

利用电子 PCR 分析草菇基因组 SSR 标记多态性

熊登坤^{1,2} 鲍大鹏^{2*} 边银丙^{1*}

(1. 华中农业大学 应用真菌研究所 湖北 武汉 430070)

(2. 国家食用菌工程技术研究中心 农业部食用菌南方资源利用重点实验室 上海市农业遗传育种重点开放实验室 上海市农业科学院食用菌研究所 上海 201403)

摘要:【目的】利用电子 PCR 分析草菇基因组 SSR 标记多态性, 并通过 PCR 验证分析结果的可靠性。【方法】利用 MISA 程序定位草菇基因组 SSR 位点并结合 Primer3.0 程序设计 SSR 分子标记引物, 运用电子 PCR 进行 SSR 分子标记多态性分析, 基于分析结果进行 PCR 验证。

【结果】随机选取 658 对 SSR 引物在草菇同核体菌株 V23-1 和 PD19 中进行真实 PCR 检测, 结果表明 28.6% 的 SSR 引物具有多态性。数据分析表明, 如果 SSR 标记来源于电子 PCR 产物长度没有差异的类型, 仅 4.8% 的 SSR 引物在真实 PCR 中表现出多态性; 如果 SSR 标记来源于电子 PCR 产物长度差异大于或者等于 3 bp 的类型, 其中至少 48.3% 的 SSR 引物在真实 PCR 中表现出多态性。【结论】PCR 验证结果表明利用电子 PCR 可以提高 SSR 多态性引物的筛选效率。

关键词: 草菇, SSR, 多态性, 电子 PCR

Polymorphism analysis of genomic SSR markers of *Volvariella volvacea* by electronic PCR

XIONG Deng-Kun^{1,2} BAO Da-Peng^{2*} BIAN Yin-Bing^{1*}

(1. The Institute of Applied Mycology, Huazhong Agricultural University, Wuhan, Hubei 430070, China)

(2. National Engineering Research Center of Edible Fungi, Ministry of Science and Technology, Key Laboratory of Edible Fungi Resources and Utilization (South), Shanghai Key Laboratory of Agricultural Genetics and Breeding, Institute of Edible Fungi, Shanghai Academy of Agriculture Science, Shanghai 201403, China)

Abstract: [Objective] To verify the polymorphism of SSR markers analyzed with electronic PCR in *Volvariella volvacea* genome by using real PCR. [Methods] SSR loci in *V. volvacea* genome were identified with MISA software and the primers of the SSR molecular markers were developed with Primer3.0 software. The polymorphism of SSR markers were analyzed with electronic PCR and then verified by PCR. [Results] The 658 pairs of randomly selected SSR primers were checked by real PCR in 2 homokaryon strains, V23-1 and PD19, and the results show that 28.6% of SSR primers have the polymorphism. When the tested SSR loci belong to the type that had no length difference analyzed by the electronic PCR, only 4.8% SSR primers had the polymorphism. When the tested SSR loci selected from the group in which the length difference among the electronic PCR products of different genome was more than or equal to 3 bp, the true polymorphous SSR primers reached

*通讯作者: ✉: 鲍大鹏: baodp@hotmail.com; 边银丙: bianyb.123@163.com

收稿日期: 2013-12-21; 接受日期: 2014-03-12; 优先数字出版日期(www.cnki.net): 2014-03-24

48.3%. **[Conclusion]** The real PCR confirmed that it is more efficient to select polymorphous SSR markers with electronic PCR.

Keywords: *Volvariella volvacea*, SSR, Polymorphism, Electronic PCR

简单重复序列(Simple sequence repeat, SSR)是一类一般由 1–6 个碱基的单元经串联重复而成的 DNA 序列^[1], 广泛存在于动植物以及真菌的基因组中^[2]。由于 SSR 标记具有需要 DNA 的量少, 对 DNA 质量要求不高, 易于 PCR 检测, 共显性遗传, 多态性高及在基因组中分布广泛的特点^[3], 广泛应用于种质资源鉴定^[4]、遗传多样性分析^[5]、基因定位^[6]、遗传图谱构建^[7]及 QTL 分析^[8]等众多方面。

