

DOI: 10.3724/SP.J.1006.2026.53084

适于玉米品种鉴定的一套三等位变异 SNP 新型标记组合

田红丽** 杨 扬** 范亚明** 易红梅 郭丹丹 王风格* 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

摘 要: 高通量测序、生物信息学以及基因型分型技术的快速发展, 为农作物品种的精准鉴定开辟了新的方法路径。本文基于全基因组水平挖掘的 78 万个初始多等位变异 SNP 位点, 利用生物信息、统计分析、靶向测序验证以及最优遗传算法等方法评估确定了一套三等位变异 SNP 新型标记组合。该组合包含的 40 个位点均匀分布于玉米的 10 对染色体上, 每对染色体为 4 个位点。利用 200 个杂交种和 270 个自交系的基因型数据对 40 个 SNP 位点进行多角度评估, 结果显示: (1) 对于杂交种和自交系, 所有位点的 PIC 值均高于二等位变异 SNP 位点的最高值 0.375, 平均值为 0.470 和 0.480, DP 平均值为 0.68 和 0.57。(2) 利用 40 个三等位和 96 个二等位变异 SNP 标记的杂交种数据进行品种的两两成对比较, 显示差异位点百分比分别集中在 50%~85% 和 50%~70%; 自交系数据分析显示, 品种间差异分别集中在 35%~75% 和 35%~60%; 三等位与二等位变异 SNP 标记相比, 品种间差异位点百分比平均值提高了 10%。(3) 对于杂交种和自交系, 40 个 SNP 位点的第一等位变异频率平均值分别为 0.58 和 0.57, 第二等位变异平均值均为 0.31, 第三等位变异平均值分别为 0.11 和 0.12。(4) 聚类结果显示, 40 个位点将 270 个自交系精准划分为 10 个杂种优势类群, 与高密度芯片划分结果一致。综上所述, 本研究报道了一套高质量、高鉴别力的三等位变异核心 SNP 标记位点组合, 为玉米品种鉴定提供了一种新型标记类型; 凭借其更高的变异解析能力, 有望与现有标记类型形成互补, 推动农作物分子鉴定技术的发展。

关键词: 玉米; 品种鉴定; SNP 位点; 三等位变异; 核心位点组合

A novel set of tri-allelic variant SNP loci suitable for maize variety identification

Tian Hong-Li**, Yang Yang**, Fan Ya-Ming**, Yi Hong-Mei, Guo Dan-Dan, Wang Feng-Ge*, and Zhao Jiu-Ran*

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

Abstract: The rapid development of high-throughput sequencing, bioinformatics, and genotyping technologies has opened new avenues for the precise identification of crop varieties. In this study, based on 780,000 initially mined multi-allelic variant SNPs from the maize whole genome, a novel set of 40 tri-allelic SNP markers was evaluated and identified using bioinformatics approaches, statistical analysis, targeted sequencing validation, and optimal genetic algorithms. These 40 SNPs are evenly distributed across the 10 maize chromosome pairs, with four markers per chromosome. A comprehensive evaluation and comparison of the 40 SNPs were conducted using genotype data from 200 hybrids and 270 inbred lines. The results showed that: (1) For both hybrids and inbred lines, the PIC values of all loci exceeded the maximum value of 0.375 typical for bi-allelic SNP loci, with average values of 0.470 and 0.480, and average discrimination power (DP) values of 0.68 and 0.57, respectively. (2) Pairwise

本研究由国家科技重大专项(2022ZD04019)和北京市农林科学院科技创新能力建设专项(KJ CX20230303)资助。

This study was supported by the National Scientific and Technological Innovation-Major Projects (2022ZD04019) and the Science and Technology Innovation Capacity Building Project of BAAFS (KJ CX20230303).

* 通信作者(Corresponding authors): 赵久然, E-mail: maizezhao@126.com; 王风格, E-mail: gege0106@163.com

** 同等贡献(Contributed equally to this work)

第一作者联系方式: 田红丽, E-mail: tianhongli9963@163.com; 杨扬, E-mail: caurwx@163.com; 范亚明, E-mail: 13718078547@163.com

Received (收稿日期): 2025-10-29; Accepted (接受日期): 2026-01-22; Published online (网络出版日期): 2026-02-06.

URL: <https://link.cnki.net/urlid/11.1809.S.20260206.0946.003>

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

comparisons using the 40 tri-allelic and 96 bi-allelic SNP markers revealed that the percentage of differential loci among hybrids mainly ranged from 50% to 85% and 50% to 70%, respectively; for inbred lines, the ranges were 35% to 75% and 35% to 60%, respectively. The average percentage of differential loci among varieties increased by 10% for tri-allelic SNPs compared to bi-allelic SNPs. (3) For both hybrids and inbred lines, the average frequencies of the first, second, and third alleles at the 40 tri-allelic loci were 0.58 and 0.57, 0.31, and 0.11 and 0.12, respectively. (4) Clustering analysis showed that the 40 loci effectively classified the 270 inbred lines into 10 heterotic groups, consistent with classifications based on high-density SNP chip data. In summary, this study reports a high-quality, highly discriminative core set of tri-allelic SNP markers, providing a novel marker type for maize variety identification. With their enhanced ability to resolve genetic variation, these markers are expected to complement existing marker systems and contribute to the advancement of molecular identification technologies in crop breeding.

Keywords: maize; variety identification; SNP loci; tri-allelic; core loci set

高通量测序、生物信息学以及基因型分型技术的快速发展,为利用更新、更高效的分子技术手段鉴别农作物品种提供了可能。近二十年,基于 SSR 分子标记玉米品种分子鉴定系列标准陆续颁布^[1-3]。利用上述标准,构建了包含玉米已知品种及 40 个 SSR 标记的标准 DNA 指纹库^[4-5]。但是,随着种业的发展和检测需求的不断变化,SSR 标记鉴定技术已在多个方面呈现不足。一是在检测通量、平台兼容性等方面不能满足品种高通量大规模鉴定需求;二是分子检测除品种真实性、纯度鉴定的常规需求外,实质性派生品种鉴定、亲缘关系分析等精准溯源鉴定需求日益剧增;三是 SSR 指纹数据是通过电泳获得的片段长度,故难以实现跨实验室、跨平台的数据整合利用。SNP 是基于单个核苷酸多态性开发的标记位点,由于其具有在基因组上分布密度高、数据易整合共享、易实现高通量检测等特征,而被关注并作为品种分子鉴定的理想标记之一。玉米系列参考基因组的公布、大量重测序数据的释放、高通量分型平台的快速发展也为玉米优异 SNP 位点的开发,以及高效精准鉴定技术体系的构建奠定了良好基础^[6-11]。但是截至目前,应用在玉米等农作物品种分子鉴定中的 SNP 标记位点均为二等位变异位点,即 ba SNPs (bi-allelic SNPs),每个位点有 2 种等位变异、3 种基因型^[12-17]。二等位变异类型的 SNP 标记,其单个位点多态性相对较低。同时,由于仅有 2 种等位变异、3 种基因型,对于品种鉴定尤其是一致性不高的样品,在检测混合样品情况下存在基因型是纯合还是杂合无法判定的情况。玉米基因组中除了分布密度最高的二等位变异 SNP 类型之外,还分布有三等位、四等位以及五等位变异(4 种碱基和缺失)的多等位变异 SNP 位点,即 ma SNPs (multi-allelic SNPs)。单个位点每多一个等位变异其多态性呈倍数增加,比如,三等位变异 SNP 标记可以出现 3 个等位变异、6 种基因型。多等位变异 SNP 位点既具有

