

玉米通用 SSR 核心引物筛选及高通量多重 PCR 复合扩增体系建立

王凤格 赵久然 戴景瑞* 易红梅 匡猛 孙艳美
于新艳 郭景伦 王璐

(中国农业大学国家玉米改良中心, 北京 100094; 北京市农林科学院玉米研究中心, 北京 100097.

* 联系人, E-mail: dajr@cau.edu.cn)

摘要 利用国内 15 个代表性自交系及 15 个来自国家区域试验的玉米杂交种进行实验室复筛, 并对玉米数据库 MaizeGDB 上已公布的 1900 个玉米 SSR (simple sequence repeats) 引物进行文献初筛, 确定了 500 个高多态性引物, 建立了玉米高多态性引物遗传图谱; 并按照在玉米全基因组均匀分布的原则, 从每条染色体上选取 10 个引物, 累计 100 个, 作为一套通用的 SSR 核心引物, 推荐供今后一般研究优先选用. 对这套核心引物按照统一的标准进行重新设计并构建基于毛细管五色荧光检测系统的高通量多重 PCR 复合扩增体系, 已经建立了两套 10 重 PCR 组合, 每套组合均由每条染色体上选取 1 个引物, 累计 10 个构成.

关键词 玉米 核心引物 SSR 多重 PCR 毛细管电泳 荧光标记

开发并筛选一套综合特性好的通用的 SSR (simple sequence repeats) 核心引物, 对玉米种质资源鉴定、杂种优势群划分、遗传多样性分析、品种纯度及真实性检测、品种 DNA 指纹数据库构建等研究均具有重要的价值. 玉米是 SSR 引物开发较早的作物之一, 早期的玉米 SSR 引物主要是从基因组文库或微卫星富集文库开发的, 开发成本高且费时费力^[1]. 随着玉米基因组测序计划的实施及大量 DNA 序列数据的公布, 通过搜索序列数据库的策略进行大规模的 SSR 引物开发成为 SSR 引物的主要来源^[2-4]. 尽管已经开发了 2000 多个玉米 SSR 引物, 然而不是所有的 SSR 引物都具有同等的应用价值, 不同的 SSR 引物在多态性水平高低、扩增质量好坏、是否共显性遗传、扩增单位点还是多位点、是否在基因组上定位及分布位置等方面差异很大, 因此, 必须根据研究的需要对这些引物进行全面评估, 才能确定其利用价值. 目前对已开发玉米 SSR 引物的评估仍然是少量的、不系统的, 削弱了 SSR 标记在玉米中的实际应用效果.

建立高通量自动化检测平台是实现 SSR 标记在实际检测中大规模应用的前提. 然而, 与 AFLP, RAPD, ISSR 等多扩增位点的引物相比, 每分析一个 SSR 引物仅能获得一个位点的基因型信息, 数据分析的通量比较低. 为了提高每次 SSR 分析的信息量, 在一些作物上探索了多重 PCR 研究, 该方法与毛细管电

泳荧光自动检测系统相结合, 可显著提高单次 SSR 分析的信息量, 降低利用 SSR 引物进行大规模分型的成本. 然而, 在玉米上进行多重 PCR 研究的较少, 构建的多重 PCR 引物组合一般仅适用于特定的研究需要, 通用性较低, 且在同一扩增组合中的引物数较少, 一般为 2~4 个^[5-7].

为解决上述问题, 本研究对已公布的玉米 SSR 引物进行全面筛选评估, 以确定一套通用的玉米核心 SSR 引物名单, 并通过核心引物进行重新设计、评估及组合, 建立基于毛细管电泳自动荧光检测的高通量多重 PCR 复合扩增系统.

1 材料与方法

() 材料. 引物的实验室复筛、新设计引物与原引物对比实验、多重 PCR 组合体系的实验验证采用的实验材料为 15 个国内常用的代表性自交系: 黄早四, 昌 7-2, 京 24, 京 89, 87-1, X178, 齐 319, 京 501, 丹 341, 太 411, 黄 C, SW1611, B73, Mo17 和 E28; 以及 15 个国家区域实验玉米杂交种: 本 2204, 长城 303, Q2104, 丹科 1 号, 奥试 3111, 东 0201, 恒丰 1 号, M3309, ND5598, CM10, ND311, 东 4326, 京品 202, 唐玉 16 和 DH3702. 15 个自交系均经过连续套袋自交提纯. 15 个杂交种名单是根据田间形态鉴定及室内一致性检测的结果, 从国家区域实验品种中选用的品种间遗传差异较大、品种内部一致性高的单交种^[8].

() 引物多态性的筛选. (1) 根据已有文献资料初筛. 根据MaizeGDB上 1900 多个SSR引物的扩增图片 (<http://www.maizegdb.org/ssr.php>) 以及包括来自 CIMMYT^[9,10]、美国^[11~13]、法国^[7]、中国^[14,15]、日本^[16]、印度^[17]等国家或机构对部分SSR引物的多态性评价, 结合考虑引物在染色体上的分布情况, 初筛出了 800 个候选引物(结果未显示). 入选指标如下: 多态性要求等位基因数(扩增谱带数) ≥ 3 或PIC值 ≥ 0.65 ; 带型分布要求不同谱带之间大小容易区分; 染色体分布位置要求引物在染色体上已经定位, 且位置是惟一的, 所有候选引物基本上能够均匀覆盖染色体全基因组. 引物扩增效果暂不作为引物是否候选的指标. (2) 实验室复筛. 利用 15 个自交系对上述 800 个候选引物的多态性进行实验室复筛, 然后对入选引物进一步利用 15 个杂交种评估引物位点的杂合子表现. 根据实验结果和文献资料上的信息对引物进行综合评价, 确定出了 500 个高多态性候选引物.

