

周 玲, 梁帅强, 林 峰, 等. 玉米二态性 InDel 位点的鉴定和分子标记开发[J]. 江苏农业学报, 2016, 32(6): 1223-1231.
doi:10.3969/j.issn.1000-4440.2016.06.005

玉米二态性 InDel 位点的鉴定和分子标记开发

周 玲, 梁帅强, 林 峰, 吕远大

(江苏省农业科学院农业生物技术研究所/江苏省农业生物学重点实验室, 江苏 南京 210014)

摘要: 随着下一代测序技术(Next generation sequencing, NGS)的发展, 可以利用高通量测序数据挖掘基因组特征的插入/缺失(InDel)位点, 这使 InDel 分子标记的开发成为可能。本研究利用 327 份玉米自交系的测序数据, 通过全基因组结构变异评估发现分布于 10 条染色体上的 25 847 个 InDel 位点。利用 *PIC* 值 ≥ 0.30 的 3 304 个位点进行系统发育树构建和群体结构分析, 其结果与利用 SNP 位点分析获得的结果高度一致, 说明这些多态性较高的 InDel 位点可以作为代表性标记用于玉米遗传组成的差异分析。采用 53 对 PCR 引物进行扩增试验, 发现这些 InDel 引物具有位点专化特征并且其杂交鉴定的平均杂合率为 0.569 2, 表明这些标记能准确鉴定出玉米杂交种的杂合度。

关键词: 玉米; 二态性 InDel 位点; 系统发育树; 群体结构

中图分类号: S513.3 **文献标识码:** A **文章编号:** 1000-4440(2016)06-1223-09

Biallelic InDel loci detection and molecular marker development in maize

ZHOU Ling, LIANG Shuai-qiang, LIN Feng, LÜ Yuan-da

(*Institute of Agro-biotechnology/Provincial Key Lab of Agro-biology, Jiangsu Academy of Agricultural Sciences, Nanjing 210014, China*)

Abstract: Next generation sequencing (NGS) technology revolutionized the throughput of sequence data generation and enabled genome-wide identification of insertion and deletion variation (InDel), which could provide potential information for molecular marker development. However, large scale characterization of InDel in maize genome was poorly understood. Here, we evaluated 327 maize inbred lines genomes and detected more than 25 847 biallelic InDel loci across ten chromosomes. 3 304 InDel loci with relative high *PIC* values were used to construct the genetic dendrogram, which was consistent with the results concluding from all called SNP loci, suggesting that these biallelic InDel loci could indiscriminately describe maize genome variations. To explore the application of the InDel loci, we developed 53 InDel markers for PCR validation. The results exhibited that the newly developed molecular markers were not only locus specific, but had the capabilities for identifying the hybrid seed heterozygosity. Taken together, our results will particularly facilitate the development of biallelic molecular marker for studies in genetics and have potential for marker-assisted breeding.

Key words: maize; biallelic InDel loci; dendrogram; population structure

玉米起源于美洲^[1], 是典型的异交作物, 表现为

极端的近交衰退和杂种优势, 因此杂种优势一直是玉米新品种选育的重要依据。尽管传统育种技术在作物遗传改良方面已经取得了显著成就, 但其选择效率低且周期长, 不能满足玉米生产中对优良品种的需求。自 20 世纪 80 年代以来, 分子标记被广泛应用于作物遗传育种中^[2-10], 高密度分子标记是基因克隆、分子育种以及全基因组关联分析中必不可少的工具。

随着测序技术的发展, 基因组数据资源呈爆炸

收稿日期: 2016-03-12

基金项目: 江苏省农业科技自主创新基金项目[CX(14)2006]; 江苏省自然科学基金项目(BK20141385)

作者简介: 周 玲(1987-), 女, 安徽合肥人, 博士, 助理研究员, 研究方向为作物遗传育种。(E-mail) zlingxiaoyao@163.com

通讯作者: 吕远大, (E-mail) lyd0527@jaas.ac.cn

式增长,利用高通量基因组数据开发插入/缺失(InDel)标记成为了可能。与单核苷酸多态性(Single nucleotide polymorphism, SNP)相比, InDel 是第二常见的多态性标记。Mills 等根据重测序结果确定了人类中的415 436个 InDel 标记,人类基因组多态性的总数超过 1.0×10^8 个,其中大约有 1.5×10^6 个是 InDel^[11]。有研究表明,较多的 InDel 会导致人类疾病^[12]。例如,在人类中最常见的遗传疾病——囊胞性纤维症,其主要是由于 *CFTR* 基因的编码区域发生了插入/缺失引起的^[13]。同样, InDel 也可以改变植物的表型,如玉米驯化基因 *Teosinte branched*、水稻增产基因 *Gn1a* 以及小麦抗条锈病基因 *Yr36*^[14-16]。因此, InDel 被认定为一个重要的分子标记来源,并成功应用于一些作物的遗传研究中^[17]。Liu 等以白菜和甘蓝的基因组为参考序列,设计 595 对 InDel 引物对 23 份甘蓝型油菜自交系进行检测,发现 461 个共显性 InDel 标记检测到 977 个等位基因,平均 2.12 个等位基因,基因多样性的变幅为 0.07 ~ 0.71, PIC 值的变幅为 0.07 ~ 0.66^[18]。冯芳军等分析了 3 个测序品种的粳粳背景,尝试筛选并利用 InDel 和 SSR 标记发现粳粳特异标记,并利用 SSR 标记对杂交早稻的种子纯度进行鉴定^[19]。