SSR 标记的单个位点多态性高,又具有 SNP 标记的数据易整合共享的特征;其拓展了传统 SNP 标记类型范围,提高了单个位点的多态性,有利于解决混合样品带来的基因型数据精准读取的问题;在与性状紧密关联的功能标记开发方面,多等位变异 SNP 位点有利于弥补与多等位变异 SNP 关联的标记位点^[18]。三等位或四等位变异的 SNP 标记位点已经被逐步开发并应用到人类司法鉴定或功能标记开发等领域^[19-22]。

本研究基于来源广泛、代表性强的 388 份玉米自交系的高质量重测序数据,首先从基因组中分布数量最多的三等位变异 SNP 标记位点出发,探索多等位变异 SNP 位点在玉米品种分子鉴定中的应用潜力。然后利用生物信息、统计分析、靶向测序验证等方法,在全基因组水平挖掘评估三等位变异 SNP 位点,并确定候选位点集合。此外,利用最优遗传算法确定了一套包含 40 个位点的三等位变异 SNP 核心位点组合,为玉米品种真实性、纯度以及亲子关系鉴定等提供了新型标记类型。

1 材料和方法

1.1 供试材料

评估 SNP 位点的玉米样品为具有代表性的 200 个杂交种和 270 个自交系。200 个杂交种为通过国家审定的主推杂交种,包含各区试组的对照品种和推广种植面积较大的品种。270 个骨干自交系为国内外各杂种优势群的代表材料^[11]。

1.2 总 DNA 提取

每个材料随机选取 30 个单株混合,采用改良 CTAB 法提取总 DNA,并除去 RNA^[15]。用 1% 的琼脂糖凝胶电泳和紫外分光光度计测定 DNA 质量和浓度(Nanodrop 2000, Thermo Scientific),工作液浓度统一调整为 100 ng μL^{-1} 。所制备的 DNA 无蛋白质、RNA 以及其他杂质污染,琼脂糖凝胶电泳显示, DNA 条带单一、无降解。

1.3 三等位变异 SNP 核心位点组合开发评估

适用于玉米品种鉴定的一套三等位变异 SNP 核心位点的全基因组挖掘评估过程如下。(1) 采用全球代表性的 388 个玉米自交系的全基因组高深度重测序数据^[11]作为原始测序数据进行分析。(2) 利用 BWA 和 GATK 软件进行序列比对、变异挖掘, 获得初始 ma SNP 位点集, 包含三等位、四等位和五等位 3 种类型位点。(3) 基于测序数据覆盖度及 GATK 软件评估每个位点的 Quality 值进行位点初筛, 获得优异位点集。(4) 从测序数据中抽取 388 个自交系的基因型评估位点的杂合度和多态性, 删除杂合度 > 0.4 且 PIC 值 < 0.1 的位点, 删除多拷贝位点和侧翼序列中含有 InDel 的位点; 基于物理和遗传图谱在全基因组均匀选取候选位点, 获得用于靶向测序验证的候选位点集。(5) 根据靶向测序要求选取侧翼序列 100 bp 的 GC 含量接近 50% 的位点, 优化设计每个位点的探针序列, 形成一套多等位变异的 SNP 标记集。(6) 基于靶向测序技术将上述位点合成探针, 并在测序平台上对 970 份广泛选取的玉米自交系进行分型分析。分别根据测序数据覆盖度、数据缺失率、杂合基因型频率、PIC 值等参数进行严格评估, 确定三等位变异 SNP 候选位点组合。(7) 利用 200 个审定杂交种和 270 个代表性自交系的数据, 根据最优遗传算法^[23-25]和相对均匀分布原则, 确定一套三等位变异 SNP 核心位点组合。(8) 设计上述候选位点组合的扩增引物, 选取京科 968、郑单 958、先玉 335 等推广种植面积较大的杂交种和代表不同杂优类群的自交系, 共计 14 个品种, 利用一代 Sanger 测序对三等位变异 SNP 位点进行验证, 并比对侧翼序列。

1.4 数据分析

数据分析对象为: (1) 基于靶向测序分型平台获得的数据, 选取 200 个国家审定的杂交种和 270 个代表性自交系的三等位变异 SNP 候选位点和核心位点组合的基因型数据; (2) 从玉米二等位变异 SNP-DNA 指纹数据库中抽取上述品种的 96 个 SNP 位点的基因型数据^[14-15], 这 96 个 SNP 位点为基于已知品种指纹数据, 利用最优遗传算法确定的一套适于玉米品种真实性鉴定的核心位点组合^[26]。利用最优遗传算法、采用 LociScan 软件, 基于三等位变异 SNP 候选位点组合, 从 1 个位点开始寻优, 逐步增加位点并计算品种区分能力, 最终确定核心位点组合^[23,25]。利用 SNP 比对统计工具软件 V1.0 (北京市农林科学院, 软著登记号: 2018SR026743) 分析核心

位点组合的多态性和品种区分能力参数 PIC (polymorphism information content) 及 DP (discrimination power), 统计每个位点 3 种等位变异的分布频率; 基于核心位点组合和二等位 SNP 位点组合的基因型数据, 分别计算品种间两两成对比较差异位点的百分比; 绘制上述统计数值的分布图。利用核心 SNP 位点组合和 270 个自交系的基因型数据, 基于 Nei1973 方法获得品种之间的遗传距离, 选择邻接法构建聚类图。

2 结果与分析

2.1 适于玉米品种鉴定的三等位变异 SNP 核心位点组合

从全基因组中挖掘获得初始 ma SNP 位点集, 包含 78 万多个等位变异位点; 然后基于测序数据质量、位点拷贝数、杂合度、多态性、侧翼序列保守性、靶向测序探针设计要求, 逐级筛选评估确定优异位点集包含 7.0 万个位点、候选位点集包含 2.8 万个位点、适于靶向测序分型技术的集合包含 2.2 万个位点。利用广泛选取的玉米品种材料, 基于靶向测序数据进行位点的严格评估过滤, 确定了三等位变异候选位点组合包含 180 个 SNP。基于 180 个 SNP 位点, 200 个杂交种和 270 个自交系的数据, 利用最优遗传算法从 1 个位点开始寻优, 并评估不同位点数目组合的品种识别率(图 1)。结果显示, 12 个位点组合已可以区分评估材料中的所有品种, 综合考虑位点的累计品种识别率、均匀分布、信息冗余以及品种发展空间, 最终确定了包含 40 个位点的核心位点组合(图 1 和附表 1)^[24-25]。40 个位点均匀分布于玉米 10 对染色体上, 每对染色体上有 4 个位点, 基本上为长短臂各 2 个位点; 其中, 15 个位点位于基因内区域(附表 1)。

图 2 展示了位点 TaSNP06 在 14 个玉米材料中的 Sanger 测序验证结果。TaSNP06 位点在 14 个材料中共分型获得 6 种基因型, 纯合及杂合基因型各 3 种; 其侧翼 30 bp 处连锁一个二等位变异 SNP 位点, 基因分型仅显示了 AA 和 GG 两种基因型。从 TaSNP06 和其连锁的二等位变异 SNP 位点的基因型数据来看, 2 个位点的结果具有一致性; 二等位变异将 14 个材料分成了 2 类, 即 AA 基因型的 10 个材料和 GG 基因型的 4 个材料; 三等位变异 TaSNP06 位点又进一步将 AA 基因型的 10 个材料区分为 4 类, 将 GG 基因型的 4 个材料区分为 2 类。