() 候选引物染色体分布及一套通用核心引物的确定. 从 MaizeGDB 上下载涵盖 SSR 引物位点比较全面的玉米分子遗传图谱 IBM2 2004 neighbors-frame, 确定高多态性候选引物在该图谱上的位置, 并采用中国农业大学自编的宏程序绘制整合了所有 500 个高多态性引物的分子遗传图谱.

根据引物在整合图谱上的位置, 依据均匀分布的原则, 在每条染色体上确定 10 个, 累计 100 个引物位点作为一套通用核心引物.

() 引物重新设计及通用 PCR 扩增程序的建立. 对 100 个核心引物位点按照统一的要求, 重新设计引物, 使得所设计的引物具有相似的扩增条件、扩增清晰稳定、扩增产物大小在期望的范围内, 便于下一步组合多重 PCR.

首先在 NCBI (<http://www.ncbi.nlm.nih.gov>)、MaizeGDB (<http://www.maizegdb.org>) 及 PlantGDB (<http://www.plantgdb.org>) 等网站上, 用 BLAST 程序搜索候选引物位点对应的原始基因组序列. 利用两个常用的引物设计软件 Primer Premier 5.0 和 Oligo 6.22 进行引物的重新设计及评估, 同一多态性位点设计的新引物用“原引物名称 + 设计者代号 + 序号”的方式重新命名.

由于新引物采用了较高的 T_m 值, 经过摸索, 确定了新引物最佳退火温度为 68 , 并建立了 68~94 二步法扩增程序: 94 预变性 5 min; 94 变性 40 s, 68

退火 80 s, 35 个循环; 72 延伸 10 min.

对同一引物位点的新设计引物与原引物采用上述统一的扩增程序进行实验检测, 比较二者在扩增带型和扩增效果的变化. 对扩增谱带的位置与预期的位置不一致或扩增效果较差的新引物重新设计. 通过了上述软件评估和实验评估的引物作为代替原引物的入选新引物, 并作为下一步的多重 PCR 组合用的候选引物.

() 荧光多重 PCR 组合的建立. 对多重 PCR 组合引物设计的具体要求为: (1) 每套引物组合是从玉米 10 条染色体上各选取 1 个, 共 10 个引物构成; (2) 组合方式为 3+3+2+2, 即按照片段大小范围, 将每套引物按上述方式分成 4 组, 不同组的引物标记不同的荧光. 要求标记相同荧光的同组合引物的扩增产物范围没有交叉; (3) 每套引物间相互作用尽可能低, $\Delta G < 13$. 10 重 PCR 组合的建立过程: 根据候选引物在玉米染色体上的位置, 从玉米 10 条染色体上各选取 1 个引物; 根据这些引物的扩增带型分布特点, 大致确定各个引物的期望设计位置; 利用 Primer Premier 5.0 和 Oligo 6.22 进行引物设计和评估, 获得符合核心引物设计要求的新引物; 对拟组合的引物利用 Primer Premier 5.0 进行多重 PCR 软件评估并进行实验评估, 排除组合时产生显著的引物间相互作用的组合类型.

PCR 扩增同新引物, 电泳检测在 ABI3730XL 遗传分析仪上进行. 引物按设计要求分别标记 FAM, NED, PET 及 VIC 四种荧光之一. 荧光引物由美国 ABI 公司合成.

2 结果与分析

2.1 高多态性引物的确定

通过文献信息和实验室复筛结果, 累计 500 个高多态引物位点入选. 由于引物扩增效果与引物序列设计的优劣有直接关系, 对扩增较差的引物位点可以根据该位点的简单重复序列两侧的保守序列重新设计引物进行扩增, 因此, 在这一步筛选中, 仅考虑了引物多态性和带型分布特点, 没有考虑引物扩增效果.

入选的引物基本符合候选标准, 部分位点由于没有完全符合条件的引物, 为填补空缺而降低条件入选了某些引物, 如引物 phi041 (位于 Bin 10.00). 没有入选的引物, 或者多态性低(如 umc2259), 或者带

型分布密集不易统计(如 nc030), 或者染色体定位位置不一致(如引物 umc1593, 定位在玉米染色体 3.05 和 7.03 两个位置).

此外, 根据杂交种的扩增结果进一步评估了引物的杂合子表现, 包括位点的杂合率和扩增是否异常(如三带、杂合带型不对称扩增等). 并进一步剔除了杂合率偏低和扩增异常的部分引物.

通过以上筛选过程最终确定的候选引物, 多态性水平较高, 平均 PIC 值达 0.74, 平均等位基因数达 5.3 (表 1). 不同研究工作可以根据其特殊要求, 从这套引物中进一步选取合适的引物, 而不必从已公布的 1900 个引物中选取, 可大大减轻研究者前期筛选的工作量.

2.2 一套通用核心引物位点的确定

为了便于根据引物在染色体上的位置进一步确定通用核心引物名单, 建立了整合所有 500 个高多态性引物的分子遗传图谱 Maize HP-SSR composite map (图 1), 其中 460 个是已定位在 IBM2 2004 neighborsframe map 上的引物, 另外 40 个是根据引物在其他图谱上的位置, 经过估测后整合到 IBM2 2004 neighborsframe map 的, 该图谱用于显示候选引物的染色体分布, 为核心引物位点的选择提供参考.