2009 年完成了玉米基因组的测序,发现在结构复杂的玉米基因组中存在大量的 InDel 变异。因此,多态性 InDel 标记在玉米研究中得到了快速发展。Lai 等对中国 5 个优良玉米自交系和兰卡斯特群代表自交系(Mo17)进行重测序后,发现了 30 178 个长度为 1 ~ 6 bp 的 InDel 位点,其中 571 个富集在基因编码区^[20]。Yan 等通过对 8 个玉米自交系的分析发现,在玉米基因的 3'UTR 端广泛存在 InDel 位点^[21]。张体付等在玉米全基因组水平对 InDel 标记进行挖掘并对其在玉米杂交种纯度鉴定中的应用进行分析^[22]。葛敏等通过生物信息学方法对玉米全基因组进行扫描,发现了可用于开发分子标记的位点,并运用 InDel 标记对玉米杂交种纯度进行了鉴定。在瑞德群代表自交系 B73 全基因组序列和二代测序组装 Mo17 邻接片段中发现约 40 000 个 InDel 位点,其中大约有 11 400 个 InDel 位点的两侧可被用于设计位点转化性引物^[23]。

虽然 InDel 标记来源广泛,基因型判别简单并快速,但是利用 InDel 标记解析玉米遗传组成差异等方面的报道较少。本研究拟采用来源不同的 327 份玉米自

交系的全基因组重测序信息,根据高多态信息含量(PIC)筛选出多态性高的 InDel 位点。利用高多态性 InDel 位点进行系统发育树构建和遗传结构推断,并与 SNP 位点获得的信息进行比较。根据 InDel 信息开发 InDel 标记,并进行 PCR 验证和玉米种子纯度分析。

1 材料与方法

1.1 供试材料及全基因组测序数据

以来源不同的 327 份玉米自交系为供试材料,从 NCBI(www.ncbi.nlm.nih.gov)的 3 个 SRA 数据库(SRA049859、SRA051245 和 PRJNA260788)中下载其全基因组的二代原始序列。采用 Q20 标准对供试材料的测序数据进行过滤,然后利用 BWA 软件处理并通过 SAMtool 软件对结果进行整理,获得 InDel 和 SNP 变异位点。以黄淮海地区大面积推广的 64 份优良玉米杂交种,改良瑞德群代表自交系 Zheng58 以及四平头群代表自交系 Chang7-2 为试验材料,用于 PCR 验证。

1.2 遗传组成差异分析

InDel 和 SNP 检测后,采用编写的 Perl 脚本和 VCFtools 软件计算最小等位频率(MAF)、多态信息含量(PIC)。利用 TASSEL V5.0 软件并使用邻接算法(N-J)^[24]计算 327 份玉米自交系之间的遗传距离并构建聚类图,进行类群划分。运用 FastStructure 软件^[25]中的 Python 脚本(Structure.py)对 327 玉米自交系的群体遗传结构进行分析。确定亚群数时,选择混合模型和相关等位基因频率模型,依次设定亚群数目(K)为 2 ~ 15,每个 K 值进行 5 次独立的运算,获得每个 K 对数似然函数的平均值,再利用 FastStructure 软件中的 Python 脚本(ChooseK.py)确定群体的最佳亚群数,依据 Patterson 等^[26]所描述的操作流程进行主成分分析(PCA)。

1.3 InDel 标记开发和筛选

在 3 ~ 4 叶期,取 64 份优良玉米杂交种以及改良瑞德群代表自交系 Zheng58 和平头群代表自交系 Chang7-2 的叶片,按照卡罗登试剂盒相关步骤提取 DNA,加适量 TE 溶解 DNA,于 4 ℃ 或 -20 ℃ 下贮藏待用。InDel 引物设计由 emboss 软件包中的 eprimer3 完成,所有引物均由上海生工生物科技有限公司合成。对 Zheng58、Chang7-2 以及 64 份杂交种中开发的 InDel 标记进行比较,筛选出扩增特异性强、扩增主带清晰的共显性标记。

