

大肠杆菌 MG1655 菌株 ERIC-PCR 图谱主带序列组成分析*

陈迎春^{1,2} 曹又方² 赵立平^{1**}

(上海交通大学生命科学技术学院生物工程系 上海 200240)¹

(山西大学生物工程实验室 太原 030006)²

摘要: ERIC-PCR 已经在细菌分类、鉴定及混合菌群分析中得到广泛应用,但对其产物形成规律的认识仍存在分歧。以大肠杆菌 MG1655 为对象,对其 ERIC-PCR 指纹图谱中 1.1kb 主带中的 DNA 片段进行了克隆、测序、基因组定位以及引物匹配分析。结果表明,这条 1.1kb 主带由分布在基因组中不同位置的 3 种序列不同的片段组成,各片段的丰度差异较大,最高为 97.89%; 3 种片段中的 2 种所在的基因组区域仅一端含有 ERIC 序列。推测对含有 ERIC 序列的基因组 DNA 进行扩增时,ERIC-PCR 是一种非随机扩增。

关键词: ERIC 序列, ERIC-PCR, *E. coli* K-12 MG1655 菌株

中图分类号: Q932 **文献标识码:** A **文章编号:** 0253-2654 (2002) 06-0028-06

* 国家高技术研究发展计划项目 (“863” 项目) (No. SZ-03-01-04)

Project of Chinese National Programs for High Technology Research and Development (No. SZ-03-01-04)

* * 联系人 E-mail: IPzhao@sjtu. edu. cn

收稿日期: 2001-09-14, 修回日期: 2001-11-16

NON-RANDOM NATURE OF GENOMIC DNA AMPLIFICATION OF *E. COLI* K-12 MG1655 VIA ERIC-PCR

CHEN Ying-Chun^{1,2}, CAO You-Fang², ZHAO Li-Ping¹

(*School of Life Science and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240*)¹

(*Shanxi Key Lab of Biotechnology, Shanxi University, Taiyuan 030006*)²

Abstract: ERIC-PCR has been widely used as genomic fingerprinting technology for classification and identification of bacterial isolates. Recently, it has been used by some labs to characterize bacterial mixtures. However, there are still disputes regarding the mechanism of product formation in ERIC-PCR. We cloned and sequenced the 1.1kb major band of the ERIC-PCR fingerprint of *E. coli* K-12 strain MG1655, the strain used for whole genome sequencing, and found that the band consisted of 3 different DNA fragments with one fragment the most abundant (93 out of 95 clones, i.e. 97.89%). Sequence analysis showed that two of the three fragments amplified from a chromosomal region where one ERIC element exist either upstream or downstream, while one fragment amplified from a region where no ERIC element was found. It was thus postulated that ERIC-PCR is not an absolutely random amplified PCR technique, especially when it is applied to those genomes containing ERIC elements.

Key words: ERIC sequence, ERIC-PCR, *E. coli* K-12 MG1655

在许多细菌基因组中发现的 ERIC 序列 (Enterobacterial Repetitive Intergenic Consensus 肠杆菌基因间重复共有序列) 长约 126bp, 具有一定的保守性, 定位在基因组中可转录的非编码区或与转录有关的区域内^[1]。

Versalovic 等以 ERIC 中部反向重复序列为结合位点, 设计了一组反向引物, 认为可用于扩增细菌染色体上两个相邻 ERIC 序列间的片段, 从而得到由一系列大小不同的片段组成的 DNA 指纹图谱, 用于区别不同细菌的基因组^[2]。这项名为 ERIC-PCR 的技术广泛应用在如利斯特菌 (*Listeria*)^[3]、植物病菌黄单胞菌属 (*Xanthomonas*) 和假单胞菌属 (*Pseudomonas*)^[4]、苜蓿根瘤菌 (*Rhizobium meliloti*)^[5] 等微生物的种属、菌株鉴定分类工作中。早期普遍认为, 能扩增出 ERIC-PCR 产物的基因组则意味着它含有 ERIC 序列。

然而 Gillings 等发现以噬菌体、真菌、植物及动物的基因组 DNA 为模板的 ERIC-PCR 反应中, 也可以扩增出相应的特征性产物^[6]。对 GenBank 等 DNA 序列数据库的检索, 并没有在噬菌体或真菌基因组中发现 ERIC 序列; 并且当提高 ERIC-PCR 反应的退火温度接近 ERIC 序列的 T_m 值 (约为 68℃) 时, 除大肠杆菌外, 所有其它供试的基因组 DNA 均不能再产生 ERIC-PCR 图谱。因此 Gillings 认为 ERIC-PCR 实质上是一种随机扩增 PCR (RAPD-PCR), 它所产生的指纹图谱并不能直接证明或代表基因组 DNA 中含有 ERIC 序列, 但这个看法一直缺少分子水平的证据。我们对已经完成全基因组测序的近 40 个基因组的分析也表明, ERIC 序列之间的距离都远远大于 ERIC-PCR 图谱中最大的约 3kb 左右的片段的距离, 因此, 其扩增的显然不是两个相邻的 ERIC 序列之间的片段。