Maize HP-SSR composite map 全长 7167 cM, 覆盖 IBM2 2004 neighborsframe map (全长 7422 cM) 的 97%, 平均引物密度为 14.8 cM, 每条染色体上分布的引物数量有一定差异, 最少的为 35 个(10 号染色体), 最多的为 59 个(1 号染色体). 大于 70 cM 的空缺区域有 5 处, 分布于 1 号染色体(81.5, 90.9, 93.6)、2 号染色体(95.8)和 9 号染色体(201). 空缺区域主要是由于该区域没有开发引物或没有入选多态性较高的

引物, 需要开发新的多态性高的引物才能填补这些区域. 某些染色体区域候选引物较多, 如在染色体 6.00 和 6.01 位置共有 11 个引物, 这些引物均完全符合候选条件, 在下一步的引物重新设计中通过比较其设计的难易进一步选择.

综合考虑引物的多态性、染色体分布及引物重新设计情况, 在每条染色体上均匀选择 10 个引物, 累计 100 个引物作为一套通用的核心引物(图 1). 这套通用核心引物多态性较高, 具有较高的鉴别能力; 在玉米染色体上分布均匀, 最大限度地降低了引物之间的连锁, 因此可作为品种鉴定、杂种优势群划分、遗传多样性分析、基因定位等研究初期优先选用的引物.

2.3 核心引物位点的重新设计及多重 PCR 组合的建立

为了进行引物重新设计, 必须获得引物位点对应的原始基因组序列. 通过对 500 个候选引物的搜索, 最终在 NCBI 上下载到 390 个引物位点的原始序列, 在 MaizeGDB 上下载到 36 个, PlantGDB 上下载到 6 个, 累计 432 个, 还有 68 个没有下载到对应的原始基因组序列. 对已下载原始基因组序列的引物位点进入下一步的引物重新设计; 对未查找到对应的原始序列的引物可考虑寻找其他网站进行序列的查询, 或者通过从玉米基因组 BAC 库中找到对应的阳性克隆后测序获得.

引物的重新设计与多重 PCR 组合的构建是同时进行的. 从 100 个通用核心引物中按照均匀分布的原则, 每个染色体上均匀选取 2 个引物, 累计 20 个引物, 通过引物重新设计及实验验证, 已构建了两套 10 重 PCR 组合.

表 1 500 个高多态性候选引物的多态性信息及在染色体上的分布数量

染色体	长度/cM	引物数目	平均密度/cM	PIC 值(最大; 最小; 平均)	平均等位基因数
1	1137	59	19.3	0.70; 0.83; 0.76	5.6
2	770	54	14.3	0.64; 0.78; 0.73	5.0
3	842	58	14.5	0.63; 0.83; 0.74	5.3
4	804	50	16.1	0.64; 0.84; 0.74	5.1
5	676	53	12.8	0.64; 0.81; 0.72	5.5
6	579	46	12.6	0.66; 0.84; 0.77	5.7
7	644	52	12.4	0.60; 0.77; 0.70	4.4
8	632	47	13.4	0.60; 0.75; 0.75	5.8
9	805	46	17.5	0.64; 0.84; 0.73	5.9
10	533	35	15.2	0.64; 0.81; 0.73	5.0
合计	7422	500	14.8	0.64; 0.81; 0.74	5.3

