

## 脉冲场凝胶电泳技术 在本草基因组计划中的应用\*

□刘明珠\*\* 白志川 (西南大学园艺园林学院 重庆 400716)

陈士林 孙超\*\* 何柳 张进

(中国医学科学院 药用植物研究所 北京 100193)  
(北京协和医学院)

**摘要:**本文着重介绍了脉冲场凝胶电泳(PFGE)在本草基因组测序中的应用,包括测定药用真菌染色体数目和基因组大小、构建基因组文库及绘制物理图谱,并对 PFGE 使用中一些待解决的问题做了简要介绍。

**关键词:**PFGE 本草基因组计划 基因组大小 基因组文库 物理图谱

doi: 10.3969/j.issn.1674-3849.2010.03.028

中医药是中国传统国粹之一,是中华民族的无价瑰宝,受到国内外医药与科技界的广泛关注。然而中医药学基础研究起步较晚,加强中医药基础研究,已成为 21 世纪中医药事业发展内在的迫切需求。近年来,随着新一代测序技术即高通量测序技术的发展,物种的 De Novo 全基因组测序变得更加快速、便捷,此时本草基因组计划的启动,必将为我国中医药事业的国际化发展起到巨大的推动作用。该计划的运作不仅有利于我国中药研究水平的全面提升,而且在物种鉴定、栽培育种、关键酶基因的发掘、活性化学成分分离及药理毒理学药物研发、合成以及生物转化、天然药物的产业化等方面都将有深入广泛的应用。

脉冲场凝胶电泳技术(Pulsed Field Gel Electrophoresis, PFGE)是以分离大片段线型 DNA 分子(50kb~12Mb)为目的的一种非常成熟的电泳技术,也是本草基因组计划实施过程中必需的技术手段。

### 一、PFGE 原理及方法

本草基因组 De Novo 测序是一项非常复杂的工作,需要使用到大量的仪器设备和技术手段,脉冲场凝胶电泳技术(Pulsed Field Gel Electrophoresis, PFGE)是其中使用较多也较重要的技术之一,经过多年不断的发展和进步,PFGE 已发展出多种类型,箱位匀强电场技术(Contour-Clamped Homogeneous Electric Field, CHEF)是现在用于微生物电泳核型分析和 DNA 大小测定等领域最受欢迎和应用最为广泛的一种。

收稿日期: 2010-06-02

修回日期: 2010-06-12

\* 国家自然科学基金面上项目(30900113):基于宏量 ESTs 的蛇足石杉转录组分析及石杉碱甲合成酶(HAS)基因的鉴定研究,负责人:罗红梅。

\*\* 联系人:刘明珠,硕士研究生,主要研究方向:植物分子生物学, E-mail: liuyanhua1004@163.com; 孙超,副研究员,主要研究方向:药用植物功能基因组学, E-mail: csun@implad.ac.cn。

PFGE 是通过在电泳槽的两个不同方向上加交替周期性电场下, DNA 分子相应以 Z 字形方式在琼脂糖凝胶连续孔洞中迁移, DNA 分子在交替变换方向的电场中做出反应所需时间取决于它的大小, 较小分子量 DNA 重新定向较快, 因而在凝胶中的迁移也较快, 于是不同大小的 DNA 分子得以被成功分离<sup>[1]</sup>。传统琼脂糖凝胶电泳技术是通过定向的电场作用, 将带负电的 DNA 分子由负极向正极直线迁移, PFGE 打破了传统电泳技术只能分离小于 50Kb DNA 分子片段的壁垒, 早在 1983 年, Schwartz 和 Cantor 就利用 PFGE 分离出超过 1 Mb 的大片段 DNA 分子<sup>[2-3]</sup>。它的这一特性使其在很多领域都有广泛的应用, 如用于流行病学方面的病原菌染色体分型鉴定, 由美国 CDC 建立的 PulseNet 就是这一方面应用的实例, 直到今天依然是热门研究领域\*。此外, 还包括微生物染色体数目与基因组大小的测定、细胞凋亡、转基因、基因定位、物理图谱、基因组文库的建立等方面的应用。本文就 PFGE 在全基因组测序方面的广泛应用进行了简要阐述。