近年来, 在生态学领域, 有报道将 ERIC-PCR 用于混合菌群^[7]甚至动物肠道^[8]、活性污泥^[9]中复杂微生物生态系统的种群结构动态分析以及功能组分的鉴定, 取得较好效果。这些工作需要我们对 ERIC-PCR 扩增产物的形成规律和性质有比较清楚的了解。

本研究中, 我们分析了用于大肠杆菌 K-12 基因组测序的标准菌株 MG1655 的 ERIC-PCR 图谱中 1.1kb 主带的组成及其序列特点, 为进一步认识 ERIC-PCR 产物形成规律、

理解 ERIC-PCR 在应用中产生的各种现象以及促进该技术在分子生态学和基因组学中的应用提供了理论依据。

1 材料与方法

1.1 实验材料

1.1.1 菌株和质粒: *E. coli* K-12 MG1655 菌株由美国威斯康星大学遗传实验室惠赠, 转化受体菌 *E. coli* DH5 α 为本室所保存, 载体 pGEM-T easy 购自 Promega 公司。

1.1.2 主要试剂及酶类: 限制性内切酶 (*Hinf* I, *Rsa* I, *Taq* I 酶), *Taq* DNA 聚合酶, T4 DNA 连接酶, PCR 产物回收试剂盒, DNA marker 等均购自 Promega 公司, IPTG 和 X-Gal 购自北方同正公司, 抗生素、LB 培养基购自华美生物工程公司。

1.1.3 ERIC-PCR 引物: 参照文献 [2]。ERIC1R 5'-ATGTAAGCTCCTGGGGATTAC-3' 和 ERIC2 5'-AAGTAAGTGAAGTGGGGTGAGCG-3'

1.2 实验方法

1.2.1 *E. coli* K-12 MG1655 基因组的提取: 参考 Versalovic 等所使用的总 DNA 提取方法^[2]。

1.2.2 PCR 扩增条件: 按文献 [2] 提供的条件进行。

1.2.3 1.1kb 条带的回收、纯化和克隆: 将 DNA 片段洗脱、回收, 连接到 pGEM-T easy 载体上, 转化到 *E. coli* DH5 α 中, 具体方法参考《分子克隆实验指南》^[10] 及 Promega 公司试剂说明书。

1.2.4 克隆的筛选、酶切分析和测序: 随机从含有 X-Gal 和 IPTG 的平板上筛选 100 个克隆, 用碱裂解法提取质粒; 对质粒进行 *Hinf* I, *Rsa* I 和 *Taq* I 酶的酶切图谱分析; 将图谱有差异的重组质粒进行 DNA 测序, 测序工作由国家人类基因组南方中心完成。

1.2.5 序列结果分析: 用 BLAST、DNAMAN 和 GeneTool 等工具和软件对测序结果进行分析, 并将各片段在 *E. coli* K-12 MG1655 的基因组中定位。

2 结果

2.1 *E. coli* K-12 株的 ERIC-PCR 扩增

扩增产物用 1.5% 的琼脂糖凝胶电泳检测, 结果表明对于不同批次提取的总 DNA, 其扩增图谱具有较高的稳定性和可重复性 (图 1); 图谱中 1.1kb 处的条带亮度最强,

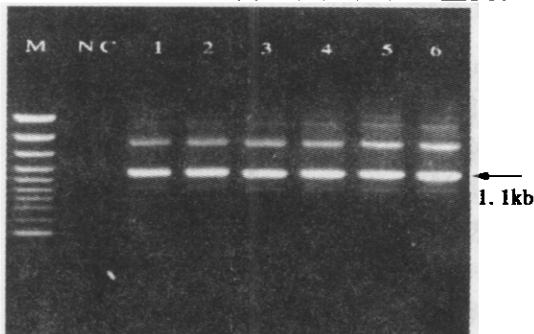


图 1 *E. coli* K-12 株的 ERIC-PCR 指纹图谱
M 100bp Marker, NC 阴性对照, 1~3: 11.14 日提取的总 DNA 的扩增产物 (3 个重复), 4~6: 11.16 日提取的总 DNA 的扩增产物 (3 个重复)