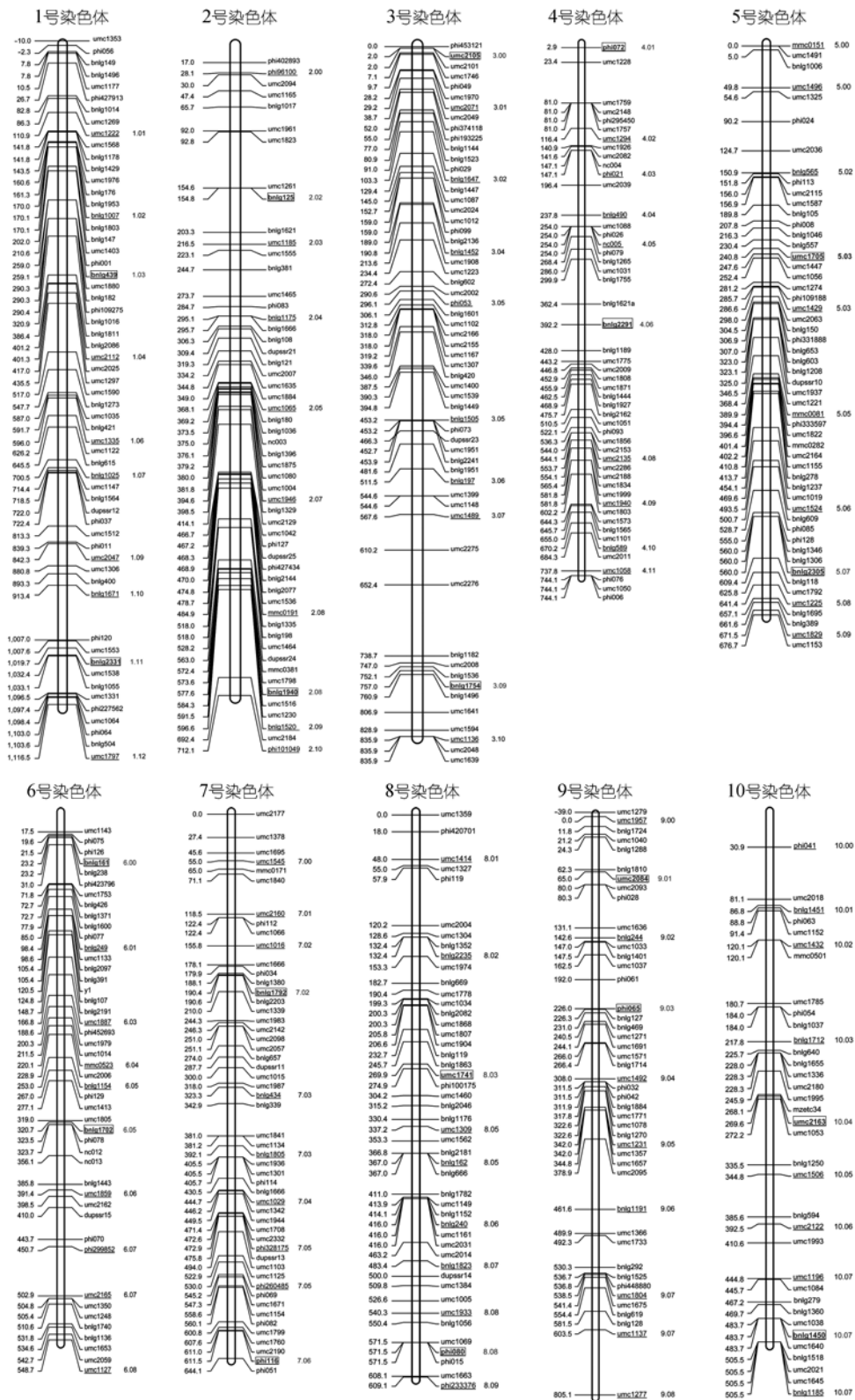


图 1 500 个玉米高多态性 SSR 引物整合分子遗传图谱(maize HP-SSR composite map)

下划线处为每条染色体上均匀选取的 10 个核心引物, 其中加框的为已组 10 重 PCR 的引物

由表 2 可见, 每套 10 重组均按照 3 + 3 + 2 + 2 的分组模式构建. 采用这种方式的好处在于: 如果使用毛细管电泳五色荧光检测系统, 完成 10 个引物的检测只需进行 1 次 10 重 PCR 扩增及 1 次电泳检测; 即使使用变性聚丙烯酰胺凝胶检测系统, 完成 10 个引物的检测也只需进行 4 次 PCR 扩增和 4 次电泳检测, 与单一扩增相比, 仍然能够提高检测效率 2.5 倍.

其中 14 个引物扩增片段区间范围是根据本课题组利用荧光引物对 96 个自交系和 1300 个杂交种的 DNA 指纹统计的结果确定的; 另外, 6 个引物仅给出了软件 Oligo 6.22 估测的该引物一个等位基因片段可能的大小. 通过利用 15 个自交系和 15 个杂交种在常规聚丙烯酰胺凝胶电泳检测结果显示, 所有这些组合均没有出现扩增片段范围交叉的情况.

表 2 10 重 PCR 组合的引物名单及组合方式

组合 ^{a)}	BIN	引物名称 ^{b)}	扩增范围	引物序列(5' 3')	标记荧光
1-1	4.01	phi072k4	413~433	GCTCGTCTCCTCCAGGTCAGG; CGTTGCCCATACATCATGCCTC	VIC
1-1	5.03	umc1705w1	273~330	GGAGGTCGTCAGATGGAGTTTCG; CACGTACGGCAATGCAGACAAG	VIC
1-1	6.00	bnlg161k8	152~208	TCTCAGTCTCCTGCTTATTGCTTTCG; GATGGATGGAGCATGAGCTTGC	VIC
1-2	9.03	phi065k9	399~419	CGCCTTCAAGAATATCCTTGTGCC ; GGACCCAGACCAGGTTCCACC	NED
1-2	2.03	bnlg125k1	219~327	GGGTACGGTTTCGTTTCCTTTGG ; TGCATCTAACAGCATCCCTTGAGC	NED
1-2	10.04	umc2163k5	154	GATGCAAGCGGGAATCTGAATC; CGACGAAATTGCTGGGGTTC	NED
1-3	1.03	bnlg439w1	319~385	AGTTGACATCGCCATCTTGGTGAC; GAACAAGCCCTTAGCGGGTTGTC	FAM
1-3	7.02	bnlg1792k8	192~250	CCCCAAAATTCAGGTGCC ; CCTCGTCTCCTACCAGAATG	FAM
1-4	8.08	phi080k15	203~233	TGAACCAACCGATGCAACTTG; TTGATGGGCACGATCTCGTAGTC	PET
1-4	3.09	bnlg1754w3	145	GGACGTCGGTACTGGCAATGG; CCACCACGCTGTCGTAGTGCTC	PET
2-1	1.11	bnlg2331k1	377~432	TTCTTTTCCTCGTTAGGCAACAG; CCAAAGCTGCCAGTTCCTAGATGAG	PET
2-1	3.00	umc2105k3	286~330	GAAGGGCAATGAATAGGCCATGAG; ATGGACTCTGTGCGACTTGTACCG	PET
2-1	7.06	phi116k3	150~171	AACTCCTGCCGGGACTCCT; CGGCCATGGATGGGATACAAATAC	PET
2-2	2.08	bnlg1940k9	407	GGCTCGTTTAAGAACGGTTGATTGC; GCACTAGACGGCTGGCAATTGG	NED
2-2	5.07	bnlg2305k4	292	CCCCTTCTCAGCACCTTG; CGTCTGTCTCCGTCCGTGTG	NED
2-2	9.01	umc2084w2	188~211	ACTGATCGCGACGAGTTAATTCAAAC; TACCGAAGAACAACGTCATTTACAGC	NED
2-3	10.07	bnlg1450k2	288~376	GCACTGAAATCTCCCATCATGTACG; TACAGCTCTTCTGGCATCGTCCG	FAM
2-3	8.03	umc1741k7	142~183	GCGCTTGGCATCTCCATGTATATC; GACCATCATCTTCCCTCGTGC	FAM
2-4	4.06	bnlg2291k4	384	GCACACCCGTAGTAGCTGAGACTTG; CATAACCTTGCTCCCAAACCC	VIC
2-4	6.05	bnlg1702k1	265	GATCCGCATTGTCAAATGACCAC; AGGACACGCCATCGTCATCA	VIC

a) 包括两套 10 重组, 每套组合分为 4 组; b) 新设计引物的命名为“原引物名称 + 设计者代号 + 序号”

