molecular marker method, but also coincide with the group results of the phylogenetic tree in terms of parent group aggregation and parental composite pattern aggregation. (4) The most commonly used heterosis patterns (regardless of reciprocal crosses) in nationally authorized maize varieties from 2000 to 2013 were Improved Reid×Tangsipingtou, Improved Reid×P, and P×Lvda red cob, with the Improved Reid×Tangsipingtou pattern most widely planted in 2017. This study provides an objective, simple, and effective analysis method to acquire the genetic background of the maize single-cross hybrids, which not only helps to accurately promote the cultivated varieties and assess their genetic diversity but also plays a significant role in ensuring national food security.

maize, single-cross hybrids, simple sequence repeat marker, heterosis group, heterosis pattern

doi: [10.1360/TB-2025-0001](https://doi.org/10.1360/TB-2025-0001)