传统的 SSR 分子标记开发需构建基因文库、杂交筛选和序列鉴定, 步骤繁琐, 具有费时、昂贵和劳动强度大的缺点^[9]。大规模测序计划的开展和物种基因组数据库的建立, 为 SSR 位点的获取提供了便捷有效的途径。利用 NCBI 等公共数据库中序列信息, 结合生物信息学分析手段可以快捷地开发 SSR 标记, 目前有多个物种的 SSR 标记已经得到开发^[10–14]。利用公共数据库中的序列信息, 可以运用生物信息学方法快速定位到大量 SSR 位点, 但是这些 SSR 位点的准确性、特异性和多态性还缺少比较全面和细致的量化研究。

电子 PCR (Electronic PCR, e-PCR)是一种常用的核苷酸序列电子分析工具, 在 DNA 片段的染色体定位、基因组辅助作图、PCR 引物辅助设计和基因克隆等方面发挥着重要作用^[15]。将电子 PCR 与 SSR 分子标记的开发结合能够充分利用基因组数据, 提高 SSR 分子标记开发效率。本研究利用草菇基因组序列信息, 定位出大量 SSR 位点, 并开发出相应的 SSR 分子标记引物。然后通过电子 PCR 方法分析 SSR 引物的扩增产物的长度、特异性以及多态性, 在此基础上, 随机选取了部分 SSR 位点进行真实 PCR 实验验证, 以明确电子 PCR 方法对 SSR 引物分析的准确性, 所获得的研究结果可以为提高 SSR 分子标记的开发效率提供科学指导。

1 材料与方法

1.1 供试材料和基因组信息

本研究用于 SSR 位点定位和电子 PCR 研究的草菇 V23-1 和 PYd21 的基因组序列来源于 NCBI 数据库 (<http://www.ncbi.nlm.nih.gov/nuccore/AMXZ00000000> 和 <http://www.ncbi.nlm.nih.gov/nuccore/ANCH00000000>)。

本研究对电子 PCR 分析结果进行 PCR 验证所用的草菇单核体菌株 V23-1 和 PD19 均由上海市农业科学院食用菌研究所菌种保藏中心保存和提供。菌株 PYd21 为草菇菌株屏优一号的原生质体单核体, PD19 为草菇菌株屏优一号的孢子单核体。

1.2 方法

1.2.1 草菇基因组 SSR 标记的开发: 以草菇基因组序列为模板, 用 MISA 程序 (<http://pgrc.ipk-gatersleben.de/misa/misa.html>)进行分析, 设定 SSR 搜索条件: 单、二、三、四、五和六碱基重复单元的最小重复数分别为 10、6、5、5、5、5。

1.2.2 SSR 标记引物的设计: 根据 SSR 位点两端的保守区域, 利用 Primer3.0 (<http://fokker.wi.mit.edu/primer3/>)设计引物, 每个 SSR 设计 3 对引物。设计引物时设置的主要参数为: GC 含量 40%–60%, 退火温度约为 55–65 °C, 引物长度为 18–22 bp, 预期扩增产物长度为 100–300 bp。

1.2.3 e-PCR 分析: 利用 e-PCR 程序 (<ftp://ftp.ncbi.nlm.nih.gov/pub/schuler/e-PCR/>)中的 re-PCR 命令对 SSR 标记的引物在草菇基因组序列中进行匹配, 结合自开发的 python 程序分析引物对电子 PCR 的扩增产物长度、特异性和多态性进行检测。

1.2.4 多态性 SSR 标记的验证: 运用 CTAB 方法^[16]提取单核体菌株 V23-1 和 PD19 的基因组 DNA 作为模板, 随机选取 SSR 标记引物进行真实

PCR 扩增, 扩增产物在聚丙烯酰胺凝胶中进行电泳分离, 用于分析相应引物的真实 PCR 与电子 PCR 结果的一致性。

2 结果与分析

2.1 SSR 位点在草菇基因组中的分布及特点

在 V23-1 菌株基因组中定位到 1 346 个 SSR 位点, 在 PYd21 基因组中定位到 1 409 个 SSR 位点, 平均每 26 kb 序列中存在 1 个 SSR 位点, 草菇基因组中 SSR 位点分布较为广泛。结合两个草菇基因组中 SSR 位点的特点来看, SSR 位点构成以单碱基重复为主(分别占 57.4%和 60.6%), 其次是三碱基重复(分别占 29.7%和 26.6%)和二碱基重复(分别占 8.2%和 8.5%), 四碱基及以上重复序列所占比例小(表 1)。