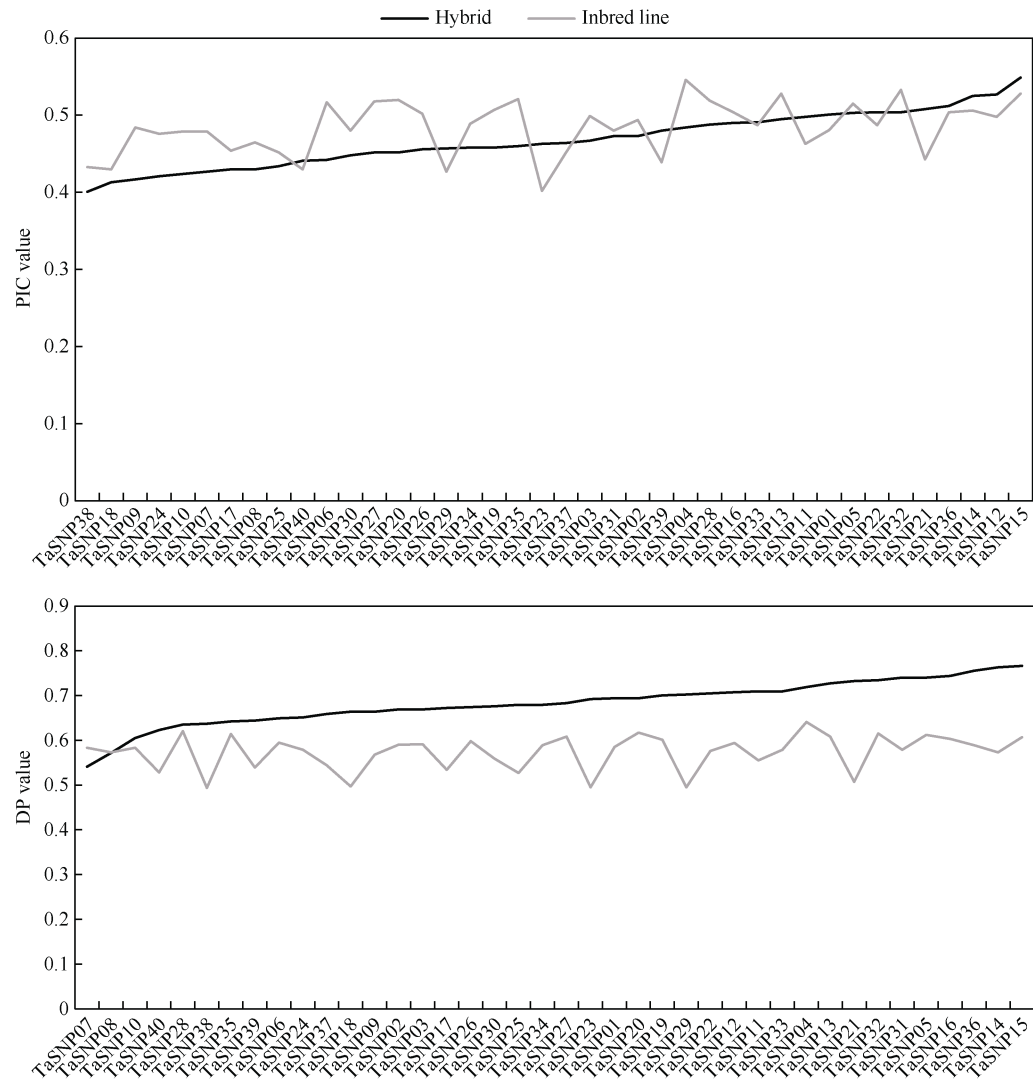
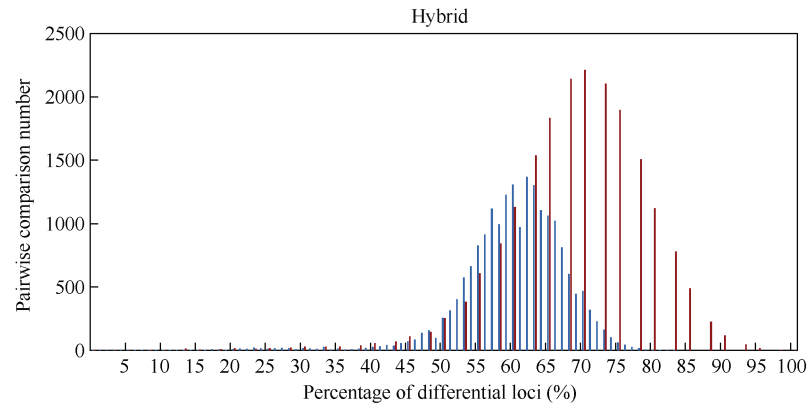


图 3 基于 200 个玉米杂交种和 270 个自交系评估 40 个 SNP 位点的 PIC 值和 DP 值

Fig. 3 PIC and DP values of 40 SNP loci evaluated using 200 maize hybrids and 270 inbred lines

横坐标为 SNP 位点, 顺序为基于杂交种数据评估值由小到大。

The x-axis represents the SNP loci, sorted in ascending order based on values evaluated from hybrid data.



(图 4)

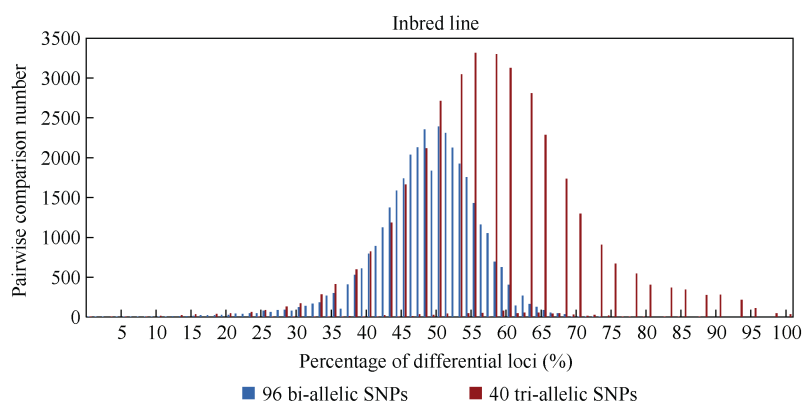


图 4 利用 96 个二等位和 40 个三等位变异 SNP 标记评估玉米杂交种、自交系两两成对比较的差异位点百分比分布图

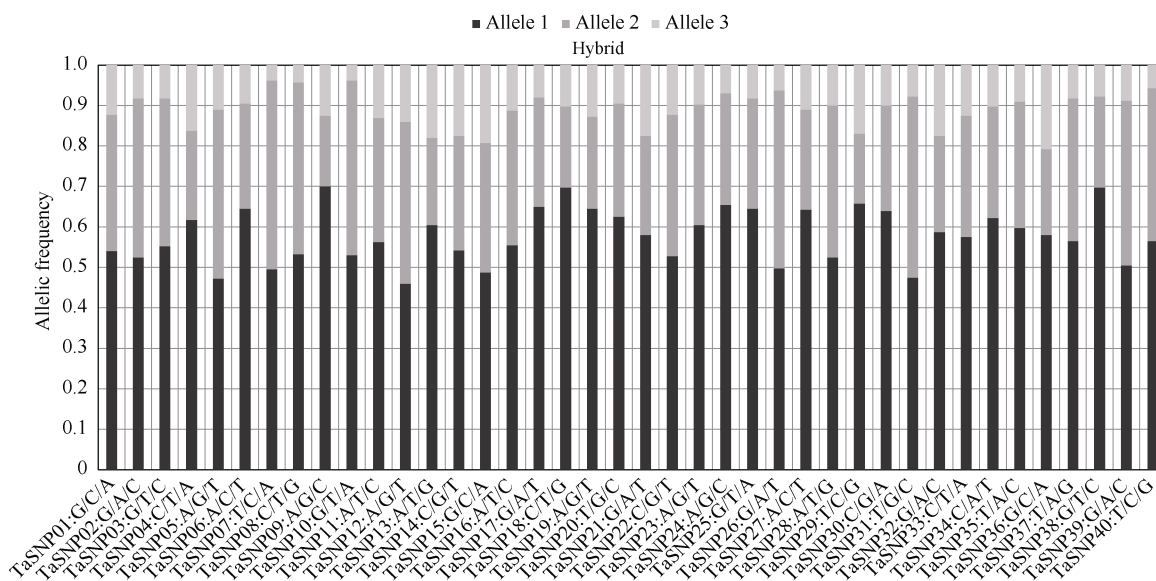
Fig. 4 Percentage distribution of differential loci in pairwise comparisons of maize hybrids and inbred lines using 96 bi-allelic and 40 tri-allelic SNP markers