我们称其为主带, 并将其做为研究对象进行下一步的回收、转化和克隆工作。

2.2 重组质粒的提取及酶切分析

随机提取的 100 个质粒, 其中 96 个明显大于空载体对照, 为连入目的片段的重组质粒, 另外 4 个与载体大小相同, 为载体自连, 弃去 (图略)。对 96 个重组质粒进行 *Hinf* I 酶切分析, 共得到 5 种类型的酶切物理图谱 (图 2), 再分别用 *Rsa* I, 和 *Taq* I 酶对 *Hinf* I 酶切图谱中结果相同的重组质粒进行酶切分析, 没有发现新类型, 进一步确认 *Hinf* I 酶切分析结果。每

种类型选一个转化子进行测序。

2.3 重组质粒测序结果及在 *E. coli* K-12 MG1655 菌株基因组中的定位

在 5 种 *Hinf* I 酶切图谱中, 第 60 号和 43 号质粒的序列相同而以相反的方向插入 pGEM-T easy 载体, 第 38 号质粒为 PCR 产物嵌合体, 因此 95 个重组质粒中共得到 3 种含不同序列的重组子, 分别命名为 pCF1, pCF3, pCF4, 其插入片段分别称为 SEQ-1、SEQ-3 和 SEQ-4。我们通过 BLAST 将这 3 个产物片段在 *E. coli* 基因组上进行定位, 在这 3 个产物片段中, SEQ-1 和 SEQ-4 所在基因组区域的 5' 端均有 ERIC 序列, 而 3' 末端则没有, SEQ-3 是 95 个重组子中唯一所在基因组区域的两端均不含有 ERIC 序列的片段。在 3 种质粒中 pCF1 在所有重组子中所占的比例达到了 97.89%, 表明产物片段 SEQ-1 的扩增量最大。

我们用 DNAMAN 软件对这 3 个序列两端引物的匹配情况进行了分析, 结果如表 2。无论是看 Identity 值还是 3' 末端的匹配情况 SEQ-1 都有明显的优势。

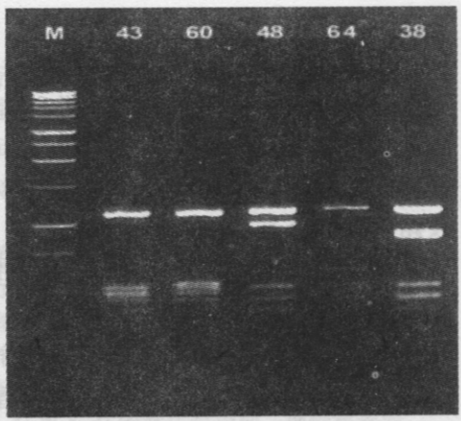


图 2 具有不同插入序列的重组子的 *Hinf* I 酶切图谱
M 1kb Marker, 43, 60, 48, 64, 38 具有不同 *Hinf* I 酶切图谱的重组质粒编号

表 1 3 种 DNA 序列在基因组中的定位及其与 ERIC 序列的关系

重组质粒	pCF1	pCF3	pCF4
产物片段	SEQ-1	SEQ-3	SEQ-4
<i>E. coli</i> 基因组中的位置 (bp)	2654492 ~ 2655651	1880528 ~ 1881422	253401 ~ 254609
重组质粒丰度	97.89% (93/95)	1.05% (1/95)	1.05% (1/95)
扩增区域 ERIC 序列存在情况	5'	无	5'

表 2 各插入片段 5' 和 3' 末端分别与引物 ERIC2 和 ERIC1R 的相似性比对

上游引物匹配	下游引物匹配	丰度
SEQ-1:ERIC2 identity= 86% 5' AACTATGTGACTGGGGTGAACG 3' 	SEQ-1:ERIC2 identity= 90% 3' CGCTCACCCTGCATTACATGG 5' 	97.89%
5' AAGTAAGTGACTGGGGTGAGCG 3'	3' CGCTCAGCCAGTCACTTACT 5'	
SEQ-3:ERIC2 identity= 59% 5' GCATGCCGTACTGGGATGAGCG 3' 	SEQ-3:ERIC1R identity= 57% 3' GTTCA.CAACAGGCACCTCGAAT 5' 	1.05%
5' AAGTAAGTGACTGGGGTGAGCG 3'	3' GTGAATCCCCAGGAGCTTACAT 5'	
SEQ-4:ERIC2 identity= 95% 5' AAGTAAGTGACTGGGGTGAACG 3' 	SEQ-4:ERIC1R identity= 45% 3' GCTCACCACGTAGTCTTTAC 5' 	1.05%
5' AAGTAAGTGACTGGGGTGAGCG 3'	3' GTGAATCCCCAGGAGCTTAC 5'	