## 二、PFGE 在本草基因组测序中的应用

随着 PFGE 在全基因组测序中的应用越来越广泛, 本草基因组测序计划的许多方面都将离不开这一技术的应用, 利用这一技术进行染色体条数和基因组大小的测定是对物种全基因组测序前的必要环节, PFGE 为本草基因测序的选择, 成本估算, 覆盖度估测和物理图谱的制作等提供了有效依据。PFGE 也是构建基因组文库及物理图谱的绘制时必须用到的技术。

### 1. 药用真菌染色体数目和基因组大小的测定

中国药典中包含了许多名贵的药用真菌, 如灵芝、虫草、猪苓、茯苓、银耳等, 这些无疑都将是本草基因计划的备选物种。相关文献中报道<sup>[4-13]</sup>, 对真菌等基因组较小的物种进行染色体数目及基因组大小的测定是进行全基因组测序的必要前期工作, 文章中将测序所得基因组大小和 PFGE 测得的基因组大小进行了比

较, 见表 1。从中可以看出, 测序所得结果和 PFGE 测定结果具有很好的一致性。另一种可用于测定药用真菌基因组大小的方法是流式细胞术(Flow Cytometry, FCM), FCM 的结果也很好的验证了 PFGE 预测的准确性<sup>[14-15]</sup>。目前, 利用 PFGE 对真菌进行染色体分型已经是非常成熟的一项技术, 早在 1993 年和 2003 年, Fierro Francisco 等<sup>[11]</sup>和 Tanesaka Eiji 等<sup>[16]</sup>就使用 PFGE 对大型真菌青霉菌(Penicillium)和金针菇(Flammulina velutipes)进行过染色体的分型。

利用 PFGE 进行药用真菌染色体数目和基因组大小的测定, 大致可以分为 5 个步骤: 样品的选择和准备, 原生质体的获得、胶栓的制备和去膜、脉冲场凝胶电泳, 染色及成像分析。

### 2. 构建基因组文库

基因组文库是将细胞的核 DNA 酶切或随机打断成一定大小的片段, 然后将全部片段随机地连接到各种基因载体上, 再转移至适当的宿主细胞中, 通过宿主细胞的大量繁殖而形成的各个片段的无性繁殖系的总集。大片段基因组文库包括噬菌体、YAC、BAC、Cosmid 和 Fosmid 等文库, 主要根据他们插入片段大小和构建的载体不同而进行分类的。当前几乎所有的模式物种及许多重要濒危物种都已建立了基因组文库。然而在传统中药研究领域, 相对还是个盲区, 除人参等少数几个物种外<sup>[17]</sup>, 其它许多重要物种都还鲜见有构建基因组文库的文献报道。

构建基因组文库也是全基因组测序中的一项重要工作, 建立不同长度插入片段的基因组文库, 为全基因组序列的拼接组装及基因组物理图谱的绘制提

表 1 部分物种 PFGE 测定结果与全基因组测序结果比较

物种名称	拉丁名	染色体数目	PFGE 测定基因组大小 (Mb)	全基因组测序测得大小 (Mb)	参考文献
巴斯德毕氏酵母	<i>Pichia pastoris</i>	4	9.7	9.43	5
柄孢霉	<i>Podospora anserina</i>	5	33.4	35.5-36	6
都柏林念珠菌	<i>Candida dubliniensis</i>	13	16-17	14.6	7,8
酿酒酵母	<i>Saccharomyces cerevisiae</i>	16	11.8	11.6	9
红球虫赤壳	<i>Nectria haematococca</i>	17	51	54.43	10
里氏木霉	<i>Trichoderma reesei</i>	7	33	34	11
产黄青霉	<i>Penicillium chrysogenum</i>	4	33.3-34.1	32.2	12,13
假丝酵母属	<i>Candida</i>	6-9	7.88-14.36	10.6-15.5	13

\* [http://www.cdc.gov/PULSENET/\[OL\]](http://www.cdc.gov/PULSENET/[OL]).

