

# 宏基因组克隆——微生物活性物质筛选的新途径\*

阎冰<sup>1,2</sup> 洪葵<sup>1\*\*</sup> 许云<sup>1</sup> 马超<sup>1</sup>

(中国热带农业科学院热带生物技术研究所热带作物生物技术国家重点实验室 海口 571101)<sup>1</sup>

(广西红树林研究中心 北海 536000)<sup>2</sup>

**摘要:** 在现有技术条件下自然界存在的微生物 95% 以上未能培养, 采用传统的分离培养筛选的途径寻找新的微生物生物活性物质受到局限; 宏基因组是特定小生境中全部微生物遗传物质的总和, 直接抽提环境样品中的总 DNA, 利用适宜的载体克隆到替代宿主细胞中构建宏基因组文库, 通过外源基因赋予宿主细胞的新性状或基于某些已知 DNA 序列筛选, 寻找新的生物活性物质或基因, 极大地扩展了微生物资源的利用空间, 增加了获得新的生物活性物质的机会。

**关键词:** 宏基因组, 环境总 DNA, 功能驱动筛选, 序列驱动筛选

**中图分类号:** Q93 **文献标识码:** A **文章编号:** 0253-2654 (2005) 01-0113-05

## Metagenome Cloning—A New Approach for Novel Microbial Bioactive Compounds Discovery

YAN Bing<sup>1,2</sup> HONG Kui<sup>1</sup> XU Yun<sup>1</sup> MA Chao<sup>1</sup>

(*Institute of Tropical Biological Sciences, Chinese Academy of Tropical Agricultural Science, State Key Lab for Tropical Crops Biotechnology, Haikou 571101*)<sup>1</sup>

(*Guangxi Mangrove Research Center, Beihai 536000*)<sup>2</sup>

**Abstract:** It is reported that more than 95% microbe species non-cultured or un-cultured yet in now days by the traditional cultivation approach, which limits discovering novel bioactive compounds from microorganisms. 'Metagenome' is the genomes of the total microbiota found in nature. Metagenome libraries were constructed by directly extracting DNA from environmental sample and transforming to surrogate host. The libraries were screened for novel bioactive compounds or genes surrounding their synthase in different strategies of function-driven or sequence-driven. These have enormously amplified the space of microbial resource utilization and enhanced the opportunity of obtain novel bioactive compounds.

**Key words:** Metagenome, Environmental total DNA, Function-driven screening, Sequence-driven screening

微生物活性物质筛选传统方法的基础是微生物分离, 获得纯培养, 其研究流程包括微生物分离纯化、大量培养、活性检测、活性物质的分离纯化直至开发利用。这一途径被证明是行之有效的, 沿着这一途径找到了许多活性物质, 但重复发现率特别高。分子微生物生态学研究表明环境中大量存在未能培养的微生物, 据估计每克土壤样品中可含有高达 4,000 种的不同微生物, 某些环境中采用现有培养技术能够培养的微生物不到 1%。最近, 一个全新的理念——“宏基因组克隆”: 直接提取特定环境中的总

\* 国家“863”海洋生物技术青年基金项目资助 (No. 2002AA628140)

海南省普通高校优秀中青年教师招聘和教学奖励基金资助

\*\* 联系人 E-mail: k1022@163.net

收稿日期: 2004-04-15, 修回日期: 2004-05-31

DNA, 克隆到可培养的宿主细胞中, 从所获得的重组克隆子中筛选活性物质和相关基因, 显示了其在挖掘和利用那些未能培养微生物的基因资源筛选新生物活性物质的潜力。

### 1 宏基因组及其研究流程

“Metagenome”是由 Handelsman 等 1998 年提出的新名词, 其定义为“the genomes of the total microbiota found in nature”<sup>[1]</sup>即生境中全部微小生物遗传物质的总和, 国内学者将其翻译为“宏基因组”, 文献中常见的“collective genome”、“environmental genome”等, 现已普遍采用“Metagenome”, 且主要指环境样品中的细菌和真菌的基因组总和。宏基因组文库既包含了可培养的又包含了未可培养的微生物基因, 避开了微生物分离培养的问题, 极大地扩展了微生物资源的利用空间, 增加了获得新的生物活性物质的机会。图 1 对采用宏基因组克隆策略与传统培养途径筛选活性物质的研究流程进行了简略比较。