1.4 PCR 扩增和电泳多态性检测

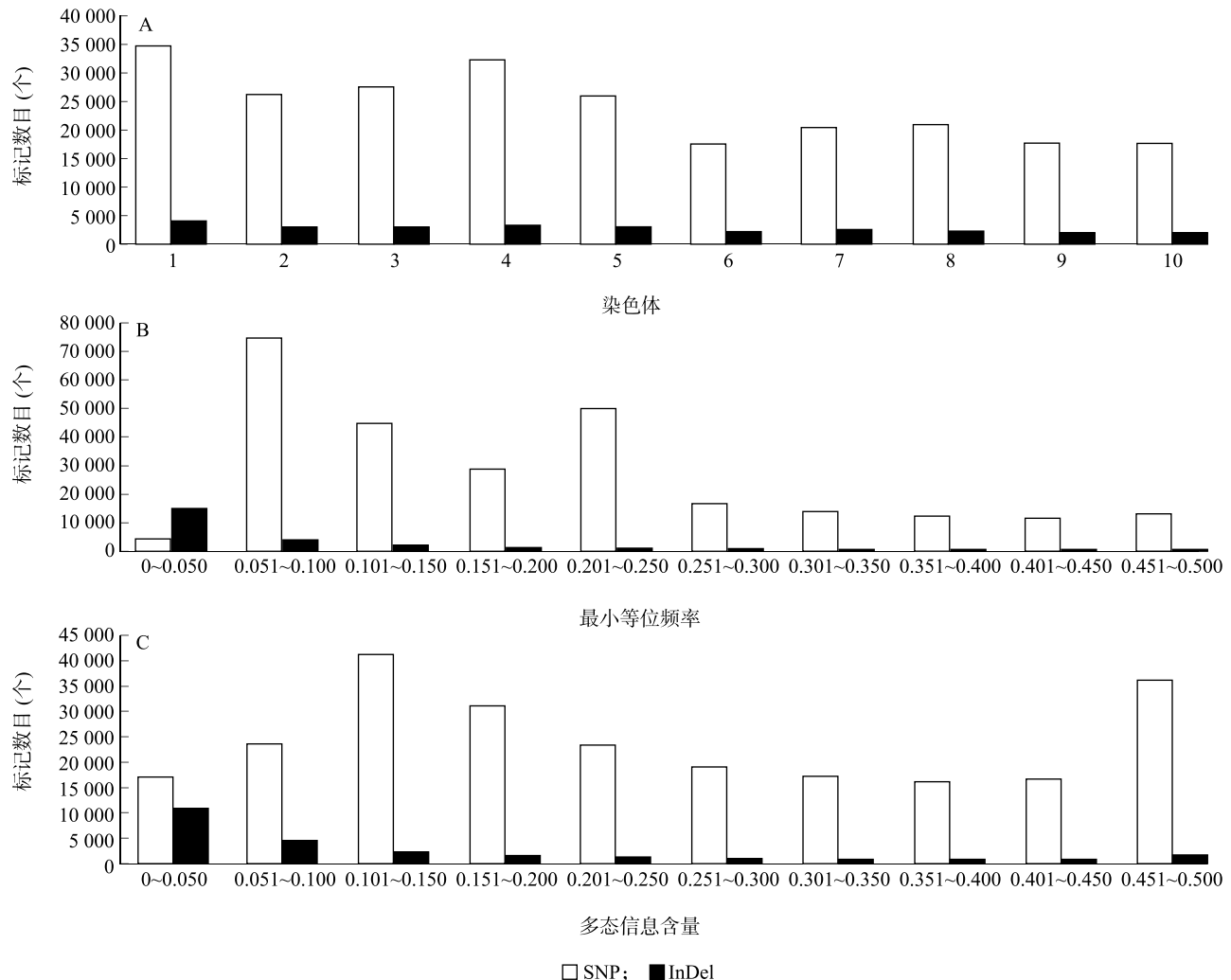
PCR 反应的总体积为 25 μl , 包含 5 pmol 两侧引物, 2.5 μl 10 \times Buffer, 1.25 nmol dNTP, 1 U *Taq* 聚合酶和 40 ng 模板 DNA, 上覆 20 μl 矿物油。扩增程序为: 在 94 $^{\circ}\text{C}$ 下预变性 3 min, 然后进入扩增循环 (94 $^{\circ}\text{C}$ 变性 30 s, 56 $^{\circ}\text{C}$ 或 58 $^{\circ}\text{C}$ 退火 40 s, 72 $^{\circ}\text{C}$ 延伸 40 s), 循环 35 次后在 72 $^{\circ}\text{C}$ 下延伸 5 min。取 5 μl PCR 产物在 2% 琼脂糖凝胶电泳中分离, 溴化乙锭显色, 凝胶成像仪上观察、照相并记录。

2 结果与分析

2.1 二态性插入或缺失 (InDel) 位点的鉴定

利用高通量测序数据开发了 25 847 个二态性 InDel 位点和 239 942 个 SNP 位点, 这些变异位点基

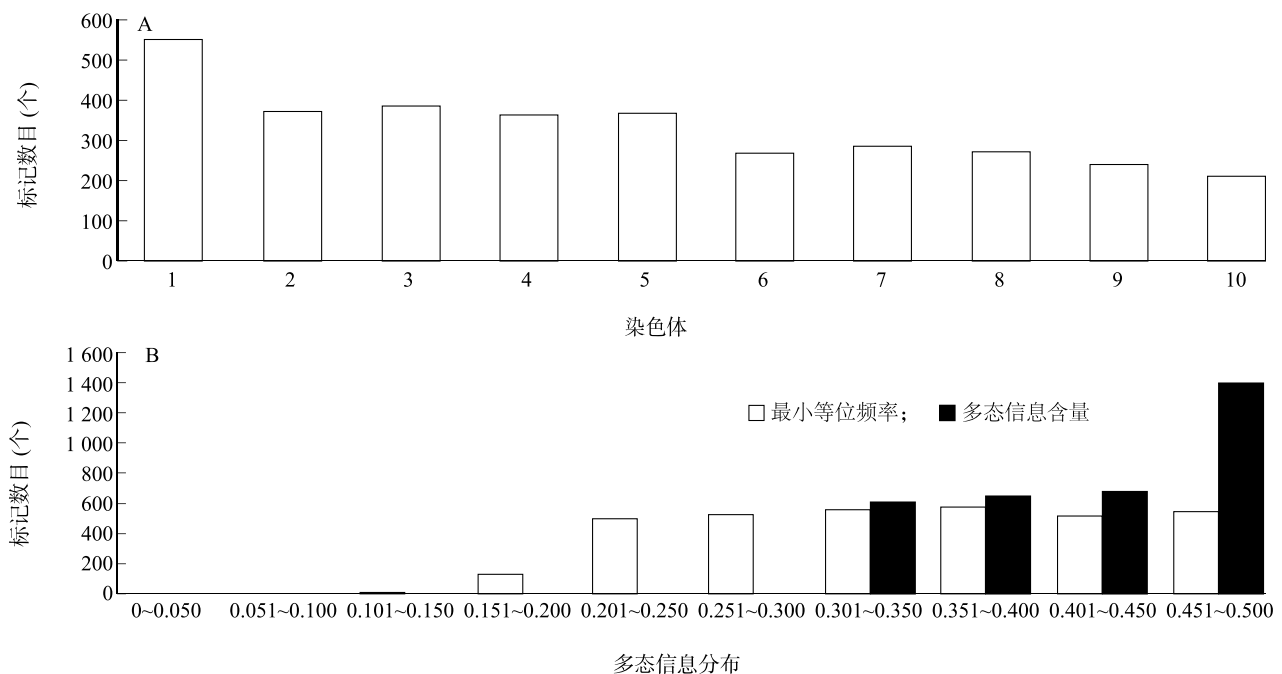
本均匀分布于 10 条染色体上 (图 1A)。利用 Perl 脚本对变异位点进行总体评价, InDel 位点的 *MAF* 的变化范围为 0~0.50, 平均值为 0.09, SNP 位点的 *MAF* 平均值为 0.24 (图 1B)。InDel 位点的 *PIC* 值变化范围为 0~0.50, 平均值为 0.24, SNP 位点的 *PIC* 平均值为 0.19 (图 1C)。结果表明, InDel 位点较 SNP 位点特异性高且稳定性好。以 *PIC* 值 ≥ 0.30 为标准, 对上述检测的总 InDel 位点进行过滤, 筛选出 3 304 个多态性高的二态性 InDel 位点, 这些位点也均匀分布于 10 条染色体上 (图 2A), 其 *PIC* 值的变化范围 0.30~0.50, 平均值为 0.42, 其中有 664 个标记的 *PIC* 值 ≥ 0.49 (图 2B)。结果 (图 2B) 表明, 过滤后的 InDel 位点不仅多态性高 (*PIC* ≥ 0.30), 且稳定性好 (*MAF* 值 ≥ 0.10)。