供必要条件。一般在 De Novo 全基因测序中应用最多的是插入片段为约 100 Kb 的 BAC 文库和 40 Kb 左右大小插入片段的 Fosmid 文库,已有大量文献报道<sup>[5,18-19]</sup>。构建 BAC 或 Fosmid 文库时,其插入片段长度都已超过普通电泳分离阈值,而 PFGE 技术的应用有效地解决了这一难题。

文库构建中主要有以下几个方面需要应用到 PFGE 技术:特定长度 DNA 插入片段的分离,已分离 DNA 插入片段的大小鉴定,建库质量鉴定(包括大小及稳定性的鉴定)。图 1 简要演示了两种文库的构建流程及 PFGE 在其中的应用。

### 3. 物理图谱的绘制

基因组物理图谱,是指一种采用分子生物学技术直接将 DNA 分子标记,基因或克隆标定在染色体实际位置上的图谱,相对遗传图谱而言,这一位置的标定是实际、可测的<sup>[20]</sup>。物理图谱的绘制方法主要有以下几种:限制性内切酶制图、BAC 文库构建制图、荧光原位杂交技术、放射性杂交细胞系技术,Happy 作图、光学图谱等。其中限制性内切酶制图法和 BAC 文库是最经典的作图技术,而后几种技术发展和应用的时间相对较短。

在限制性内切酶和 BAC 文库绘制物理图谱时,PFGE 是最常用的,也是必需的技术之一,主要用于分离酶切后片段,鉴定目的片段长度等。图 2 简要演示了 PFGE 在其中的应用。

## 三、亟待解决的问题

### 1. 样品制备复杂且易污染

利用 PFGE 分离药用真菌染色体时,其完整染色体的制备过程非常复杂且极易受人为因素的影响。虽然相关文献书籍中介绍的标准操作流程很多,但其操作技巧性大,结果重复性差。保证染色体的完整性以及不受污染是样品成功制备的关键,成功的样品是得到清晰锐利染色体带型基础。

### 2. 分辨力不足

虽然随着 PFGE 有关仪器设备及试剂等的不断进步和发展,现在已经能比较好的分离出  $\geq 5\text{Mb}$  的染色体条带,但是相对于这一领域的应用需求,还远远不够,有关两篇文献的比较很好的说明了这一问题,2000 和 2003 年,Doolyi Kim 和 Eiji Tanesaka 等均使用 PFGE 对大型真菌 *Flammulina velutipes* 进行过染色体的分型,结果差异显著,后者明显优于前者<sup>[16,21]</sup>,但 PFGE 在分离超大染色体时带型不清晰,不能成带而易成团,而对大小相近的染色体的分离效果亦不明显。

### 3. 成本较高且耗时

目前,PFGE 仪器主要是美国伯乐公司生产,相关试剂也有许多都是伯乐独家出售,如最大的 DNA 标准品粟酒裂殖酵母 *S. pombe* 染色体 DNA,可用于分离 10Mb DNA 的脉冲场电泳专用琼脂糖等,这些仪器配件和试剂等都价格不菲,实验成本很高。