以第一套 10 重 PCR 组合的构建过程为例. 通过重新设计, 与原引物相比, 新引物在各项序列指标上均有了明显改善, 不同新引物之间具有了相似的序列特征, 便于采用相同的扩增条件进行扩增(表 3). 通过对新设计引物与原引物进行对比实验, 表明新引物扩增谱带均在预期的位置, 且扩增效果均有改善. 从引物间相互作用看, 软件 Oligo 6.22 评估显示引物间相互作用最高 ΔG 仅 11.4. 引物进行 10 重 PCR 扩增, 并在 ABI3730XL 进行毛细管电泳, 实验显示这套引物组合扩增正常, 引物间无明显相互作用(图 2).

3 讨论

3.1 SSR 引物位点的多态性

以往的研究表明, SSR 引物多态性高低与重复单位的重复次数及重复类型、基因组不同区域、开发引物的来源数据库、检测所用材料等有关.

Sharopova 等人^[1]在对玉米 SSR 引物开发的研究中, 发现重复单位的重复次数与引物多态性存在较强的正相关, 在开发的 1051 个引物中, 随着重复次

数的增加, 多态性水平显著提高. Masi 等人^[18]在对大豆 SSR 的研究中, 发现二碱基和三碱基重复的多态性比三碱基以上的重复类型多态性高, 重复单位长的比短的多态性高.

研究发现, 不同染色体区域引物开发密度存在明显差异. Sharopova 等人^[1]发现在玉米第 2, 3, 5, 7, 8 和 10 号染色体上, 着丝点区域比染色体末端区域标记密度大, 认为这种差异可能与不同染色体区域的重组率不同有关. Nagaki 等人^[19]通过对两个玉米着丝点 BAC 克隆的序列分析, 显示着丝点区主要由短的富含 C 的卫星序列和着丝点特异的反转座子序列组成. Macaulay 等人^[20]在对大麦 SSR 引物研究中, 发现 SSR 在基因组上的分布不均匀, 着丝点区域标记密度比其他区域高. 从本研究筛选的 500 个玉米高多态性引物的分布看, 着丝点附近区域入选引物数量一般比较多, 而在第 2 号染色体的长臂、第 4 号染色体的短臂、第 5 号染色体的短臂、第 9 号染色体的长臂和第 10 号染色体短臂的末端区域入选的引物数量少或缺, 在其他区域也存在不同程度的标记富集

表 3 新设计引物与原引物序列软件评估信息对比

引物名称	产物大小	TM 值(U; L; P) ^{a)}	GC 含量(U; L; P)	引发效率(U; P)	错误引发(U; P)	二聚体(U; L; UL)	发夹(U; L)	3'端 ΔG
bnlg439	201~253	79.2; 77.9; 82.2	50; 41.4; 41.5	492; 483	91; 131	4; 8; 5	3; 3	7.9; 8.5
bnlg439w1	319~385	75; 75.4; 85.6	50; 56.5; 47.9	469; 522	117; 99	3; 3; 4	3; 3	6.1; 6.7
bnlg125	323~422	66.6; 67.9; 83.7	47.6; 43.5; 43.7	377; 388	81; 146	2; 2; 2	0; 0	6.7; 6.6
bnlg125k1	219~327	74.9; 75.1; 83.9	52.2; 50.0; 45.7	465; 445	31; 64	3; 4; 3	3; 3	8.8; 8.2
bnlg1754	215	68.5; 68.6; 87.9	50.0; 55.0; 55.8	450; 401	86; 136	4; 6; 3	0; 3	8.8; 8.2
bnlg1754w3	145	75.4; 75.9; 88.0	61.9; 63.6; 56.6	447; 431	88; 58	6; 3; 3	0; 3	8.4; 8.2
phi072	142~162	78.6; 78; 78	50; 42.9; 34.5	533; 484	103; 113	6; 6; 4	3; 0	6.4; 9.7
phi072k4	413~433	74.3; 75.2; 83.7	66.7; 54.5; 42.9	432; 463	69; 90	3; 4; 3	3; 3	8.2; 9.4
umc1705	57~114	70.7; 76.4; 82.9	50; 50; 49.5	411; 480	193; 81	4; 6; 3	0; 3	6.4; 6.4
umc1705w1	273~330	73.9; 74.4; 87.4	59.1; 54.5; 52.6	425; 464	98; 92	2; 6; 3	0; 0	8.4; 6.7
bnlg161	129~185	70.6; 70.4; 82.7	41.7; 39.1; 45.1	379; 410	115; 82	2; 4; 4	0; 3	6.9; 6.3
bnlg161k8	152~208	75.2; 74.8; 83.3	48.0; 54.5; 45.4	462; 428	38; 111	4; 4; 5	3; 3	9.0; 8.5
bnlg1792	113~171	68.0; 68.4; 81.9	50.0; 45.0; 44.4	431; 444	120; 97	4; 2; 4	0; 0	6.4; 7.0
bnlg1792k8	192~250	73.1; 72.8; 84.4	57.9; 56.5; 47.2	471; 426	0; 65	4; 2; 4	0; 3	9.4; 6.9
phi080	243~273	75.9; 75.1; 87.9	54.5; 54.5; 58	464; 407	78; 69	4; 4; 3	3; 4	5.5; 6.3
phi080k15	203~233	74.9; 74.0; 89.4	52.4; 52.2; 59.4	458; 449	71; 71	4; 4; 3	0; 3	6.7; 5.5
phi065	132~153	70.8; 72; 83.6	50; 50; 47	379; 415	33; 76	4; 4; 3	3; 3	5.5; 6.7
phi065k9	399~419	75.2; 75.7; 88.4	50.0; 66.7; 54.2	497; 436	94; 88	4; 3; 3	4; 3	9.4; 9.4
umc2163	145	72.2; 72.1; 81.7	41.7; 41.7; 43.4	496; 462	86; 83	3; 5; 3	3; 5	6.7; 7.0
umc2163k5	154	73.6; 73.4; 82.5	50.0; 55.0; 44.8	495; 462	83; 63	4; 4; 4	3; 3	6.6; 7.9