杂交种的差异位点百分比范围分别为 1%~100%和 1%~84%，平均值分别为 70%和 60%，分别集中在 50%~85%和 50%~70%；差异位点百分比 $\geq 50\%$ 的分别占 96.81%和 94.32%， $\geq 60\%$ 的分别占 86.31%和 57.61%， $\geq 70\%$ 的分别占 52.88%和 7.32%， $\geq 80\%$ 的分别占 14.09%和 0.11%。基于 40 个三等位变异和 96 个二等位变异 SNP 位点统计显示，自交系的差异位点百分比范围分别为 1%~100%和 1%~73%，平均值分别为 58%和 48%，分别集中在 35%~75%和 35%~60%；差异位点百分比 $\geq 30\%$ 的分别占 98.69%和 97.78%， $\geq 40\%$ 的分别占 94.58%和 89.97%， $\geq 50\%$ 的分别占 78.33%和 46.29%， $\geq 60\%$ 的分别占 43.64%和 3.71%， $\geq 70\%$ 的分别占 15.45%和 0.03%。基于杂交种和自交系的 2 套数据分析结果，与 96 个

二等位变异 SNP 标记相比，40 个三等位变异品种间差异位点百分比的平均值提高了 10%，集中分布范围的最高值提高了 15%。

2.3 40 个 SNP 位点的等位变异频率、类群划分能力

基于 200 个杂交品种和 270 个自交系基因型数据，对 40 个位点 3 种等位变异频率的分布情况进行统计(图 5)。每个位点的 3 个等位变异按照频率由大到小排列，杂交种和自交系中第一等位变异频率分布范围分别为 0.46~0.70 和 0.48~0.68，平均值分别为 0.58 和 0.57；第二等位变异频率分布范围分别为 0.17~0.47 和 0.19~0.45，平均值均为 0.31；第三等位变异频率分布范围分别为 0.04~0.21 和 0.05~0.22，平均值分别为 0.11 和 0.12。



(图 5)

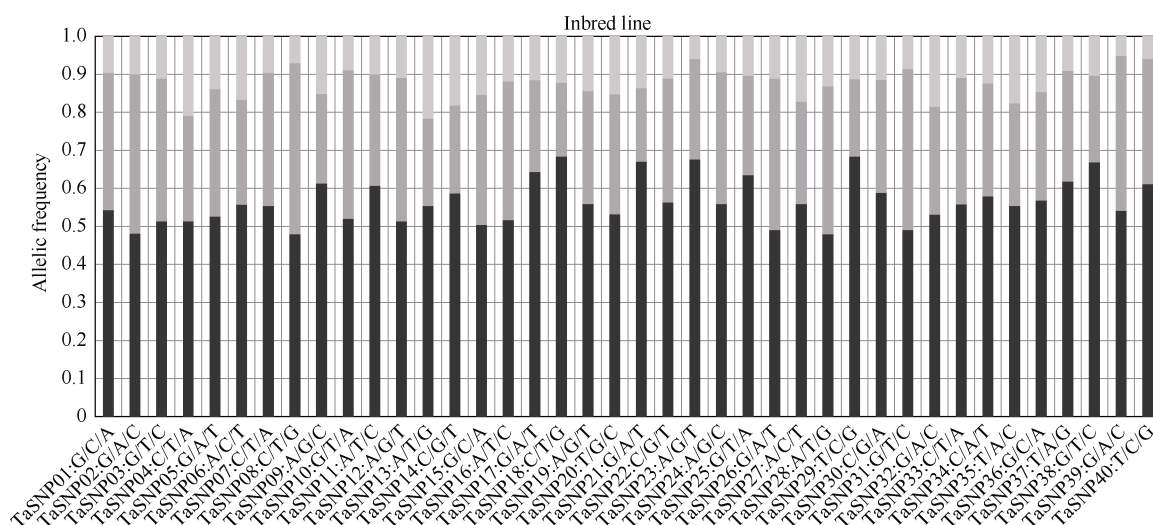


图 5 基于玉米 200 个杂交种和 270 个自交系评估 40 个 SNP 位点的 3 种等位变异频率

Fig. 5 Frequency distribution of three alleles at 40 SNP loci evaluated using 200 maize hybrids and 270 inbred lines

横坐标为 40 个 SNP 位点及其 3 种等位变异, 例如, TaSNP01:G/C/A 指等位变异 1 为 G, 等位变异 2 为 C, 等位变异 3 为 A。

The x-axis shows the 40 SNP loci and their corresponding tri-allelic variants (e.g., TaSNP01:G/C/A, where G, C, and A represent variants 1, 2, and 3, respectively).

基于邻接法的聚类结果, 将 270 个自交系精准划分为 10 个杂种优势类群, 分别是塘四平头群 (Tangsi ping tou)、旅大红骨群 (Lvda red cob)、瑞德群 (Reid)、改良瑞德群 (Improved Reid)、兰卡斯特群 (Lancaster)、混合 1 群 (Mixed 1)、混合 2 群 (Mixed 2)、

爱荷华黄马齿群 (Iodent)、P 群和 X 群; 其中, 塘四平头与混合 1 群、混合 2 群在一个主分支上, 改良瑞德群与 X 群在一个主分支上。类群划分结果与基于 61,214 个二等位变异 SNP 标记划分结果一致^[11], 与系谱记录相符 (图 6)。

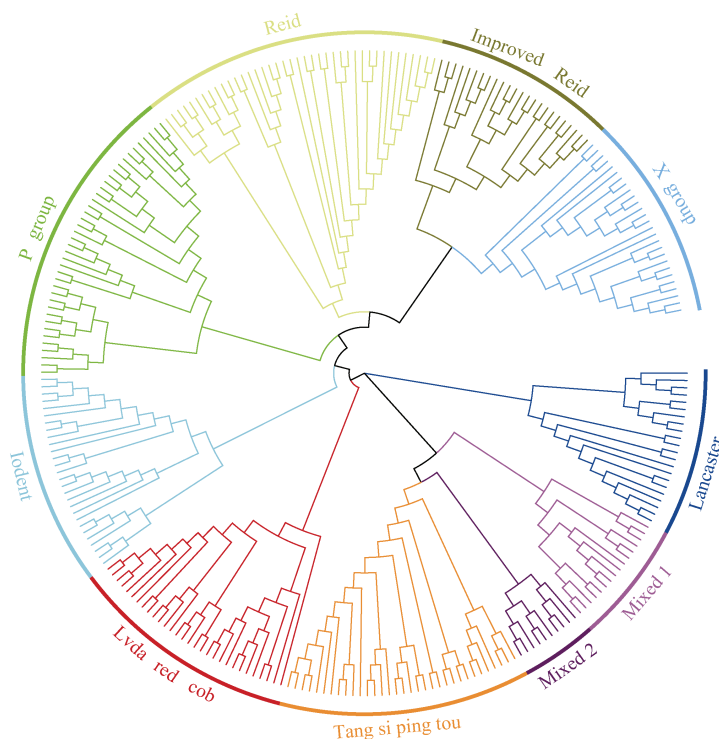


图 6 基于玉米 270 个自交系的 40 个 SNP 位点数据构建的杂优类群聚类图

Fig. 6 Clustering diagram of heterosis groups constructed using the data of 40 SNP loci of 270 maize inbred lines