A: InDel 和 SNP 位点在基因组上的分布; B: InDel 和 SNP 位点上的最小等位基因频率分布; C: InDel 和 SNP 位点上的多态信息含量分布。

图1 327 份自交系基于 25 847 个 InDel 和 239 942 个 SNP 位点上的遗传多态性

Fig.1 The genetic diversity of 327 maize inbred lines based on 25 847 InDels and 239 942 SNPs on ten chromosomes in maize



A: 3 304 个 InDel 位点在基因组上的分布; B: 327 份自交系基于 3 304 个 InDel 位点上的最小等位频率和多态信息含量分布。

图 2 327 份自交系基于 3 304 个 InDel 位点的多态性分布

Fig. 2 Polymorphic information content among 327 maize inbred lines based on 3 304 InDels

2.2 系统发育及群体结构分析

基于 3 304 个多态性高的二态性 InDel 位点, 采用 N-J 法构建了系统发育树。结果(表 1、图 3A)显示, 327 份玉米材料被划分为 8 个不同的类群; 第 I 类为改良瑞德群(Improved Reid), 代表自交系为 478、Zheng58、5003 和 Tie7922 等; 第 II 类为瑞德群(Reid), 代表自交系为 B73 和 A632 等; 第 III 类为兰卡斯特群(Lancaster), 分为 2 个亚群, 代表自交系分别是 Mo17 和 OH43; 第 IV 类为四平头群(SiPingTou), 代表自交系为 Chang7-2、Huangzaosi、Huangyesi 和 Huangyesi3 等; 第 V 类为 PB 群, 代表自交系为 P138、Qi319、Shen135、Shen137 和 Yu87-1 等; 第 VI 类为热带种质群(Tropical), 代表自交系为 Suwan1611 和 Four-row 等; 第 VII 类为爱阿华马齿种质群(Iodent), 代表自交系为 207 等; 第 VIII 类为其他纬度种质群(Other latitude), 代表自交系有 F2 和 F7 等。划分的结果与系谱来源高度一致, 说明多态性高的二态性 InDel 可以代替总 InDel 进行类群划分。这些二态性 InDel 与 SNP 间遗传距离的相似性高达 84.40%, 表明多态性高的二态性 InDel 划分类群的结果与 SNP 划分类型的

结果也是高度一致的。

利用 FastStructure 软件分析了多态性高的二态性 InDel 位点的群体结构, 结果(图 3C)显示, 327 份玉米材料也被划分为 8 个亚群, 其中有 238 份(72.78%)被划分到 8 个亚群中, 而剩余的 89 份(27.21%)则没有明确的类群归属特性, 形成了一个混合群, 其推断的群体结构与 SNP 推断的结构(表 1)高度一致, 系谱来源也一致。因此, 本研究利用多态性高的二态性 InDel 所揭示的遗传变异与 SNP 获得的信息是无差异的, 具有一定的代表性。

根据群体结构分析结果, 本研究还基于多态性高的二态性 InDel 对 327 份材料进行了 PCA 分析, 将 8 个亚群按 PC1 (Principal component 1) 和 PC2 (Principal component 2) 的坐标轴方向进行显示, 结果(图 3B)表明, 8 个亚群基本被划分开, 爱阿华马齿种质群、其他纬度种质群和兰卡斯特群与其他 5 个群之间几乎没有重叠区域, 但是其他 5 个类群间有重叠分布区域, 尤其是瑞德群与改良瑞德群之间, 以及热带种质群与四平头群之间的重叠分布区域较多。

表 1 327 份自交系群体分析结果的统计
Table 1 Summary of clustering analysis among 327 maize inbred lines

类群	239 942 个 SNP 位点		25 847 个 InDel 位点		3 304 个多态性高的 InDel 位点		代表自交系
	样本大小	占总样本比例 (%)	样本大小	占总样本比例 (%)	样本大小	占总样本比例 (%)	
四平头	16	4.89	15	4.59	15	4.59	Chang7-2、Huangzaosi、Huangyesi3、Huangyesi
瑞德	45	13.76	46	14.07	39	11.93	B73、A632
改良瑞德	37	11.31	18	5.50	27	8.26	478、Zheng58、5003、Tie7922
兰卡	30	9.17	31	9.48	28	8.56	Mo17、OH43
PB	14	4.28	31	9.48	15	4.59	Qi319、P138、Shen137、Shen135、Yu87-1
热带	79	24.16	78	23.85	74	22.63	Suwan1611、Four-row
爱阿华马齿	17	5.20	16	4.89	19	5.81	207
其他维度	20	6.12	20	6.12	21	6.42	F2、F7
混合	70	21.00	72	22.01	89	27.21	