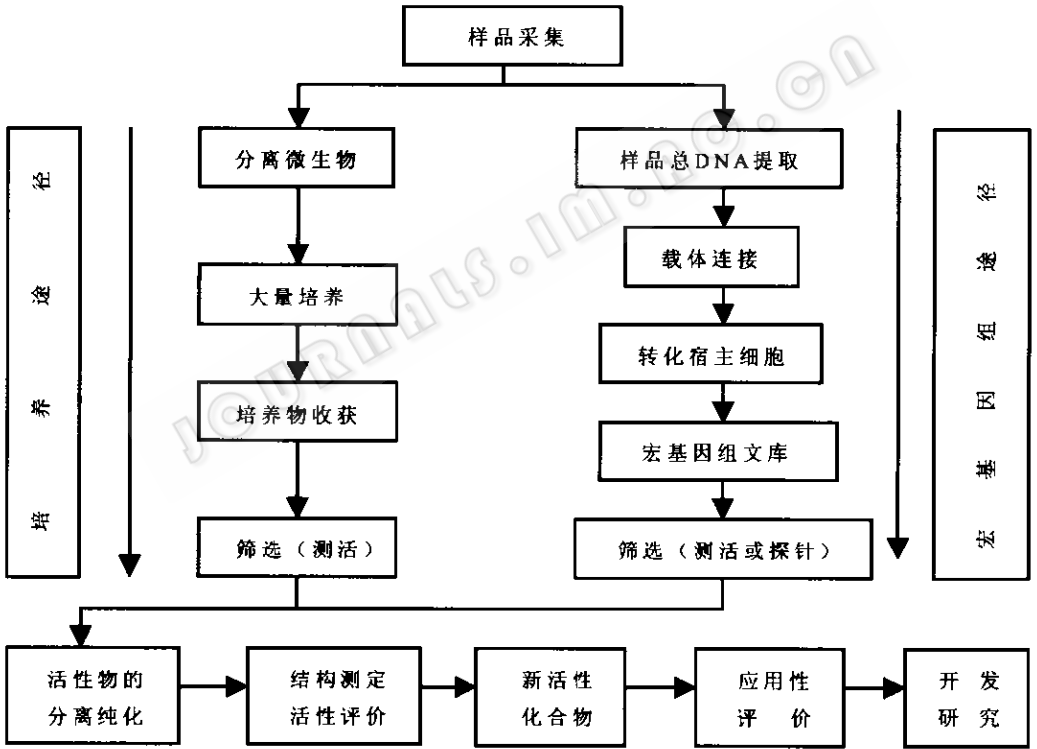


图 1 采用传统培养途径和宏基因组途径筛选新的微生物活性物质研究流程的简略比较

### 2 宏基因组克隆筛选生物活性物质研究概况

1991 年 Pace 等在研究海洋微小浮游生物时, 直接提取浮游生物群体的总 DNA 经酶切后通过 λ 噬菌体载体克隆到大肠杆菌中构建了 eDNA 文库, 并从中筛选到了 rRNA 基因, 展示了从环境中克隆任何生物体基因的可能性。采用宏基因组克隆最先筛选到的生物活性物质是具有抗菌活性的五种新的小分子物质 TerragineA、B、C、D、E, 由