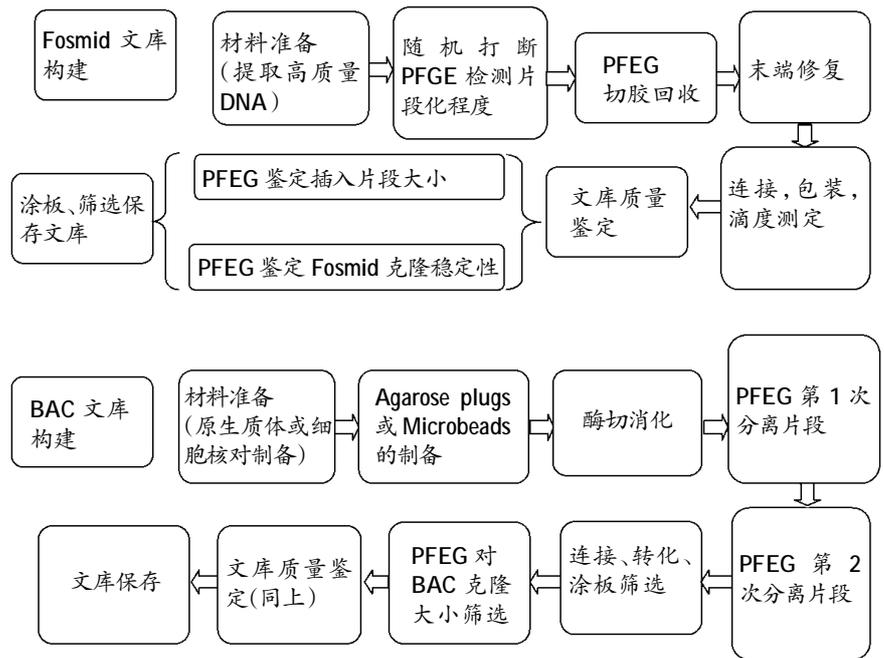


图 1 Fosmid 及 BAC 文库的简要流程图及 PFGE 在其中的应用

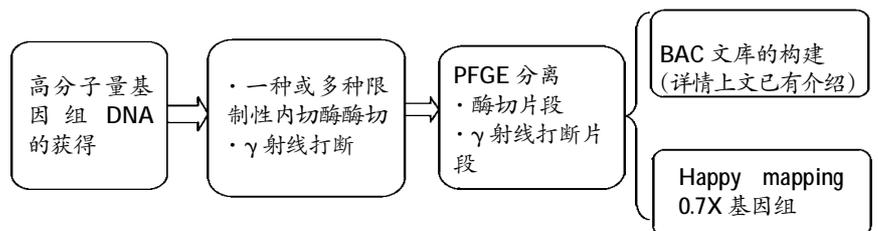


图 2 PFGE 在物理图谱构建中的应用

对于使用 PFGE 分离大片段的 DNA 或完整染色体,除了分离效果可能会不尽如人意外,分离速度也非常慢,比如分离 $\geq 1\text{Mb}$ 的 DNA,需要耗时多达 5 天或以上<sup>[22]</sup>。

#### 四、展 望

随着全基因组测序技术的发展日新月异,相关设备仪器的更新换代也越来越快,可以预见在不久的将来,测序技术还会有更大的突破,进行大规模大基因组物种的测序也将成为可能。本草基因组计划的实施依赖于仪器设备和技术的更新发展,而 PFGE 在其中的应用也必将日益广泛,本草基因组计划工程浩大,是一项长期的工作,其中用到 PFGE 的地方非常多,虽然 PFGE 依然存在着一些问题和不足,但这也必将促使 PFGE 的进一步革新和进步。