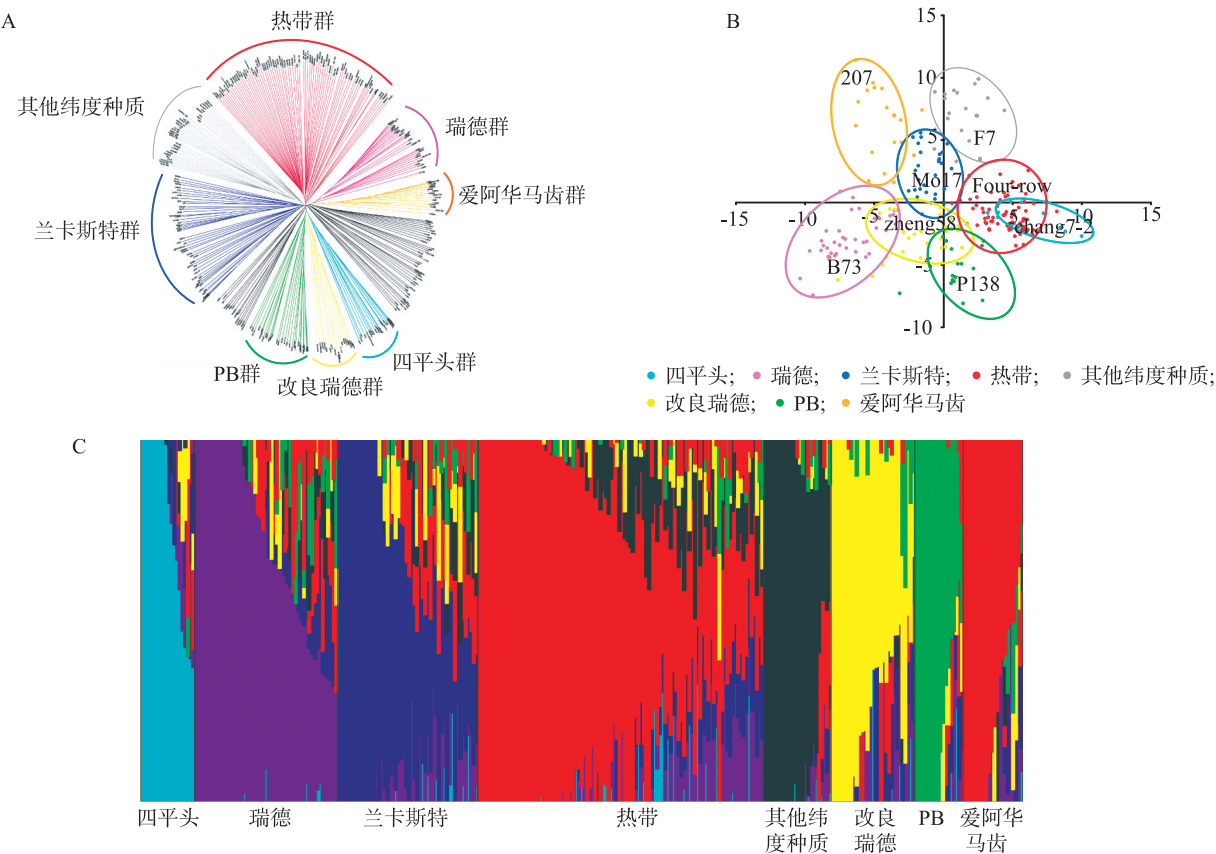


图 3 327 份玉米自交系基于 3 304 个 InDel 位点的群体遗传分析
Fig.3 Population analysis for 327 maize inbred lines based on 3 304 biallelic InDels

2.3 基于 InDel 标记鉴定杂交种的杂合度

基于 3 304 个多态性高的二态性 InDel 位点,本研究设计了 53 对 InDel 引物(表 2),对 64 份玉米杂交

种、自交系 Zheng58 和 Chang7-2 进行 PCR 扩增,各对引物均获得了扩增产物,所有引物都可以清楚揭示自交系 Zheng58 和 Chang7-2 DNA 之间的多态性,且仅

扩增 1 条条带,而在杂交种间大多数 InDel 引物扩增 2 条特异条带。因此,利用开发的 InDel 标记,通过计算 64 份杂交种的杂合率(即多少个杂种表现为杂合位点)鉴定杂交种的杂合度。例如 JAAS3704 和

JAAS3888 标记鉴定 64 份杂交种的杂合率分别为 0.687 5 和 0.703 1(图 4)。开发的 53 个 InDel 标记不均匀分布在玉米基因组上(表 2),其鉴定杂交种的杂合率为 0.245 3~0.903 2,平均杂合率为 0.569 2。