草菇 SSR 位点中碱基构成情况如图 1 所示, 在单碱基重复 SSR 位点中, 以多聚 A/T 为主; 在二碱基重复 SSR 中, 所占比例较大的是 AG/CT、AT/AT 重复, 而 CG/CG 重复所占比例较小; 三碱基重复 SSR 的主要基序是 ACC/GGT、ACG/CGT、AGC/GCT。二碱基及以上重复类型 SSR 基序的重复次数主要在 10 次以内, 与其它真菌基因组中

SSR 基序的重复规律相一致^[11]。

2.2 SSR 标记引物的设计及电子 PCR 分析

利用 Primer3.0 对 V23-1 基因组序列中定位到的 SSR 位点进行引物设计, 共得到 4 020 对引物。将这些引物分别在 V23-1 和 PYd21 基因组序列中用电子 PCR 程序中的 re-PCR 命令进行匹配, 结合自开发的 python 程序进行分析, 结果发现, 4 020 对 SSR 引物在 V23-1 基因组中生成特异性产物的概率高达 90.4% (3 633 对 SSR 引物) 这可能与 PCR 引物序列来源于 V23-1 菌株基因组有关, 但仍有 8.5% 的引物(342 对)生成非特异扩增产物, 无法生成扩增产物的概率为 1.1% (45 对); 与之比较 4 020 对 SSR 引物在 PYd21 基因组中, 可以生成特异性产物的概率只有 65.5% (2 632 对), 生成非特异扩增产物的引物占 7.9% (317 对), 没有生成任何扩增产物的 SSR 引物占 26.6% (1 071 对)。进一步对 PYd21 基因组中没有生成 PCR 产物的 1 071 对 SSR 引物进行分析, 发现有 120 对引物的正向引物和反向引物分别匹配到不同 contig 上, 这可能是由于测序数据不完整造成的; 另外有 951 对引物中的一条引物或两条引物没有匹配到序列, 这可能是由于引物序列的变异造成的。

表 1 不同重复基序类型 SSR 在草菇基因组中出现的频率

Table 1 The frequency of different motif type

重复基序类型 Motif type	V23-1 基因组 Genome V23-1		PYd21 基因组 Genome PYd21	
	数目 Count	比例 Rate (%)	数目 Count	频率 Rate (%)
Mononucleotide	773	57.4	854	60.6
Dinucleotide	110	8.2	120	8.5
Trinucleotide	400	29.7	375	26.6
Tetranucleotide	19	1.4	15	1.1
Pentanucleotide	14	1.0	16	1.1
Hexanucleotide	30	2.2	29	2.1
Total	1 346	100.0	1 409	100.0

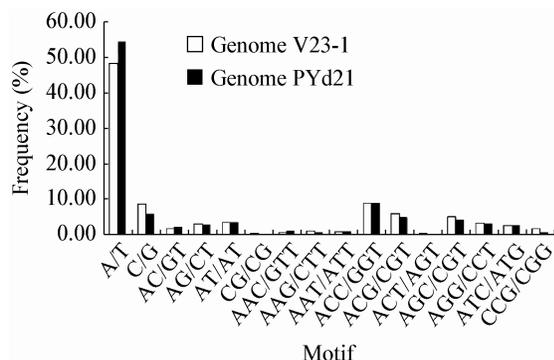


图1 草菇基因组 SSR 主要重复基序的分布频率
Figure 1 The frequency distribution of main repeat motif in two genome

经电子 PCR 分析, 2 617 对 SSR 引物在草菇的两套基因组中均能够匹配获得特异性产物, 其中有 1 368 对 SSR 引物经匹配获得的扩增产物存在长度差异, 在分析的两套基因组间具有多态性, 对应于 519 个 SSR 位点。进一步对引物电子 PCR 生成的多态性扩增产物进行分析, 在 1 368 对引物中, 长度差异值小于 3 bp 的引物有 807 对, 占总数的 59.0%, 长度差异值小于 30 bp 的引物有 1 242 对, 占总数的 90.8% (图 2)。另外 1 249 对引物预测生成的条带在 V23-1 和 PYd21 基因组之间没有多态性。