Tangsi ping tou: 塘四平头群; Lvda red cob: 旅大红骨群; Iodent: 爱荷华黄马齿群; P group: P 群; Reid: 瑞德群; Improved Reid: 改良瑞德群; X group: X 群; Lancaster: 兰卡斯特群; Mixed 1: 混合 1 群; Mixed 2: 混合 2 群。

3 讨论

3.1 ma SNP 标记位点在玉米品种鉴定中的应用优势

随着基因组学相关技术的快速发展,各种类型分子标记在农作物品种鉴定中发挥着日益重要的作用,如简单重复序列多态性的 SSR 标记、小型片段插入缺失的 InDel 标记和单核苷酸多态性的 SNP 标记。SSR 标记主要优势为单个位点多态性高,但在全基因组分布密度低且主要分布在基因间隔区;利用毛细管电泳平台进行基因型分型,由于重复序列特征难以实现自动化精准基因型读取。InDel 标记兼容毛细管电泳、KASP (kompetitive allele specific PCR)和固液相芯片分型平台的优点,主要缺点为单个位点仅有 2 个等位变异,且为插入缺失变异,因而单个位点多态性低。SNP 标记具有分布密度高、覆盖全基因组的特征,且分布区域包含基因内和基因间隔区,同时由于单个核苷酸变异特征易实现高效精准基因型读取整合。目前利用的 SNP 标记均为二等位变异类型,单个位点的分辨率低是其主要缺点。随着玉米品种数量剧增,出现了同质化品种难分辨、套牌侵权难识别、种质来源难追溯等难题。因而急需各种新型标记和检测方法的开发探究,以达到标记位点高分辨率、检测规模高通量、指纹数据易共享能追溯的目的。

ma SNP 在基因组中分布广泛且均匀,尤其是在重组热点区域其富集现象更显著;多等位变异 SNP 标记可以是 3、4 和 5 个等位变异,以 3 个为主。多等位变异 SNP 标记位点在品种鉴定中的显著优势主要体现在以下方面。第一,与二等位变异 SNP 和 InDel 标记相比,单个位点的多态性和品种区分能力明显提高。即使三等位变异的 SNP 标记,其单个位点基因型可出现 6 种,杂合基因型为 3 种,完全解决了二等位变异标记仅有 1 种杂合基因型的困境。研究结果显示,40 个 SNP 标记位点的 PIC 值均高于二等位变异 SNP 标记,品种间差异位点百分比的平均值提高了 10%,这与单个位点等位变异数目的增加直接相关。因而 ma SNP 标记有利于提高品种真实性、纯度和亲缘关系的鉴定效率,有利于解决采用混株制备试验样品带来的基因型数据精准读取的问题。第二,ma SNP 标记可以解析更多遗传变异,尤其是稀有变异。40 个三等位变异 SNP 位点在杂交种和自交系中显示可以解释稀有变异的频率介于 0.04

到 0.22 之间;四等位变异和五等位变异的 SNP 位点则能解释更多的变异类型,进而精准解析复杂性状。第三,单核苷酸变异特征保障基因型数据的高效精准采集。由于所开发的多等位变异位点仍然为同一个 SNP 位点内的单个核苷酸变异,并非紧密连锁的几个 SNP 组成的标记类型,因而其基因型数据采集与二等位变异标记相同,具有简便、精准的特征。

3.2 ma SNP 标记应用在品种鉴定中需要研究和解决的关键问题

农作物品种分子鉴定技术的研究应用主要有 3 个阶段,上世纪 90 年代基于第一代标记,如 RFLP (限制性片段长度多态)和 RAPD (随机扩增多态)的鉴定技术,21 世纪初兴起的基于第二代标记 SSR 的鉴定技术,以及近十年研究应用焦点基于第三代标记 SNP 的鉴定技术^[27]。各标记鉴定技术在不同阶段、不同检测需求中发挥了优势作用,随着高通量测序和基因组学技术的快速发展,为利用更新更高效的分子鉴定技术有效鉴别农作物品种提供了可能。因而,复杂多样的基因组中的新型标记类型需要被探索应用,如 ma SNP、序列变异标签标记等。本研究基于全基因组测序数据初步探索评估多等位变异 SNP 标记在品种鉴定中的应用潜力,其成熟应用仍面临一系列亟待研究和解决的问题。

如何开发高质量 ma SNP 位点是首要的技术挑战。为了筛选出真实、稳定的多等位变异 SNP 位点,需避免由于测序、比对错误或复杂的基因组区域产生的假阳性位点。这需要利用高质量参考基因组和深度重测序数据,需要开发精准的生物信息学流程,需要设定严格的过滤参数,如测序深度、位点质量值、等位变异频率等,需要通过 Sanger 或靶向测序对位点进行验证。评估一套高质量 ma SNP 位点集合,还需要进一步利用遗传算法确定出最优位点集合。其次,建立 ma SNP 标记高效基因型分型平台是其实现广泛应用的关键。鉴于其等位变异数量特征,基于毛细管电泳的突变阻滞扩增系统 PCR 技术 (amplification refractory mutation system PCR, ARMS-PCR)和基于测序的靶向测序技术 (genotyping by target sequencing, GBTS)可作为适宜的分型平台^[28-32]。ARMS-PCR 技术是一种根据 SNP 位点设计 3'末端与 SNP 位点碱基互补或错配的、长度不同的特异 PCR 引物,通过电泳方法检测扩增产物的长度,从而获得 SNP 基因型的方法。该方法适用于中低通量 ma SNP 标记的基因型分型,尤其是针对自

交系类型的鉴定, 具有简便、快捷、低成本的特征。靶向测序技术是近些年来成熟应用的分型技术之一, 对于区分 ma SNP 及其侧翼连锁标记优势明显, 并且具有高通量、可规模化检测的特征。最后, ma SNP 标记的指纹采集、统计比对和管理面临新的复杂性, 需基于标记和分型平台特征, 结合已成熟应用的多等位变异 SSR、ba SNP 标记的经验, 开发简便、高效的配套软件工具以及算法, 建立全流程式指纹采集分析管理方案。

概而言之, 随着测序技术的迭代、生物信息学方法的创新和深度学习算法的应用, ma SNP 标记位点的挖掘、分型效率将大幅提升, 凭借其更高的多态性和更精准的解析能力, 有望与 SSR、ba SNP、InDel 形成互补, 推动农作物分子鉴定进入一个全新的精准时代。

4 结论

本文基于全基因组挖掘的 78 万个初始 ma SNP 位点, 利用生物信息、统计分析、靶向测序验证以及遗传算法等方法, 评估确定了一套包含 40 个位点的高质量、高鉴别力的三等位变异 SNP 核心位点组合, 为玉米品种鉴定提供了一种新型标记类型, 凭借其更高的变异解析能力, 有望与现有标记类型形成互补, 推动农作物分子鉴定技术发展。