a) U, 上游引物序列; L, 下游引物序列; P, 扩增产物; UL, 上下游引物序列

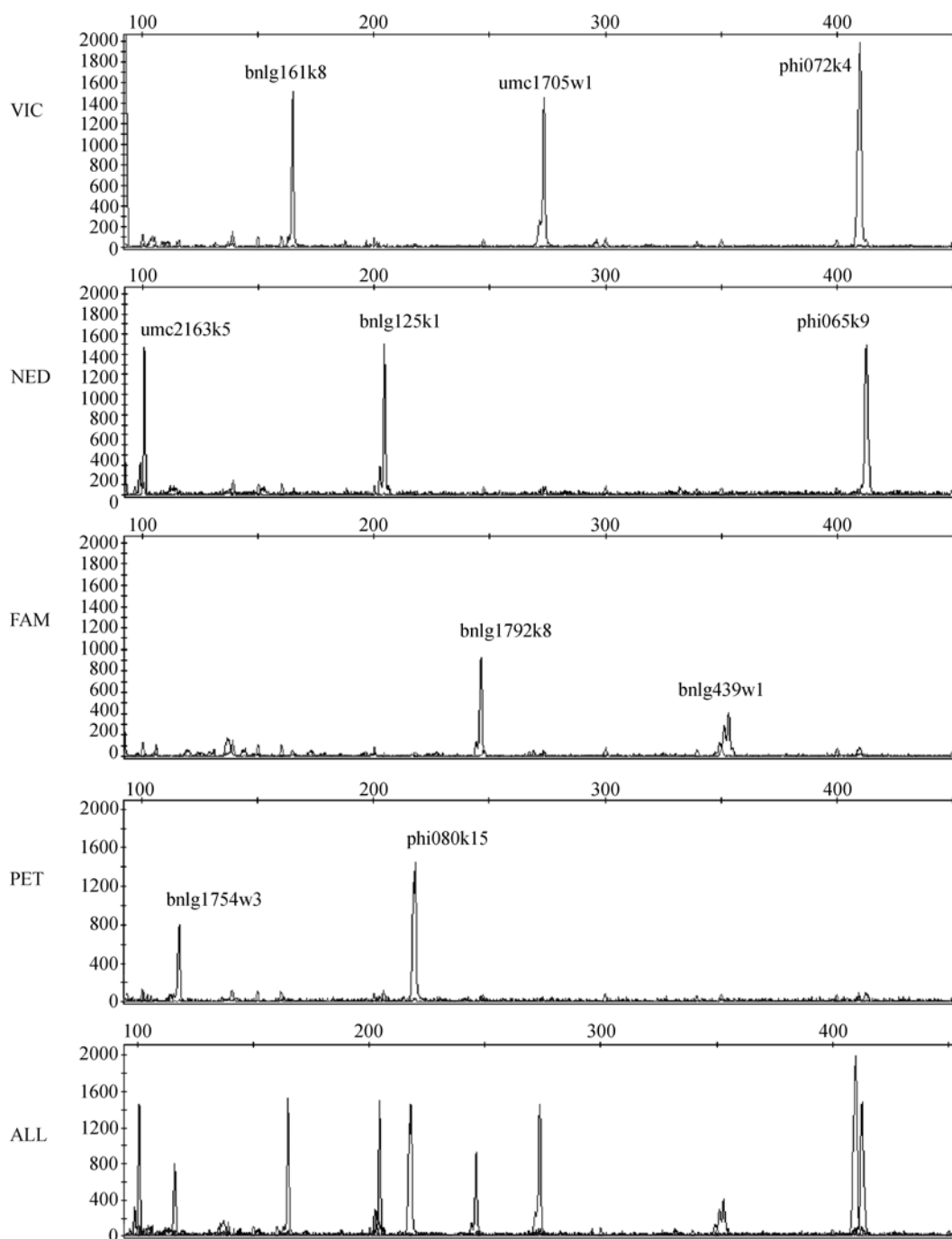


图2 第1套10重PCR组合引物在ABI3730XL遗传分析仪上的电泳效果

区和标记空缺区(图 1), 与前人的研究结果相吻合. SSR 引物的不完全均匀分布为选取一套全基因组上均匀分布的高多态性的通用引物带来了一定的障碍, 因此, 对某些低引物密度区域只能通过降低引物多态性入选标准, 才能尽量保证引物在染色体上的均

匀分布.