2.3 SSR 引物多态性检测结果

根据电子 PCR 分析结果, 共选取 658 对 SSR 引物用于真实 PCR 检测, 其中从 1 249 对预测无

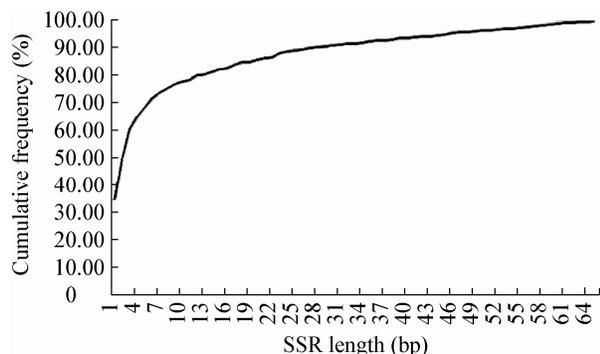


图2 SSR 在草菇基因组中的长度差异分布
Figure 2 The cumulative frequency of SSR length diversity in different genome

多态性的 SSR 引物中随机选取 146 对, 从 1 368 对预测有多态性的 SSR 引物中随机选取 281 对, 其它随机选取的 231 对引物为电子 PCR 扩增产物长度未知。然后以单核体菌株 V23-1 和 PD19 的基因组 DNA 为模板进行真实 PCR, 扩增产物通过聚丙烯酰胺凝胶电泳进行检测(图 3)。结果显示, 这 658 对 SSR 引物均能够在真实 PCR 中扩增出特异性条带, 这说明电子 PCR 在预测 SSR 位点的特异性上具有较高的可靠性, 其中共有 188 对引物(28.6%)扩增产物具有多态性。

进一步对不同来源的 SSR 引物具有的真实多态性进行分析, 结果表明, 146 对电子 PCR 扩增产物无长度差异性的引物中有 7 对在真实 PCR 检测中表现出多态性, 其余 139 对(95.2%)和电子 PCR 分析结果一致, 没有多态性 PCR 产物产生, 这表明电子 PCR 分析给出的没有多态性的预测结果具有很高的可信度。在 281 对电子 PCR 扩增产物有长度差异性的引物中有 112 对可以通过真实 PCR 检测到多态性, 平均多态性概率为 39.9%。进一步对电子 PCR 分析显示的不同长度差异值的引物进行细化分析, 当预测差异值大于或者等于 3 bp 时,

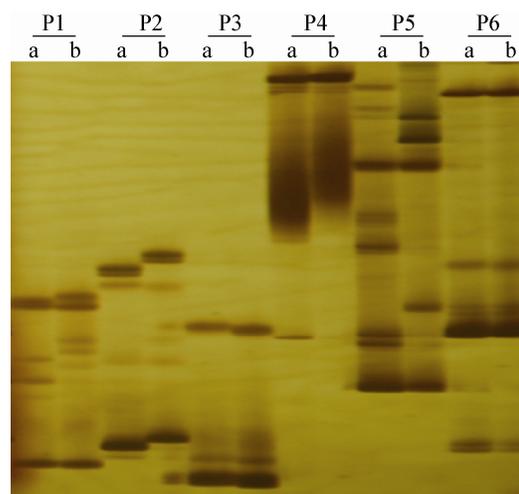


图3 部分 SSR 引物的 PCR 扩增产物
Figure 3 The PCR product of partial SSR primers
注: a: 菌株 PD19; b: 菌株 V23-1.
Note: a: The strain PD19; b: The strain V23-1.