加拿大 TerraGen Discovery 公司从以链霉菌为宿主的宏基因组文库中筛选到<sup>[2]</sup>。德国哥廷根大学的 Daniel R. 实验室从土壤样品直接提取 DNA, 利用表达载体 pSK<sup>+</sup> 转化到大肠杆菌中成功构建了宏基因组文库, 用 4-羟基丁酸作唯一的碳源和能源筛选到了 5 个能够稳定利用 4-羟基丁酸的克隆子, 随后的研究中先后获得了新的生物素合成操纵子、脂(酯)酶、琼脂糖酶、乙醇脱氢酶、甘油脱水酶等<sup>[3-7]</sup>。由于细菌合成某些活性物质的基因经常以基因簇的形式出现, 美国威斯康星大学的 Handelsman 实验室利用细菌人工染色体载体 pBeloBAC II 构建了两个土壤样品的宏基因组 BAC 文库, 筛选出了一个分泌脱氧核糖核酸酶、1 个有抗菌活性、2 个分泌酯酶、8 个分泌淀粉酶、29 个有溶血活性的克隆子, 随后对 1 个棕色重组克隆子进行了亚克隆研究, 获得了 2 个具有广谱抗菌作用的新抗生素 TurbomycinA 和 B 及其合成酶基因簇<sup>[8, 9]</sup>。美国康乃尔大学的 Baker 实验室采用 Comsid 载体构建宏基因组文库, 筛选到了紫色菌素和脱氧紫色菌素、长链-N-酰基氨基酸抗菌化合物及其合成酶基因簇、长链-酰基脂肪酸苯酚及其合成酶基因簇<sup>[10]</sup>。此外, Curtis<sup>[11]</sup>和 Pie 等分别从土壤和甲虫宏基因组文库中获得了新的抗肿瘤活性物质聚酮的合成酶基因簇, DiazTorres<sup>[12]</sup>采用人唾液、Cottrell<sup>[13]</sup>采用海水构建宏基因组文库分别筛选到了新的几丁质酶和四环素抗性基因等等。目前已采用土壤、海水、海洋浮游生物、海绵、甲虫、人唾液等环境样品成功构建了宏基因组文库, 已筛选到脂酶/酯酶、蛋白酶、淀粉酶、氧化酶、几丁质酶、核酸酶、膜蛋白、4-羟基丁酸代谢酶系、生物素合成酶系、色素、抗菌抗肿瘤活性物质及抗生素抗性基因等(表 1)。

表 1 几个已构建的宏基因组文库特点及其筛选到的活性克隆

DNA 来源	载体	克隆子数	插入片段大小	宿主	活性类型	参考文献
土壤	pBeloBAC II	24, 576	≥40 kb	大肠杆菌	核酸酶、淀粉酶、抗菌	[8, 9]
土壤	pWE15	35, 000	30 ~ 40 kb	大肠杆菌	生物素合成酶	[4]
土壤	pSK <sup>+</sup>	400, 000	3.0 ~ 5.6 kb	大肠杆菌	乙醇氧化还原酶	[6]
土壤	pWE15	1, 532	25 ~ 40 kb	大肠杆菌	淀粉酶、琼脂糖酶、酰胺酶	[5]
土壤	pBluescriptSK <sup>+</sup>	930, 000	5 ~ 8 kb	大肠杆菌	4-羟基丁酸代谢酶系	[3]
土壤	SuperCos I	700, 000	—	大肠杆菌	抗菌	[10]
土壤	—	1, 020	—	链霉菌	抗菌	[2]
人唾液	TOPO-XL	450	0.8 ~ 3.0 kb	大肠杆菌	四环素抗性	[14]
海水	Δ-Zap II	825, 000	1.8 ~ 6.1 kb	大肠杆菌	几丁质酶	[13]

### 3 宏基因组文库的构建

宏基因组文库的构建沿用了分子克隆的基本原理和技术方法, 并根据具体环境样品的特点和建库目的采用了一些特殊的步骤和对策。

**3.1 土壤样品总 DNA 的提取** 获得高质量土壤样品中的总 DNA 是宏基因组文库构建的关键之一。既要尽可能地完全抽提出样品中的 DNA, 又要保持其较大的片段以获得完整的目的基因或基因簇, 另外土壤中存在腐殖酸类物质强烈抑制分子克隆操作过程中多种酶的活性, 须尽量除去。提取方案大致上可以分为两类, 一类是原位裂解法, 将土壤样品直接悬浮在裂解缓冲液中处理, 继而抽提纯化, 此法操作容易、成本低、DNA 提取率高、偏差小, 但由于强烈的机械剪切作用, 所提取的 DNA 片段较小(1 ~ 50 kb), 且腐殖酸类物质也难以完全去除; 另一类是异位裂解法, 先采用物理方法将微