对不同作物的研究均表明, 从未知功能的EST序列及已知功能的基因内开发的引物的多态性水平总体上低于从全基因组开发的引物^[21,22], 表明非编码区的SSR引物多态性总体高于编码区的SSR引物. 这

可能是由于对特定等位基因组合的选择降低了育种资源库中等位基因的数目,进而降低了与之连锁的 SSR 的多态性。

此外,对引物多态性高低的评估受所用的材料类型及材料范围的影响较大^[12,23],许多引物尽管在栽培种间的多态性不高,当范围扩大到野生型或近源种时,多态性水平可能会提高。本研究确定的高多态性引物是对普通玉米品种评估获得的,对基于普通玉米的相关研究具有重要参考价值,当研究材料范围扩大后,在普通玉米上确定为低多态性或无多态性的引物仍可能具有应用价值,应根据研究的特殊需要对这部分引物进一步评估。

3.2 引物选择标准和多重 PCR 组合

许多研究者针对所研究的物种各自提出了 SSR 引物选择标准,如马铃薯^[22]、向日葵^[24]、大麦^[20]、大豆^[18]、小麦^[25]、玉米^[26]、水稻^[27]、葡萄^[28]、油菜^[29]、西红柿^[30]等,综合这些选择标准,基本包括了以下 4 点:多态性高、扩增质量高、在全基因组上均匀分布、复合扩增的潜力。此外,在玉米^[5-7]、向日葵^[31]、棉花^[32]、水稻^[33]、大豆^[34]等一些作物上,也探索了多重 PCR 组合的研究。然而,受限于已开发引物的数量和质量,获得特定物种的一套完全符合上述选择标准的高质量的核心 SSR 引物并建立其复合扩增体系,其难度仍然是比较大的。

由于玉米数据库上已公布的 SSR 引物是由不同的研究单位设计的,设计者采用的设计标准不完全相同,因此,设计的质量存在差异,适宜的扩增条件各不相同,基于这些引物进行多重 PCR 组合时限制较大:引物的 TM 值差别较大,因此最适退火温度不同;引物扩增产物大小范围比较集中,较难选择到合适的多重组合方式;即使已满足了上述两个条件,部分引物组合还会出现引物间较强的相互作用。从玉米上已进行的多重 PCR 组合研究看^[5-7],这些研究主要考虑了引物扩增片段范围及引物间相互作用,而对所用的 SSR 引物没有进行系统的多态性的筛选,部分引物多态性较低,且没有完全考虑引物在玉米染色体上的均匀分布,因此,获得的多重 PCR 引物组合一般仅适用于特定研究目的,且一般只能组 2~4 重,尚未建立起稳定的 5 重以上的 PCR 组合。

本研究为了获得一套玉米上通用的 SSR 核心引物组合,采用了分步筛选,逐步解决策略:首先对全基因组上所有已开发的 1900 个 SSR 引物位点进行

多态性筛选,获得全部 500 个高多态性玉米 SSR 引物;然后从高多态引物中按照均匀覆盖全基因组的原则进一步选择 100 个引物作为一套通用核心引物;再对这套核心引物进行重新设计,一方面改善其扩增质量,一方面使其适于 10 重 PCR 复合扩增的要求,且每套 10 重 PCR 引物组合均是从玉米每条染色体上选取 1 个引物构成的,最终期望建立起 10 套 10 重 PCR 扩增引物组合。这一研究策略与以往的研究相比,最大的不同在于:该策略将多态性的筛选与扩增质量的筛选分开进行;将引物的设计与多重 PCR 组合结合进行。因为多态性的高低是扩增的微卫星位点固有的特征,只有通过广泛筛选才能将玉米全基因组中多态性高的位点挑选出来,与引物设计的好坏无关;而引物扩增质量则与引物设计的好坏直接相关,通过引物重新设计可以得到改善。

本研究建立的 10 重 PCR 组合是在全基因组均匀分布的高多态性的核心引物的基础上进行设计的,所用的引物具有广泛的通用性,在基于 ABI3730XL 毛细管电泳五色荧光检测系统上可以实现 10 重 PCR 扩增和 10 重电泳检测,具有较高的检测通量。因此,本研究获得的多重 PCR 引物组合将在较大范围内具有应用价值。

致谢 中国科学院遗传与发育研究所王斌研究员对本论文进行了审阅和指导,在此表示感谢。本工作为北京农业育种基础研究创新平台和北京市自然科学基金(批准号:YZPT02-06)资助项目。

参 考 文 献

- 1 Chin E C L, Senior M L, Shu H, et al. Maize simple sequence repetitive DNA sequences: Abundance and allele variation. *Genome*, 1996, 39: 866—873
- 2 Kantety R V, Rota M L, Matthews D E, et al. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Mol Biol*, 2002, 48: 501—510[DOI]
- 3 Sharopova N, McMullen M D, Schultz L, et al. Development and mapping of SSR markers for maize. *Plant Mol Biol*, 2002, 48: 463—481[DOI]
- 4 Parida S K, Kumar K A R, Dalal V, et al. Unigene derived microsatellite markers for the cereal genomes. *Theor Appl Genet*, 2006, 112: 808—817[DOI]
- 5 Gethy J G, Labate J A, Lamkey K R, et al. SSR variation in important US maize inbred lines. *Crop Sci*, 2002, 42 (3): 951—957
- 6 王凤格, 赵久然, 郭景伦, 等. 中国玉米新品种 DNA 指纹库建立系列研究. 多重 PCR 技术在玉米 SSR 引物扩增中的应用. *玉米科学*, 2003, 11(4): 3—6
- 7 Clerc V L, Bazante F, Baril C, et al. Assessing temporal changes in genetic diversity of maize varieties using microsatellite markers.