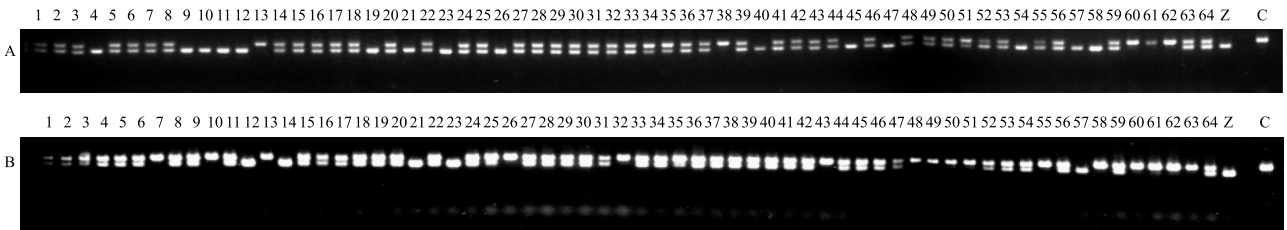
表 2 InDel 标记信息

Table 2 The information of InDel markers

标记	引物序列		染色体	扩增产物长度 (bp)		差异大小 (bp)
	正向引物 (5'→3')	反向引物 (5'→3')		Zheng58	Chang7-2	
JAAS3528	GTGGCGACGACAATGCTATGG	ACTGTGGAGGGTGCATCTGTG	8	314	443	129
JAAS3530	TAGCGAGTTCGAGGCATCAGG	ATGACCATCAGTTGCGTCTTCG	8	378	312	66
JAAS3532	CCCAACACGGTCTGCATCTTG	CACCACACTAGCGTCAACCAAC	8	295	326	31
JAAS3533	GGCATCGTAGCATCCCTCTCC	CATCAAAGCAAGAAGGCAACCC	8	165	201	36
JAAS3534	CGCCAGCACCCAACAAATGTC	CATCCGATCATCACCACACAGG	8	232	214	18
JAAS3537	CGTGCTGCTAGGTTGATGCC	CGGCTTGCTGGCTTTAGATCC	8	272	311	39
JAAS3538	CTCCAAGTTTGGCGTCAATCTG	CTCACCCTGCATCAGACCTAC	8	424	373	51
JAAS3539	ACGCTGCTTCTGCTACATCATC	CTCCGAGTATGCTTCTGCCATC	3	407	379	28
JAAS3540	GTCACATGCCTGGCTACATTCC	CTGCTGCTTGCCACTTGAAGAG	2	320	304	16
JAAS3541	TAGCCACGTTCCGAGAGACTG	GACACACCTGCCAACTGAGATC	7	378	350	28
JAAS3542	GCCATAGCGACCACCAACAATC	GACCACAGGGATGGGAAACTTG	3	199	186	13
JAAS3543	GTCAAGGAGGCTCAGAACCATG	GCCAGAATGACACCCTGATAGC	1	208	220	12
JAAS3546	CACGGATCGCCATCAGTTGAC	GCTGCTGTTTGTGCGCTGAG	8	276	257	19
JAAS3549	GACAGAGTGACAGCAAGGTTGG	GTGCGTGATGACCAGGATGC	10	191	276	85
JAAS3554	TCAAACAGGGTGGACGAAGAGG	TCACTACGCACTCGATCCCTTC	8	213	198	15
JAAS3555	CTATGCTCTTCTTTCCGACTG	TTCCTTACCTTGCTGTTCTTG	3	390	440	50
JAAS3635	AGGATGGTGATGATGGGAATGC	TTGTGGGTAGACTGCTCATTGC	2	217	236	19
JAAS3637	CCGAGTTGTTGCCTCAAATCCC	CTGCTTGCTGGTGGTTGTTCC	5	187	171	16
JAAS3639	GTGGCTTTATTGCTGGTCGTC	TTCAGTCCACGCATAGCATAGG	8	152	126	26
JAAS3641	TCACTACGCACTCGATCCCTTC	TCAAACAGGGTGGACGAAGAGG	8	213	198	15
JAAS3642	CATCAAAGCAAGAAGGCAACCC	GGCATCGTAGCATCCCTCTCC	8	165	201	36
JAAS3650	CGATGGTGGTCTTCTTCCTCAG	GTGATGTCGTCCATTCCGTGTG	8	151	135	16
JAAS3700	TTCTCATGGCTCGACTGTAACTG	TTAGCAACTGGAGGGTGTGTGAC	1	270	253	17
JAAS3702	GGATGCCCACTAACGAAATTGTACC	GGTGTGCTGCTAGTCAGTTTCATG	1	245	217	28
JAAS3703	CGCCCTGTGCTTGATTGTTTC	ATCGCTTAAACGCCAAAGTGCTC	1	197	240	43
JAAS3704	AGCCAACTGGATCTGTGGTGTAAG	GGTAGGTGCCTCTAGGTGTTGTAC	1	226	267	41
JAAS3705	TGGTAGCAGCGATGGTGCTAAG	CCAACCAACCAGACTCAAGAATGC	1	134	119	15
JAAS3706	GCCCTTTATCAAACCAACCCACAC	GCGTCAGAGCTAGTTCTATGGTC	1	122	151	29
JAAS3707	CATCTACTGCCAAGTCAATGCCAAC	TCGTGGAGGTTGTAAATGAGGATGG	1	245	270	25
JAAS3708	CCAGAGGAGATTAGAGGGACTAGGG	ATCCGCCGACCTCATTACATAGC	1	173	194	21
JAAS3713	TTCACCTGGTGCCTTGTTATCC	TATGTGTTGGAGTATGGCTCAGACG	1	324	255	69
JAAS3714	GCTCAGAATGGAACGGCATCATG	GCTTCTTCCGCTGAACTGCTAAC	10	168	283	115
JAAS3715	CCTACCCACAGACGCCCTTTG	TGGCTCTTTACGTGGCTATTCTCTC	10	158	113	45

续表 2 Continued 2

标记	引物序列		染色体	扩增产物长度 (bp)		差异大小 (bp)
	正向引物 (5'→3')	反向引物 (5'→3')		Zheng58	Chang7-2	
JAAS3724	AGAACACGACCGAGACACCAATC	GAGCACATCCAGACACCATTGAAG	10	128	290	162
JAAS3728	AGTGGCTTTATGCAGTTGGTTGTAC	TTCTGTCTTGGTGCTTTCCTAC	2	394	566	172
JAAS3730	GCCCCACTGACCACTGTGAAGAG	AGAACAAGCTGAGTAGGACAGAACC	2	101	224	123
JAAS3731	GAAGAGGAGGCTAATTCAGCAAGTG	CAACCACACCACTCTATGCAACATG	2	150	215	65
JAAS3735	ACGATCACCGGGAGCACATAATAG	GAGGCGTGGTTGGGTCATGG	2	163	138	25
JAAS3743	CCAATGATGGTGAGGAAGGTAGGG	GCTTCAGGCTCTGGCAGGTG	3	111	79	32
JAAS3753	AGCACGGACAGTCTCAAGGTTG	GCTGACCGATGCCTCCTGAAG	4	297	336	39
JAAS3773	ACCATCCACCGAACCAATCAGAC	TCTCTAGCAAACGCCCAAACG	6	250	197	53
JAAS3780	GCTCGTCACTCACCGTCGTC	TAGGCCCACACCAACAGTTCATG	8	244	320	76
JAAS3785	GTCGGCACAAGGATGAATGAAGATG	TTGTAAACCCCTTCCCACTCTAACC	8	169	125	44
JAAS3786	CATTTGTCTGCTGTCACCTGTCAG	AGTCGTGGATCGTCAGAAGGAATC	8	145	102	43
JAAS3787	CGGTCAGCAGCATGAATGAATGAG	CAGAGAGGGACGGAAATCAGACAG	8	179	199	20
JAAS3790	AACCTTTTGTGGGAGCAGTTACTTG	TGTTGGTGTGCGCCCTGTTGTC	8	131	104	27
JAAS3791	TCCAGTCCCACCCAAATGATGAAC	TGTTGATGTAGCATAGGAGACCAC	9	172	383	211
JAAS3793	TTGCTTCGTATCACCAATCGTCATG	GGGTCCCCAGCTCGTACTCC	9	227	390	163
JAAS3794	TGTTTCGTCCTCTCTGTTGAATGG	GGGGTGATTGAAGTTGGGCAAAC	9	108	85	23
JAAS3797	ACCCCTACTCCCTCCTCATAACG	GTTGCTTGTGCTGCTGCTCTG	9	200	166	34
JAAS3798	CCGCCATGTCATTGTCAGCTTG	GAGATGGACACCGCCGAGAAG	9	260	282	22
JAAS3886	ACCCTCTCCCTCCATCAAG	CCAGGTTGCCCGTCGAGAAG	1	268	185	83
JAAS3888	AAGGGCGAGGTTGCTCTTGTGTC	GTCCCACGACTCCTTGTCTCTG	1	287	262	25