References

- [1] 中华人民共和国农业农村部, 全国植物新品种测试标准技术委员会. 玉米品种鉴定 DNA 指纹方法: NY/T 1432-2007. 北京: 中国农业出版社, 2007.
Ministry of Agricultural and Rural Affairs of the People's Republic of China, National Technical Committee for Testing of New Plant Varieties. Maize Variety Identification Molecular Techniques: NY/T 1432-2007. Beijing: China Agriculture Press, 2007 (in Chinese).
- [2] 中华人民共和国农业农村部种子管理局, 全国植物新品种测试标准技术委员会. 玉米品种鉴定技术规程 SSR 标记法: NY/T 1432-2014. 北京: 中国农业出版社, 2014.
Seed Management Bureau of the Ministry of Agricultural and Rural Affairs of the People's Republic of China, National Technical Committee for Testing of New Plant Varieties. Protocol for the Identification of Maize Varieties-SSR Marker Method: NY/T 1432-2014. Beijing: China Agriculture Press, 2014 (in Chinese).
- [3] 中华人民共和国农业农村部, 全国农作物种子标准化技术委员会. 主要农作物品种真实性和纯度 SSR 分子标记检测 玉米: GB/T 39914-2021. 北京: 中国标准出版社, 2021.
Ministry of Agricultural and Rural Affairs of the People's Republic of China, National Technical Committee for Standardization of Crop Seeds. Variety Genuineness and Purity Testing of Main Crops with SSR Markers-Maize: GB/T 39914-2021, Beijing: Standards Press of China, 2021 (in Chinese).
- [4] 王风格, 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.
Wang F G, Yang Y, Yi H M, et al. Construction of an SSR-based standard fingerprint database for corn variety authorized in China. Sci Agric Sin, 2017, 50: 1-14 (in Chinese with English abstract).
- [5] 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.
Wang F G, Zhao J R, Yang Y. SSR Fingerprinting Data of National Approval of Maize Cultivars. Beijing: China Agricultural Science and Technology Press, 2015 (in Chinese).
- [6] Lai J S, Li R Q, Xu X, et al. Genome-wide patterns of genetic variation among elite maize inbred lines. Nat Genet, 2010, 42: 1027-1030.
- [7] Chia J M, Song C, Bradbury P J, et al. Maize HapMap2 identifies extant variation from a genome in flux. Nat Genet, 2012, 44: 803-807.
- [8] Jiao Y P, Zhao H N, Ren L H, et al. Genome-wide genetic changes during modern breeding of maize. Nat Genet, 2012, 44: 812-815.
- [9] Semagn K, Babu R, Hearne S, et al. Single nucleotide polymorphism genotyping using kompetitive allele specific PCR (KASP): overview of the technology and its application in crop improvement. Mol Breed, 2014, 33: 1-14.
- [10] Bukowski R, Guo X S, Lu Y L, et al. Construction of the third-generation *Zea mays* haplotype map. Gigascience, 2018, 7: gix134.
- [11] Tian H L, Yang Y, Yi H M, et al. New resources for genetic studies in maize (*Zea mays* L.): a genome-wide Maize6H-60K single nucleotide polymorphism array and its application. Plant J, 2021, 105: 1113-1122.
- [12] Tian H L, Wang F G, Zhao J R, et al. Development of maizeSNP3072, a high-throughput compatible SNP array, for DNA fingerprinting identification of Chinese maize varieties. Mol Breed, 2015, 35: 136.
- [13] 田红丽, 杨扬, 王璐, 等. 兼容型 maizeSNP384 标记筛选与玉米杂交种 DNA 指纹图谱构建. 作物学报, 2020, 46: 1006-1015.
Tian H L, Yang Y, Wang L, et al. Screening of compatible maizeSNP384 markers and the construction of DNA fingerprints of maize varieties. Acta Agron Sin, 2020, 46: 1006-1015 (in Chinese with English abstract).
- [14] Tian H L, Yang Y, Wang R, et al. Screening of 200 core SNPs and the construction of a systematic SNP-DNA standard fingerprint database with more than 20,000 maize varieties. Agriculture, 2021, 11: 597.
- [15] 中华人民共和国农业农村部种业管理司, 全国农作物种子标准化技术委员会. 玉米品种真实性鉴定 SNP 标记法: NY/T 4022-2021. 北京: 中国农业出版社, 2021.
Seed Management Department of the Ministry of Agricultural and Rural Affairs of the People's Republic of China, National Technical Committee for Standardization of Crop Seeds. Maize (*Zea mays* L.) Variety Genuineness Identification-SNP Based Method: NY/T 4022-2021. Beijing: China Agricultural Publishing House, 2021 (in Chinese).
- [16] 中华人民共和国农业农村部种业管理司, 全国农作物种子标

准化技术委员会. 水稻品种真实性鉴定 SNP 标记法: NY/T 2745-2021. 北京: 中国农业出版社, 2021.

Seed Management Department of the Ministry of Agricultural and Rural Affairs of the People's Republic of China, National Technical Committee for Standardization of Crop Seeds. Rice (*Oryza sativa* L.) Variety Genuineness Identification–SNP Based Method: NY/T 2745-2021. Beijing: China Agriculture Press, 2021 (in Chinese).

[17] 中华人民共和国农业农村部种业管理司, 全国农作物种子标准化技术委员会. 小麦品种真实性鉴定 SNP 标记法: NY/T 4021-2021. 北京: 中国农业出版社, 2021.

Seed Management Department of the Ministry of Agricultural and Rural Affairs of the People's Republic of China, National Technical Committee for Standardization of Crop Seeds. Wheat (*Triticum aestivum* L.) Variety Genuineness Identification–SNP Based Method: NY/T 4021-2021. Beijing: China Agriculture Press, 2021 (in Chinese).

[18] Westen A A, Matai A S, Laros J F J, et al. Tri-allelic SNP markers enable analysis of mixed and degraded DNA samples. *Forensic Sci Int Genet*, 2009, 3: 233–241.

[19] Phillips C, Amigo J, Carracedo Á, et al. Tetra-allelic SNPs: informative forensic markers compiled from public whole-genome sequence data. *Forensic Sci Int Genet*, 2015, 19: 100–106.

[20] Gao Z H, Chen X G, Zhao Y C, et al. Forensic genetic informativeness of an SNP panel consisting of 19 multi-allelic SNPs. *Forensic Sci Int Genet*, 2018, 34: 49–56.

[21] Phillips C, Amigo J, Tillmar A O, et al. A compilation of tri-allelic SNPs from 1000 Genomes and use of the most polymorphic loci for a large-scale human identification panel. *Forensic Sci Int Genet*, 2020, 46: 102232.

[22] Yang S Q, Shi H T, Zhang Z F, et al. Detection of tri-allelic single nucleotide polymorphisms of *ABCB1* and *CRP* genes by penta-primer amplification refractory mutation system-polymerase chain reaction. *Genet Test Mol Biomarkers*, 2022, 26: 43–48.

[23] 王凤格, 赵久然, 杨扬. 基于遗传算法的植物品种真实性鉴定位点筛选方法. 中国专利: ZL 201310629676.2, 2013-11-29.

Wang F G, Zhao J R, Yang Y. The Method of Locus Selecting for Plant Variety Authenticity Identification Based on Genetic Algorithm. Chinese Patent: ZL 201310629676.2, 2013-11-29 (in Chinese).

[24] Yang Y, Tian H L, Wang R, et al. Variety discrimination power: an appraisal index for loci combination screening applied to plant variety discrimination. *Front Plant Sci*, 2021, 12: 566796.

[25] Yang Y, Tian H L, Yi H M, et al. LociScan, a tool for screening genetic marker combinations for plant variety discrimination. *Crop J*, 2024, 12: 583–593.

[26] 田红丽, 杨扬, 范亚明, 等. 用于玉米品种真实性鉴定的最优核心 SNP 位点集的研发. 作物学报, 2024, 50: 1115–1123.

Tian H L, Yang Y, Fan Y M, et al. Development of an optimal core SNP loci set for maize variety genuineness identification. *Acta Agron Sin*, 2024, 50: 1115–1123 (in Chinese with English abstract).

[27] 赵久然, 王凤格, 田红丽, 等. 适于农作物品种分子身份鉴别和确权鉴定的检测方法. 中国专利: ZL 201710527354.5, 2017-06-30.

Zhao J R, Wang F G, Tian H L, et al. Detection Methods for Identity Distinguish and Intellectual Property Confirmation of Crop Varieties Using Molecular Markers. Chinese Patent: ZL 201710527354.5, 2017-06-30 (in Chinese).