真实 PCR 检测到的多态性概率均高于 39.9% 的平均值, 这提示在运用电子 PCR 开发多态性 SSR 引物时, 从长度差异值大于或者等于 3 bp 的 SSR 位点中选择候选引物, 将能够获得较高的成功率。另外在 231 对电子 PCR 扩增产物长度未知的引物中有 69 对(29.9%) 在真实 PCR 检测中表现出多态性, 与随机选取引物进行筛选获得多态性引物的概率(28.6%)相当(表 2)。

2.4 SSR 引物多态性与 SSR 位点长度的关系

由于在电子 PCR 预测中表现出多态性的引物是开发多态性引物的主要候选引物, 所以本研究进一步对随机选取的 281 对具有多态性的 SSR 引物进行详细分析, 首先对 SSR 位点的长度与真实 PCR 多态性比例之间的关系进行了统计分析, 结果表明(表 3), SSR 位点的长度越长, 真实 PCR 表现出多态性的可能性越高。当 SSR 位点的长度小于 15 bp

时, 从中筛选得到多态性引物的概率为 29.5%, 低于平均多态性产生概率(39.9%)。当 SSR 的长度大于或者等于 15 bp 时, 获得多态性引物概率(42.2%–84.0%)均高于 39.9% 的平均值, 其中当 SSR 位点长度大于 24 bp 时, 根据其位点设计的引物可具有高达 84.0% 的多态性。

2.5 SSR 引物多态性与 SSR 位点基序类型的关系

在 281 对引物电子 PCR 扩增产物有长度差异性的 SSR 中, 单碱基及三碱基重复类型 SSR 引物有 257 对, 占总数的 91.5%。47 对三碱基重复类型 SSR 引物中经 PCR 检测具有多态性的引物有 21 对(44.7%), 不具有多态性的引物有 26 对(55.3%)。210 对单碱基重复类型 SSR 引物中经 PCR 检测具有多态性的引物有 75 对(35.7%), 不具有多态性的有 135 对(64.3%)。可见在草菇基因组占大部分的

表 2 SSR 引物多态性检测结果

Table 2 Verification for the polymorphism of SSR primers

电子 PCR 分析显示的长度差异值 The length difference displayed by e-PCR (bp)	引物总数 Count of primers	有多态性 With polymorphism	比例 Rate (%)	无多态性 Without polymorphism	比例 Rate (%)
0	146	7	4.8	139	95.2
1	74	6	8.1	68	91.9
2	37	8	21.6	29	78.4
3	29	14	48.3	15	51.7
4	12	6	50.0	6	50.0
5	18	11	61.1	7	38.9
6	14	9	64.3	5	35.7
>6	97	58	59.8	39	40.2
Unknown	231	69	29.9	162	70.1
Total	658	188	28.6	470	71.4

表 3 SSR 多态性与长度间的关系

Table 3 The correlation between SSR polymorphism and length

SSR 长度 Length of SSR (bp)	引物总数 Count of primers	有多态性 With polymorphism	比例 Rate (%)	无多态性 Without polymorphism	比例 Rate (%)
10–14	166	49	29.5	117	70.5
15–19	64	27	42.2	37	57.8
20–24	26	15	57.7	11	42.3
>24	25	21	84.0	4	16.0
Total	281	112	39.9	169	60.1

单碱基重复类型 SSR 的多态性不如三碱基重复类型 SSR。进一步分析单碱基重复类型 SSR 的多态性与重复基元的相关性, 发现 168 对 A/T 重复型 SSR 标记引物中检测到多态性的引物有 56 对 (33.3%), 没有检测到多态性的引物有 112 对 (66.7%); 42 对 C/G 重复型 SSR 标记引物中检测到多态性的引物 19 对 (45.2%), 没有检测到多态性的引物 23 对 (54.8%)。相对于 A/T 重复型 SSR, 更适合从 C/G 重复型 SSR 中筛选多态性 SSR 标记。

3 讨论

海量的基因组数据为定位 SSR 位点提供了便利, 生物信息学的分析方法为开发 SSR 分子标记提供了有效的工具。本研究以草菇基因组为研究对象, 首次运用电子 PCR 和真实 PCR 相结合的方法, 细致地分析了各种类型 SSR 位点在两个单核体菌株之间的特异性和多态性。结果显示从电子 PCR 产物长度差异超过 3 bp 的 SSR 引物中筛选得到具有多态性 SSR 标记的效率相对较高; 而且 SSR 长度越长, 筛选得到具有多态性 SSR 标记的可能性越高。这些研究结果为高效利用数据库信息提供了指导。