#### 参考文献

- 1 J· 萨姆布鲁克, D·W· 拉塞尔. 分子克隆实验指南. 3 版. 黄培堂, 译. 北京: 科学出版社, 2002: 429~450.
- 2 Schwartz DC, Cantor CR. Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell*, 1984, 37: 67~75.
- 3 Schwartz DC, Saffran W, Welsh J, et al. New techniques for purifying large DNAs and studying their properties and packaging. *Cold Spring Harb. Symp. Quant. Biol*, 1983, 47 (Part 1):189~195.
- 4 De Schutter K, Lin YC, Tiels P, et al. Genome sequence of the recombinant protein production host *Pichia pastoris*. *Nature Biotechnology*, 2009, 27:561~566.
- 5 Espagne E, Lespinet O, Malagnac F, et al. The genome sequence of the model ascomycete fungus *Podospora anserina*. *Genome Biology*, 2008, 9:R77.
- 6 Magee BB, Sanchez MD, Saunders D, et al. Extensive chromosome rearrangements distinguish the karyotype of the hypovirulent species *Candida dubliniensis* from the virulent *Candida albicans*. *Fungal Genet Biol*, 2008, 45:338~350.
- 7 Jackson AP, Gamble JA, Yeomans T, et al. Comparative genomics of the fungal pathogens *Candida dubliniensis* and *Candida albicans*. *Genome Res*, 2009, 19(12):2231~2244.
- 8 Argueso JL, Carazzolle MF, Mieczkowski PA, et al. Genome structure of a *Saccharomyces cerevisiae* strain widely used in bioethanol production. *Genome Res*, 2009, 19:2258~2270.
- 9 Coleman JJ, Rounsley SD, Rodriguez-Carres M, et al. The Genome of *Nectria haematococca*: Contribution of Supernumerary Chromosomes to Gene Expansion. *PLoS Genet*, 2009, 5 (8):e1000618. doi:10.1371/journal.pgen.1000618.
- 10 Martinez D, Berka RM, Henrissat B, et al. Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). *Nature Biotechnology*, 2008, 26:553~560.
- 11 Fierro F, Gutiérrez S, Diez B, Martín JF.. Resolution of four large chromosomes in enicillin-producing filamentous fungi: the penicillin gene cluster is located on chromosome II (9.6 Mb) in *Penicillium notatum* and chromosome I (10.4 Mb) in *Penicillium chrysogenum*. *Mol Gen Genet*, 1993, 241:573~578.
- 12 Cornman RS, Chen YP, Schatz MC, et al. Genomic Analyses of the Microsporidian *Nosema ceranae*, an Emergent Pathogen of Honey Bees. *PLoS Pathog*, 2009, 5(6):e1000466.
- 13 Butler G, Rasmussen MD, Lin MF, et al. Evolution of pathogenicity and sexual reproduction in eight *Candida* genomes. *Nature*, 2009, 459:657~662.
- 14 O' Sullivan D, Tosi P, Creusot F, et al. Variation in genome organization of the plant pathogenic fungus *Colletotrichum lindemuthianum*. *Curr Genet*, 1998, 33:291~298.
- 15 Solieri L, Cassanelli S, Croce MA, et al. Genome size and ploidy level: New insights for elucidating relationships in *Zygosaccharomyces* species. *Fungal Genetics and Biology*, 2008, 45 :1582~1590.
- 16 Tanesaka E, Kinugawa K, Okabe K, et al. Electrophoretic karyotype of *Flammulina velutipes* and its variation among monokaryotic progenies. *Mycoscience*, 2003, 44:67~69.
- 17 Hong CP, Lee SJ, Park JY, et al. Construction of a BAC library of Korean ginseng and initial analysis of BAC-end sequences. *Mol Gen Genomics*, 2004, 271:709~716.
- 18 Huang SW, Li RQ, Zhang ZH, et al. The genome of the cucumber, *Cucumis sativus* L. *Nature Genetics*, 2009, 41:1275~1281.
- 19 Espagne E, Lespinet O, Malagnac F, et al. The genome sequence of the model ascomycete fungus *Podospora anserina*. *Genome Biology*, 2008, 9:R77.
- 20 覃瑞, 宋发军, 宋运淳. 植物基因组比较作图研究进展. *细胞生物学杂志*, 2004, (2):68~72.
- 21 Kim D, Tamai Y, Azuma T, et al. Analysis of the electrophoretic karyotype of *Flammulina velutipes*. *The Japan Wood Research Society*, 2000, 46:466~469.
- 22 Herschleb J, Ananiev G, Schwartz DC. Pulsed-field gel electrophoresis. *Nature protocols*, 2007, 2(3):677~684.

#### Application of Pulsed Field Gel Electrophoresis in the Herbal Genome Program (HerbGP)

Liu Mingzhu<sup>1</sup>, Bai Zhichuan<sup>1</sup>, Chen Shilin<sup>2</sup>, Sun Chao<sup>2</sup>, He Liu<sup>2</sup>, Zhang Jin<sup>2</sup>

1. Horticulture and Landscape College, Southwest University, Chongqing 400716, China;
2. Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100193, China)

**Abstract:** Pulsed Field Gel Electrophoresis (PFGE), which was developed in the 1980s, is mainly used to separate large linear DNA molecules (50 kb -12 Mb). With the development and implementation of the herbal genome program (HerbGP), PFGE is playing an increasingly critical role. In this study, we focus on the PFGE application in herbal genomic sequencing, including the measurement of the chromosome number and genome size of medicinal fungi, as well as construction of the genomic library and physical mapping. We also discuss the problems of how to use PFGE.

**Keywords:** PFGE, HerbGP, Genome size, Genome library, Physical map

(责任编辑:王 瑀,责任译审:张立崑)