Z 和 C 泳道分别为 Zheng58 和 Chang7-2 条带;1 ~ 64 泳道分别为 64 个杂交种样本条带。
图 4 高多态性 InDel 标记 JAAS3704 (A) 和 JAAS3888 (B) 在 64 份杂交种中的琼脂糖凝胶检测结果

Fig. 4 The agarose gel electrophoresis of PCR with high polymorphism InDel markers JAAS3704 (A) and JAAS3888 (B) in 64 hybrids

3 讨论

插入和缺失 (InDel) 在 DNA 或氨基酸序列水平上发生的频率仅次于残基替换。随着比较基因组学的深入研究和玉米基因组测序及重测序计划的完成,玉米 InDel 多态性标记得到了更快的发展。已有研究发现,在玉米基因编码区及 3'UTR 端广泛存

在 InDel 位点^[27-29],这些位于基因编码区及调控区的 InDel 位点可能与基因的功能存在联系,更容易开发为功能性 InDel 标记。因此,在玉米育种研究中,开发 InDel 位点对于分子标记辅助育种非常重要。本研究以来源不同的 327 个基因组序列信息为背景,利用生物信息学方法挖掘 25 847 个二态性

InDel 位点,其均匀分布于玉米基因组中。为了降低高通量 InDel 位点进行分子设计时的成本并提高工作效率,以 PIC 值 ≥ 0.30 为标准筛选出 3 304 个二态性 InDel 位点,其 PIC 值的变化范围为 0.30 ~ 0.50,平均值为 0.42,说明这些二态性 InDel 位点具有极高的多态性。Lu 等利用 SNP 标记对 770 份玉米自交系的遗传多样性进行分析,结果表明,1 034 个 SNP 位点的 PIC 值为 0.003 ~ 0.375,平均值为 0.259^[30]。吴金凤等利用 1 041 个 SNP 位点对 53 份自交系进行类群划分,SNP 位点多态信息含量值为 0.186 ~ 0.375,平均值为 0.345^[31]。与前人基于 SNP 标记划分类群的结果相比较,本研究筛选出的二态性 InDel 位点具有较高的多态性。

根据系统发育和群体结构分析,将 327 份供试材料划分为改良瑞德、瑞德、兰卡斯特、四平头、热带、PB、爱阿华马齿以及其他纬度种质群,与其系谱相吻合,与前人的报道也高度一致^[30,32-34]。尽管旅大红骨群(LRC)在中国玉米自交系类群划分中也很重要,但由于旅大红骨材料非常少且其多样性可能较为丰富,因此没有被划分到一个具体的类群中。另外,与 SNP 位点获得的信息进行比较,发现多态性高的二态性 InDel 所揭示的遗传变异与 SNP 获得的结果基本无差异,证明这些二态性 InDel 位点在揭示遗传组成差异上是具有代表性的。而群体间的 PCA 分析结果表明,在玉米自交系的 8 大类群中,瑞德群和改良瑞德群之间的关系较近且有区域重叠,这与其系谱关系是相吻合的。而四平头群与瑞德群、改良瑞德群、PB 群、兰卡群、爱阿华马齿群和其他纬度种质群之间的关系都较远,仅与热带群间的关系较近且有较高的重叠区域,表明四平头群中的一些自交系是通过热带血缘改良的。

根据多态性高的二态性 InDel 信息,本研究开发了 53 个 InDel 标记,经过 PCR 试验验证后,发现这 53 个 InDel 标记可用于鉴定 64 个杂交种的杂合度,具有位点专化性的特点,鉴定结果准确可靠。尽管本研究目前提供的标记数目有限,但这些标记信息为解析玉米不同群体间的遗传差异、基因组变异以及标记辅助育种等方面奠定了基础。

参考文献:

- [1] 唐祈林, 荣廷昭. 玉米的起源与演化 [J]. 玉米科学, 2007, 15(4): 1-5.
- [2] 崔彩红, 房伟强, 李兴锋, 等. 携带抗秆锈病基因 *Sr25*、*Sr26* 小麦新种质的分子标记辅助选育 [J]. 山东农业科学, 2015, 47(4): 13-17.
- [3] 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 [J]. 中国农业科学, 2003, 36(6): 622-627.
- [4] 黄如葵, 黄熊娟, 梁家作, 等. 利用辐射诱变及分子标记辅助筛选进行苦瓜种质创新 [J]. 南方农业学报, 2015, 46(3): 376-380.
- [5] 袁力行, 傅骏骅, WARBURTON M, 等. 利用 RFLP、SSR、AFLP 和 RAPD 标记分析玉米自交系遗传多样性的比较研究 [J]. 遗传学报, 2000, 27(8): 725-733.
- [6] 余玲, 李爱宏, 潘存红, 等. 分子标记辅助选择培育抗病优质晚粳稻品种扬粳 806 [J]. 江苏农业科学, 2014, 42(8): 75-77.
- [7] REIF J C, HAMRIT S, HECKENBERGER M, et al. Trends in genetic diversity among European maize cultivars and their parental components during the past 50 years [J]. Theor Appl Genet, 2005, 111(5): 838-845.
- [8] 杨淑珂, 林红珍, 陈博聪, 等. 分子标记辅助选择在水稻新品种(系)培育中的应用进展 [J]. 山东农业科学, 2014, 46(1): 137-141, 147.
- [9] 王风格, 赵久然, 余花娣, 等. 玉米自交系遗传关系的 SSR 标记分析 [J]. 中国农学通报, 2005, 21(2): 10-14.
- [10] 刘驰, 马增凤, 秦钢, 等. 利用分子标记辅助选择培育水稻广亲和保持系 [J]. 南方农业学报, 2015, 46(3): 365-369.
- [11] MILLS R E, LUTTIG C T, LARKINS C E, et al. An initial map of insertion and deletion (InDel) variation in the human genome [J]. Genome Research, 2006, 16(9): 1182-1190.
- [12] MIKI Y, SWENSEN J, SHATTUCK-EIDENS D, et al. A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1* [J]. Science, 1994, 266(5182): 66-71.
- [13] COLLINS F S, DRUMM M L, COLE J L, et al. Construction of a general human chromosome jumping library, with application to cystic fibrosis [J]. Science, 1987, 235(4792): 1046-1049.
- [14] CLARK R M, WAGLER T N, QUIJADA P, et al. A distant upstream enhancer at the maize domestication gene *tb1* has pleiotropic effects on plant and inflorescent architecture [J]. Nat Genet, 2006, 38(5): 594-597.
- [15] ASHIKARI M, SAKAKIBARA H, LIN S, et al. Cytokinin oxidase regulates rice grain production [J]. Science, 2005, 309(5735): 741-745.
- [16] FU D, UAUY C, DISTELFELD A, et al. A kinase-*START* gene confers temperature-dependent resistance to wheat stripe rust [J]. Science, 2009, 323(5919): 1357-1360.
- [17] WU D H, WU H P, WANG C S, et al. Genome-wide InDel marker system for application in rice breeding and mapping studies [J]. Euphytica, 2013, 192(1): 131-143.
- [18] LIU B, WANG Y, ZHAI W, et al. Development of InDel markers for *Brassica rapa* based on whole-genome re-sequencing [J]. Theor Appl Genet, 2013, 126(1): 231-239.

- [19] 冯芳君. SSR 和 InDel 标记在水稻籼粳亚种分化与杂种纯度鉴定中的应用[D]. 武汉:华中农业大学, 2006.
- [20] LAI J, LI R, XU X, et al. Genome-wide patterns of genetic variation among elite maize inbred lines[J]. *Nature Genetics*, 2010, 42(11): 1027-1030.
- [21] FU Y, WEN T J, RONIN Y I, et al. Genetic dissection of intermated recombinant inbred lines using a new genetic map of maize[J]. *Genetics*, 2006, 174(3): 1671-1683.
- [22] 张体付, 葛 敏, 韦玉才, 等. 玉米功能性 Insertion/Deletion (InDel) 分子标记的挖掘及其在杂交种纯度鉴定中的应用[J]. *玉米科学*, 2012, 20(2): 64-68.
- [23] 葛 敏, 蒋 璐, 张晓林, 等. 利用 Insertion/Deletion (InDel) 分子标记检测玉米互交种混杂的原理及应用[J]. *分子植物育种*, 2013(1): 37-47.
- [24] SAITOU N. The neighbor-joining method; a new method for reconstructing phylogenetic trees[J]. *Mol Biol Evol*, 1987, 4(4): 406-425.
- [25] RAJ A, STEPHENS M, PRITCHARD J K. fastSTRUCTURE: variational inference of population structure in large SNP data sets[J]. *Genetics*, 2014, 197(2): 573-589.
- [26] PATTERSON N, PRICE A L, REICH D. Population structure and eigenanalysis[J]. *PLoS Genetics*, 2006, 2(12): e190.
- [27] FU Y, WEN T J, RONIN Y I, et al. Genetic dissection of intermated recombinant inbred lines using a new genetic map of maize[J]. *Genetics*, 2006, 174(3): 1671-1683.
- [28] LAI J, LI R, XU X, et al. Genomewide patterns of genetic variation among elite maize inbred lines[J]. *Nat Genet*, 2010, 42(11): 1027-1030.
- [29] 吕远大, 李 坦, 石 丽, 等. 基于全基因组重测序信息开发玉米 H99 自交系特异分子标记[J]. *作物学报*, 2014, 40(2): 191-197.
- [30] LU Y, YAN J, GUIMARÃES C T, et al. Molecular characterization of global maize breeding germplasm based on genome-wide single nucleotide polymorphisms[J]. *Theor Appl Genet*, 2009, 120(1): 93-115.
- [31] 吴金凤, 宋 伟, 王 蕊, 等. 利用 SNP 标记对 51 份玉米自交系进行类群划分[J]. *玉米科学*, 2014, 22(5): 29-34.
- [32] WANG F G, TIAN H L, ZHAO J R, et al. Development and characterization of a core set of SSR markers for fingerprinting analysis of Chinese maize varieties[J]. *Maydica*, 2011, 56: 7-17.
- [33] WANG R, YU Y, ZHAO J, et al. Population structure and linkage disequilibrium of a mini core set of maize inbred lines in China[J]. *Theor Appl Genet*, 2008, 117(7): 1141-1153.
- [34] 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[J]. *中国农业科学*, 2012, 45(11): 2107-2138.

(责任编辑:王 妮)