SSR 标记在菌株 V23-1 和 PD19 间进行 PCR 扩增检测出的多态性与其在基因组 V23-1 和 PYd21 中经电子 PCR 分析获得的多态性结果并不完全吻合, 这种差异可能是菌株 PD19 和 PYd21 间的差异造成的。同核体菌株 PD19 是由菌株屏优一号经单孢分离得到的, 而同核体菌株 PYd21 是由菌株屏优一号经原生质体单核化获得, 两者虽均来源于亲本屏优一号, 但其基因组序列必然存在一定差异。总体来看, 检测结果符合分析结果, 恰好可以证明电子 PCR 分析结果在不同菌株中的普适性。研究结果表明, 结合电子 PCR 分析可以极大地提高草菇基因组中多态性 SSR 标记的筛选效率。

参考文献

- [1] Tautz D. Hypervariability of simple sequences as a general source for polymorphic DNA markers[J]. *Nucleic Acids Research*, 1989, 17(16): 6463-6471.
- [2] Bhargava A, Fuentes FF. Mutational dynamics of microsatellites[J]. *Molecular Biotechnology*, 2010, 44(3): 250-266.
- [3] Gupta PK, Varshney RK. The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat[J]. *Euphytica*, 2000, 113(3): 163-185.
- [4] Donges K, Schlobinski D, Cremer E, et al. Six newly developed microsatellite markers of *Laccaria amethystina*, using an improved CSSR approach[J]. *Mycological Progress*, 2008, 7(4): 285-290.
- [5] Xiao Y, Liu W, Dai Y, et al. Using SSR markers to evaluate the genetic diversity of *Lentinula edodes*' natural germplasm in China[J]. *World Journal of Microbiology and Biotechnology*, 2010, 26(3): 527-536.
- [6] Liu C, Wang J, Huang T, et al. A missense mutation in the VHYNP motif of a DELLA protein causes a semi-dwarf mutant phenotype in *Brassica napus*[J]. *Theoretical and Applied Genetics*, 2010, 121(2): 249-258.
- [7] Labbé J, Zhang X, Yin T, et al. A genetic linkage map for the ectomycorrhizal fungus *Laccaria bicolor* and its alignment to the whole-genome sequence assemblies[J]. *New Phytologist*, 2008, 180(2): 316-328.
- [8] Foulongne-Oriol M, Rodier A, Savoie JM. Relationship between yield components and partial resistance to *Lecanicillium fungicola* in the button mushroom, *Agaricus bisporus*, assessed by quantitative trait locus mapping[J]. *Applied and Environmental Microbiology*, 2012, 78(7): 2435-2442.
- [9] 张增翠, 侯喜林. SSR 分子标记开发策略及评价[J]. *遗传*, 2004, 26(5): 763-768.
- [10] Cheng X, Xu J, Xia S, et al. Development and genetic mapping of microsatellite markers from genome survey sequences in *Brassica napus*[J]. *Theoretical and Applied Genetics*, 2009, 118(6): 1121-1131.
- [11] 刘林, 李成云, 杨静, 等. 灰盖鬼伞基因组中微卫星序列的组成[J]. *西南农业学报*, 2006, 19(1): 131-135.
- [12] 李明卫. 大型真菌黑松露 EST-SSR 信息分析[D]. 杭州: 浙江农林大学硕士学位论文, 2012.
- [13] Murat C, Riccioni C, Belfiori B, et al. Distribution and localization of microsatellites in the Perigord black truffle genome and identification of new molecular markers[J]. *Fungal Genetics and Biology*, 2011, 48(6): 592-601.
- [14] Vincenot L, Nara K, Sthultz C, et al. Extensive gene flow over Europe and possible speciation over Eurasia in the ectomycorrhizal basidiomycete *Laccaria amethystina* complex[J]. *Molecular Ecology*, 2012, 21(2): 281-299.
- [15] Schuler GD. Electronic PCR: bridging the gap between genome mapping and genome sequencing[J]. *Trends in Biotechnology*, 1998, 16: 456-459.
- [16] 张红, 秦莲花, 谭琦, 等. 用改进的 CTAB 法提取香菇基因组 DNA[J]. *上海大学学报*, 2006, 12(5): 547-550.