3 讨论

ERIC-PCR 所扩增的片段大多为基因组中 ERIC 序列附近的 DNA 片段, 而并非如 Versalovic 设想的为两个 ERIC 序列之间的 DNA 片段。这主要是因为, 在 4,600kb 左右的大肠杆菌基因组中含有约 30 个 ERIC 序列, 平均每两个 ERIC 序列间的距离约为 1×10^5 bp, 这超出了 PCR 反应的有效扩增范围。

对含有 ERIC 序列的细菌基因组进行 ERIC-PCR 扩增时, 其指纹图谱中的条带也非 Gillings 认为的是一种简单的随机扩增 (RAPD-PCR) 结果。分析 *E. coli* K-12 MG1655 菌株中 1.1kb 条带发现, 它由丰度不等的多个片段组成, 其中大部分的扩增片段是从基因组中 ERIC 序列附近得到的, 这些扩增片段的一端含有 ERIC 序列, 而另一端没有 ERIC 序列。

在所有的扩增片段中, 仅有一少部分序列两端都不含有 ERIC 序列, 如重组质粒 pCF3 中的 SEQ-3, 这可能是由于 SEQ-3 的两端与 ERIC 引物的相似性较高, 以致于被引物识别, 从而在 PCR 反应中被扩增出来, 这是随机扩增的结果。但 96 个重组质粒中仅发现一例, 所占比例很小。

可以推断, 如果细菌基因组中含有 ERIC 序列, 那么 ERIC-PCR 的扩增产物通常一端含有 ERIC 序列, 而另一端则是随机匹配; 如果细菌基因组中不存在 ERIC 序列, ERIC-PCR 就成为一种较特殊的随机扩增: 其引物较长 (22 个寡聚核苷酸), 因此可产生具有较高可重复性和一定特异性的指纹图谱。因此我们认为 ERIC-PCR 是一种半随机性质的 PCR。

1.1kb 条带转化的克隆中, pCF1 的克隆数占到总体的 97.89%, 可以认为 pCF1 中的插入片段 SEQ-1 是 PCR 反应的主要产物; 同时表明在 ERIC-PCR 扩增中存在优先扩增的现象, 即引物与模板的结合机会不均一。所以当对含有 *E. coli* K-12 MG1655 菌株的混合菌群样品进行 ERIC-PCR 扩增时, 在混合样品的指纹图谱中的 1.1kb 处的条带中, 代表 *E. coli* K-12 MG1655 菌株的应该是序列 1 (SEQ-1), 对此, 我们进行了实验验证, 结果将另文发表。对 ERIC-PCR 产物形成规律的比较准确的认识, 为在分子生态学和基因组学中的应用提供了理论依据。

参考文献

- [1] Hulton C S J, Higgins C F, Sharp P M. *Mol Microbiol*, 1991, 5: 825 ~ 834.
- [2] Versalovic J, Koeuth T, Lupski J R. *Nucleic Acids Research*, 1991, 24: 6823 ~ 6831.
- [3] Jersek B, Tcheneva E, Rijpens N, *et al.* *Letters in Applied Microbiology*, 1996, 23: 55 ~ 60.
- [4] Louws F J, Fulbright D W, Stephens C T, *et al.* *Appl Environ Microbiol*, 1994, 60: 2286 ~ 2295.
- [5] Frans J B R U U N. *Appl Environ Microbiol*, 1992, 58: 2180 ~ 2187.
- [6] Gillings M, Holley M. *Letters in Applied Microbiology*, 1997, 25: 17 ~ 21.
- [7] Giovanni G D Di, Watrud L S, Seidler R J, *et al.* *Microb Ecol*, 1999, 37: 129 ~ 139.
- [8] Song G, Wei G, Zhao L, *et al.* *Book of Abstracts of the Asia Pacific Conference On Plant Tissue Culture & Agribiotechnology* 19 ~ 23 November 2000 Singapore G16 166.
- [9] Zhao L, Cao P, Wang J, *et al.* 101st Annual Meeting of American Society of Microbiology, 2001, Poster Presentation.
- [10] 萨姆布鲁克 J, 弗里奇 E F, 曼尼阿蒂斯 T. *分子克隆实验指南* (第二版). 北京: 科学出版社, 1993.