[28] Newton C R, Graham A, Heptinstall L E, et al. Analysis of any point mutation in DNA: the amplification refractory mutation system (ARMS). *Nucleic Acids Res*, 1989, 17: 2503–2516.

[29] Ye S, Dhillon S, Ke X, et al. An efficient procedure for genotyping single nucleotide polymorphisms. *Nucleic Acids Res*, 2001, 29: E88-8.

[30] Latorra D, Campbell K, Wolter A, et al. Enhanced allele-specific PCR discrimination in SNP genotyping using 3' locked nucleic acid (LNA) primers. *Hum Mutat*, 2003, 22: 79–85.

[31] 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. 中国农业科学, 2020, 53: 2983–3004.

Xu Y B, Yang Q N, Zheng H J, et al. Genotyping by target sequencing (GBTS) and its applications. *Sci Agric Sin*, 2020, 53: 2983–3004 (in Chinese with English abstract).

[32] Guo Z F, Yang Q N, Huang F F, et al. Development of high-resolution multiple-SNP arrays for genetic analyses and molecular breeding through genotyping by target sequencing and liquid chip. *Plant Commun*, 2021, 2: 100230.

附表 1 适于玉米品种鉴定的一套三等位变异 SNP 核心位点信息
Table S1 Detailed information of the tri-allelic SNP core loci suitable for maize variety identification

位点 Locus	染色 体 Chr.	物理位置 Physical location	等位 变异 1 Allele 1	等位 变异 2 Allele 2	等位 变异 3 Allele 3	上游侧翼序列(60 bp) Upstream flanking sequence (60 bp) (5'–3')	下游侧翼序列(60 bp) Downstream flanking sequence (60 bp) (5'–3')
TaSNP01	1	26,694,070	A	C	G	CCAGCCAGTCTGCTTCAACCTTT GCCGGTGCCCTCGGCCATTGG AATGACATCTCCATA	ATATCATTGGGGGTTTCCTCGAGT TGCTGGCAAATCTGTGATTTGAC TGTGGTGAACAT
TaSNP02	1	70,264,325	C	A	G	CCTCGATCCTTGAGAGATTTTC ATATCCATTGACCAATAAGCAT TCATTCATCACCTG	GTGTCAATAAAACCTAGAGCCCCT AGTCTTTTGGCCTAGTTAGCATCT CCCACCTCGCCA
TaSNP03	1	70,370,458	C	T	G	CGGAACCTCTGGAACGGAGCC TCGAACAAAAAAGAATATA GAACAATATAAGCAC	GTTAGCAGATGCGATGACGACTA GAAGGGAATGGGAACCGGGACC ACGGCAAGCAGACAC

(续附表 1)

位点 Locus	染色 体 Chr.	物理位置 Physical location	等位 变异 1 Allele 1	等位 变异 2 Allele 2	等位 变异 3 Allele 3	上游侧翼序列(60 bp) Upstream flanking sequence (60 bp) (5'-3')	下游侧翼序列(60 bp) Downstream flanking sequence (60 bp) (5'-3')
TaSNP04	1	268,027,869	T	C	A	TTCTTGTGAATTGTGATAGAGCC TTGTGAGAGTGCTAAAGAGAAG AGTGTGTGCTGCGTT	TTGTGATCTTGATGCAAGACACCC ATCCATTGTAAAGCTAGCATTAGC CTTCAACGTGTG
TaSNP05	2	56,077,965	T	A	G	TGCCAACCCAGTAAACACCAAGT TCTACAGGCAAGGCAGCATAACA AGACTTTCCTTTTGAG	ATAGAATTTGCGAAAGATAGCGCA TATGGCGTGCAAAATTGGCTGTGA TAAAGCAATTAC
TaSNP06	2	191,937,511	T	A	C	AATATGTTCCATAGAATTACCTTA GGTTTGTCTCTCTTCACGCAGC AGTTCTAATGATC	TGTACTTCTTGACTGTTGCCATTC TTAATAAGAAACCGTCCAGTCATT TGATAATGCTTA
TaSNP07	2	192,091,927	A	T	C	ACTATCTACACAGAGACGGATAG CATTAGTTACAGCAGAACATGTA CTGCAATATACTGT	TGGAGAGAAGGGGGCATCAGCGC GCGGAGTGGCAACAGATAAACGG CATCCACGGCTGCT
TaSNP08	2	237,697,548	G	T	C	GTTGATATTTCTATTACGGATGTGG TCACCACAACCGAATCTTTAGGC ATGAACCTCTTCCA	GCTTCTTACTATGCCAGATCACAC ACTAGATGATAATCGGTTATAACC ATTCATCTATAA
TaSNP09	3	47,899,705	G	A	C	GAAGAGCAAGTAGCCATATTTTT GCATGTTCATAGGAAACAATGAG CGATTTCGTGTCATT	CAATGTCCTTTCGCCGCTCTACCG AGACAATTAGTCATTATTTCCAAG AGGTCCTAGCAG
TaSNP10	3	129,111,971	A	G	T	AGTGCTGTTTCCTCACCTCAGCC AGTCCACAAGTAGCAAGTAGCA ACTATAGTCCTAAAG	CTAAAGTACTAATACTACTAGCCAG ACATAACTAATTCCATATGCTTTG GGAGACAATGCA
TaSNP11	3	222,001,675	C	A	T	CCAACATTTGAGTGGGGGCCTA AATAGAGGCTAAATGACTACTAG ACCTTTGTAGCTACA	TAAAAATGTCCACTGAAACTAGAA AATTGAGTGGGGGCTGTGCCCC CGCTAAGCCCAACG
TaSNP12	3	224,178,922	T	G	A	TTGCACCTCCTCAATTTTGTTATA GCTTCGTCATTTACTAGCACCAG CGTGTCGTGCCCCG	TGACAAAAGAAATAGTGAAGAA GACCAGGAGCAGCCTTGCGGTGG TGGCCAGGTGCAAC
TaSNP13	4	3,788,094	G	A	T	AAATAGAAAATCTTGACCTTCAG CATCATGCTAAACAGGTGAACCA TGAACACTGTTAGA	ACTTTATGCACATTTTTATTGTACA TATATTTTTTTCCTAATTAACCTCT GCCAGAGCCC
TaSNP14	4	10,076,723	G	C	T	GCATCCTCTCCTAGATGCTCTCT TATGTTACTAGTATCAAGTTTCTA GCCAAAATGATGC	GGTCTTATTTCTACAATACATAGT GATGGCCACACTATAAATGCAGGT GATGTGCTCTC
TaSNP15	4	10,645,740	A	C	G	TATAGAGGAGCATCAAACGTGAC GATGGTCACTCGTGGACTATTG ACAGGCATCAAACC	TGTGAATGATGCACGAGTCCTATC CAACAAGCTCAAATCATAGGCATT AGATAGTCCGAT
TaSNP16	4	86,481,785	C	A	T	CAAATATGAGAAATCACTATATG TCATTCTTATTTTGAGGCCTAG GATTTTAGCTCATT	TGTGTAGCTACAACAAGTGTGAC ACAATAGCTGTAGAAGGGGTGAG TATGGATCATGATC
TaSNP17	5	1,487,478	T	G	A	ACCCGCCCCCTTCCGTGCGAC GAATAGTTGGCGAACACTGCAT CGAGTTACGAGACATT	AGTCAAGACACGAGAAGGGGAA AGCATCGGTGACGCTAAGAACA TCAAACAGTTGGGAT

(续附表 1)