- Theor Appl Genet, 2005, 110: 294—302[DOI]
- 8 廖琴, 孙世贤. 中国玉米新品种动态 2004 年国家级玉米品种区试报告. 北京: 中国农业科学技术出版社, 2004
- 9 Xia X C, Reif J C, Melchinger A E, et al. Genetic diversity among CIMMYT maize inbred lines investigated with SSR markers. Lowland tropical maize. Crop Sci, 2004, 44(6): 2230—2237
- 10 Marilyn L, Warburton M L, Xia X C, et al. Genetic characterization of CIMMYT inbred maize lines and open pollinated populations using large scale fingerprinting methods. Crop Sci, 2002, 42(6): 1832—1840
- 11 Smith J S C, Chin E C L, Shu H, et al. A evaluation of the utility of SSR loci as molecular markers in maize(*Zea mays* L): Comparisons with data from RFLPs and pedigree. Theor Appl Genet, 1997, 95: 163—173[DOI]
- 12 Matsuoka Y, Mitchell S E, Kresovich S, et al. Microsatellites in *Zea*-variability, patterns of mutations, and use for evolutionary studies. Theor Appl Genet, 2002, 104: 436—450[DOI]
- 13 Lu H, Bernardo R. Molecular marker diversity among current and historical maize inbreds. Theor Appl Genet, 2001, 103: 613—617[DOI]
- 14 李晓辉, 李新海, 李文华, 等. SSR 标记技术在玉米杂交种种子纯度测定中的应用. 作物学报, 2003, 29(1): 63—68
- 15 滕文涛. 利用分子标记划分玉米杂种优势群及定位农艺性状 QTL. 博士学位论文. 北京: 中国农业大学, 2002
- 16 Enoki H, Sato H, Koinuma K. SSR analysis of genetic diversity among maize inbred lines adapted to cold regions of Japan. Theor Appl Genet, 2002, 104: 1270—1277[DOI]
- 17 Bantte K, Prasanna B M. Simple sequence repeat polymorphism in Quality Protein Maize (QPM) lines. Euphytica, 2003, 129: 337—344[DOI]
- 18 Masi P, Zeuli P L S, Donini P. Development and analysis of multiplex microsatellite markers sets in common bean (*Phaseolus vulgaris* L). Mol Breed, 2003, 11: 303—313[DOI]
- 19 Nagaki K, Song J, Stupar R M, et al. Molecular and cytological analyses of large tracks of centromeric DNA reveal the structure and evolutionary dynamics of maize centromeres. Genetics, 2003, 163: 759—770
- 20 Macaulay M, Ramsay L, Powell W, et al. A representative, highly informative 'genotyping set' of barley SSRs. Theor Appl Genet, 2001, 102: 801—809[DOI]
- 21 Leigh F, Lea V, Law J, et al. Assessment of EST- and genomic microsatellite markers for variety discrimination and genetic diversity studies in wheat. Euphytica, 2003, 133: 359—366[DOI]
- 22 Ghislain M, Spooner D M, Rodríguez F, et al. Selection of highly informative and user-friendly microsatellites (SSRs) for genotyping of cultivated potato. Theor Appl Genet, 2004, 108: 881—890[DOI]
- 23 Lübberstedt T, Dussle C, Melchinger A E. Application of microsatellites from maize to teosinte and other relatives of maize. Plant Breed, 1998, 117: 447—450[DOI]
- 24 Zhang L S, Clerc V L, Li S, et al. Establishment of an effective set of simple sequence repeat markers for sunflower variety identification and diversity assessment. Can J Bot, 2005, 83(1): 66—72[DOI]
- 25 Perry D J. Identification of Canadian durum wheat varieties using a single PCR. Theor Appl Genet, 2004, 109: 55—61[DOI]
- 26 George M L C, Regalado E, Li W, et al. Molecular characterization of Asian maize inbred lines by multiple laboratories. Theor Appl Genet, 2004, 109: 80—91[DOI]
- 27 Singh R K, Sharma R K, Singh A K, et al. Suitability of mapped sequence tagged microsatellite site markers for establishing distinctness, uniformity and stability in aromatic rice. Euphytica, 2004, 135: 135—143[DOI]
- 28 This P, Jung A, Boccacci P, et al. Development of a standard set of microsatellite reference alleles for identification of grape cultivars. Theor Appl Genet, 2004, 109: 1448—1458[DOI]
- 29 Tommasini L, Batley J, Arnold G M, et al. The development of multiplex simple sequence repeat (SSR) markers to complement distinctness, uniformity and stability testing of rape (*Brassica napus* L.) varieties. Theor Appl Genet, 2003, 106: 1091—1101
- 30 Bredemeijer G M M, Arens P, Wouters D, et al. The use of semi-automated fluorescent microsatellite analysis for tomato cultivar identification. Theor Appl Genet, 1998, 97: 584—590[DOI]
- 31 Tang S X, Kishore V K, Knapp S J. PCR-multiplexes for a genome-wide framework of simple sequence repeat marker loci in cultivated sunflower. Theor Appl Genet, 2003, 107: 6—19
- 32 Liu S, Saha S, Stelly D, et al. Chromosomal assignment of microsatellite loci in cotton. J Hered, 2000, 91: 326—332[DOI]
- 33 Coburn J R, Temnykh S V, Paul E M, et al. Design and application of microsatellite marker panels for semiautomated genotyping of rice (*Oryza sativa* L). Crop Sci, 2002, 42(6): 2092—2099
- 34 Narvel J M, Chu W C, Fehr W R, et al. Development of multiplex sets of simple sequence repeat DNA markers covering the soybean genome. Mol Breed, 2000, 6: 175—183

(2006-07-12 收稿, 2006-11-11 接受)