位点 Locus	染色 体 Chr.	物理位置 Physical location	等位 变异 1 Allele 1	等位 变异 2 Allele 2	等位 变异 3 Allele 3	上游侧翼序列(60 bp) Upstream flanking sequence (60 bp) (5'-3')	下游侧翼序列(60 bp) Downstream flanking sequence (60 bp) (5'-3')
TaSNP18	5	60,276,448	G	C	T	CTCTGTCTCTTTTAGTCTTTTGG AGGAGGAACAAAGAACCAGTG GGGTTCGAAAGGAG	AGTAGGCCGACAAGTAGCATCGC ATCACGTTCTCCTAATGAATGACT AGCATGAGTCAAA
TaSNP19	5	91,897,631	G	A	T	CCCACCAATGATATCCCTTATGC TCAGCACTATACCCAAGAAAAA GCACACTCAACTGAC	GAGCAATCAACTTAGTGCCTCAT GAGGTGGGAGCAGCATATAAA CACAACCTAGACAT
TaSNP20	5	186,888,339	C	T	G	TACTCAGCAGGTAATGATGCTAT AAAAGCATCAGACTGCCTGGGC GCCATATCAGCCTTC	ACTGGTGCATATCAGCCACTGCTT TCACCTCCCCCACCCTCTCACTCT CACCTGATATCT
TaSNP21	6	75,706,844	T	G	A	GGAATACTTCGGTTGAAGTCGTG GAACCTCGTTCCTACGCGATGAA TGGGTGGGGGAACA	GTGGGACTTTTGGTTTACATTCA AACAAAAGATAGTGAAGGGGTTT TGAGTCTGTCTGTT
TaSNP22	6	149,861,566	T	G	C	TAGGCAAGTTAGTAGAGGTGTAT TTCAATGACATTGTATTGAAAAC ATGTGAGTCCAACA	TCCCATTGGGATCTTGAAACCGT TTTACAGCCTCCCCCAAACTG CATCTTCAACAT
TaSNP23	6	151,920,452	G	T	A	AATTAAGTGTTTTGTTCACAAA ACACATACATTCAAAGTGTAAGA CCAAATACCACCAC	GAATGGTTGAAGCTACCCAGGGA CATCCATAATTAAAGAACCAAAAA CAAATAAAAGTAG
TaSNP24	6	167,615,957	C	A	G	GAATGAAGGAAACAATTAATTG GACCTTCTGAGCCTGGCCTAATT GCAGAGGTATGCAAT	TGGACAGCAGTTGAGGCCATCAA ACTCAGATTGAGAGGCCGGTTGG AAACCGGTATGACG
TaSNP25	7	20,793,475	A	T	G	GCTCTAAGTTACCCTCATTCTATA GTGAAAAACATTCTCATATTAAG AGCAACAGCTCCTG	CTTGACAACGAGGAAGGGTGTGT GACCCGTGGCCTTTGTGGCTGAC CTCAACTACGTAGA
TaSNP26	7	23,775,501	T	G	A	CGACATACTCGTCGCTGGTCTAC ACCACATTTGTCTCGAGTTCTT CGCCGACAAGTACC	TCGTCGCGCTCCTCTGGCGCCATC GAGTCTCTTCATGTGTGTCGTAAC TCTCACCGTTGC
TaSNP27	7	25,104,386	T	A	C	ATTTAGTTTTAAATTCCATATCAA TCCACGCTATAATAACACATGTAT AAGGAGTATATC	AAACAGGCCACACATGGAATGT TTGAGATATGCATTCAAATGCATT CCTCCGCATCCTG
TaSNP28	7	91,528,461	G	T	A	GATCCATTTTTTGAGGAACGGTG GGTTGCTTCATCGTTGGTCCCTA TAGTGTGGAGATTC	TGCTTTCCAAGGCAGTCAATTCAT CCAACGGGAGTGACGACACATTT CATCGAGGTAGTC
TaSNP29	8	117,427,483	G	T	C	GCTTGTTGAAAAGGCGAATCC AAAGAAATGCTAACGTTGCAGA AGAATGAAGGCCAATT	TGCCTTAGATTGCGTAAGCATGAG AAAAGCAAGTGGCTACAGAAACG ACCGCTACAGATA
TaSNP30	8	141,530,330	A	C	G	CTTCACCGCTCTCAAACCTGACAT CTATCTCCGAACAGAACATACAA ATAATCAAACCAAT	AGTTTACTGAATCAAGTGCAATTC GCCGTAAGAACAAACCTGTAAGG TACGATATTTGTG
TaSNP31	8	141,766,718	C	G	T	TATAGCGAGGACAAGACATGAG CAACTTCTGGTCAGGTAGAAAA GTACGGTTGGACAGCA	CATGAAGCACACAAAAAGGGATT AGATAAAATCATGGCTGCAGTA GGTTGATTAGGTCT

(续附表 1)

位点 Locus	染色 体 Chr.	物理位置 Physical location	等位 变异 1 Allele 1	等位 变异 2 Allele 2	等位 变异 3 Allele 3	上游侧翼序列(60 bp) Upstream flanking sequence (60 bp) (5'-3')	下游侧翼序列(60 bp) Downstream flanking sequence (60 bp) (5'-3')
TaSNP32	8	172,879,909	A	G	C	TCCAATGCCAGTCTCTCGAGGG AACTAGAGTTTAGGCTAGCCTCT AGTTTCTATCTTGTT	CACTATTCTATGGACCAATGGATT TATTCAGTGGTCCCTGGTTCCAAG CACGAGAGCAAA
TaSNP33	9	46,962,286	T	C	A	CTTATATTTATCACATTGCTGTTT TTTTTACTGTTGTGCGGACCAAT GTTTTTTTAATTT	TTGAGAACAACAGTGCAACATGC GTAAGTTAGAGAAGGAGCAGCAA GGTAATCCATGTCC
TaSNP34	9	88,827,030	T	C	A	TGCAAAATTCTCACTCGCTAACC CTAATCCTCATTTCTCTAAAATAAT CGTTGAGAAAGCA	CGGACATGTCTGGTGCAGAAAGT GGACGAGCACCGAACATGTCCAA TGCACCCAACCTGGG
TaSNP35	9	97,080,329	A	T	C	ATGGGAAAATTCACAGCGAATC TGCTTTGCCCATAGGTAATGAAC GAATGCCATCGATGG	ACACCACTTAGATATGGAAACTGC CAGCCGACCAGCTGGGCTGGTTC TTTCTAACTTCGC
TaSNP36	9	149,325,574	A	G	C	ACTGGACTTCGAATCCCTTCTTC TGCTCTGCAATAATTGCAAGCAC AAAAGAAGGCTAAA	CAAGGAATTGCTCCTTCCTTCTAG AGGTCTTTGCGCTGGGACGTTGA CGAAATTATCTCG
TaSNP37	10	16,222,443	A	T	G	ATGATAGGGCTCCAACATATGGG GTCCTATATTTAGGACCTGAGAG GCTGAGTATTATAA	TATTCGATGATTTTTTGTGAAATAA GACCTCTGTTGGAGCAAGCTTTTA GTGTTAGAGTC
TaSNP38	10	113,185,518	T	G	C	AAAAAAGGATCAGGCACTCTGA TAAACTAATAGCAAGGTGATTGC AAATCAGTATGTCCA	CCTACTAAATCTGTAGATCTGTAC ACAAAAATTGCAAAGGATGGCAT CTTAACAACGCAG
TaSNP39	10	145,237,355	C	G	A	TAGTTGATCTGATCCCTAACCTT CAATCTTTGCACAAATGGGTAAT ACAACTTCTGACT	ATTGTAGCTTTTAAAGGCCAACTC CAAGAATGACATGCAAAAGACCC GAAGAAGTGTTTT
TaSNP40	10	146,434,632	G	C	T	TGGTTATGGGCCTAACGGGCAG AAGTATGTATATATAAAAATATTA GAATATGTTTCATCA	ACAGTACATATTAGACTCTCGCTG ACTGTTGTTGTACCATATATTTGTA ATATAGATCAT