生物细胞从土壤中分离出来,然后采用较温和的方法抽提DNA,如先采用尼可登介质密度梯度离心分离微生物细胞,然后包埋在低熔点琼脂糖中裂解,脉冲场凝胶电泳回收DNA。此法可获得大片段DNA(20~500 kb)且纯度高,但操作繁琐,成本高,有些微生物在分离过程中可能丢失,温和条件下一些细胞壁较厚的微生物DNA抽不出来。徐平等采用了用微波法快速提取放线菌基因组DNA,该法曾被Orsini用于土壤细菌基因组DNA的提取<sup>[14]</sup>。

**3.2 载体** 载体的选择主要是针对有利于感兴趣的目标基因的扩增、表达及在筛选细胞毒类物质时表达量的调控等。很多微生物活性物质是其次生代谢产物,代谢途径由多基因簇调控,尽量插入大片段DNA以获得完整的代谢途径多基因簇是很有必要的。目前多采用细菌人工染色体(BAC)和粘粒(Cosmid)载体,前者插入片段大(可达350 kb),但克隆效率低,后者插入片段中等(20~40 kb),克隆效率高。BAC和Cosmid载体在宿主细胞中的稳定性高,但拷贝数低,宏基因的扩增困难,表达量低。为了提高宏基因的表达便于重组克隆子活性检测,有研究者直接利用表达载体构建宏基因组文库。表达载体可插入的宏基因片段一般小于10 kb,适合于筛选单一基因或小的操纵子产物。外源基因的表达受宿主细胞的遗传类型、细胞基质、细胞的生理状态及初级代谢产物等的影响,利用穿梭载体扩大宿主范围有利于促使和提高外源基因的表达。为提高和调控外源基因的拷贝数与表达量,常需构建不同类型的载体,如Handelsman实验室构建的superBAC-X载体系列,以pBeloBAC II为骨架,添加各种复制原点,调控其拷贝数及宿主范围,利于外源基因的表达及调控其表达量<sup>[15]</sup>。

**3.3 宿主** 宿主菌株的选择主要考虑转化效率、重组载体在宿主细胞中的稳定性、宏基因的表达、目标性状(如抗菌)缺陷型等。研究经验表明不同微生物种类所产生的活性物质类型有明显差异,不同的研究目标应选择不同的宿主菌株,如70%的抗生素来源于放线菌,如以寻找抗菌抗肿瘤活性物质为目标,选择链霉菌为宿主较理想,而筛选新的酶则采用大肠杆菌为宜。

## 4 宏基因组文库的筛选

由于环境样品中微生物种类繁多,宏基因组文库容量一般较大,活性克隆子的筛选是新活性物质筛选的瓶颈,根据研究目的,可从生物活性水平、化合物结构水平以及DNA序列水平设计不同的筛选方案。

**4.1 生物活性水平的筛选** 又称为功能驱动筛选(function-driven screening),根据重组克隆产生的新活性进行筛选。采用各种活性检测手段检测挑选活性克隆子,进而对其深入研究。如在各种选择性平板上通过可见性状筛选到了产生脂酶、淀粉酶、七叶苷水解酶、甘油脱水酶及抗菌活性的克隆子<sup>[3,6,8]</sup>。这一策略以生物活性为线索能够发现全新的活性物质或基因,能够快速鉴别有开发潜力的克隆子,但工作量大,效率低,并且受检测手段的局限。

**4.2 化合物结构水平的筛选** 在一定条件下不同结构的物质在色谱中有不同的峰值,通过比较转入和未转入外源基因的宿主细胞或发酵液抽提物的色谱图筛选产生新结构化合物的克隆子。如Wang(2001)等通过对重组克隆子的细胞抽提物与对照宿主细胞

抽提物的高压液相色谱图谱进行比较,发现了2个产生新结构化合物的克隆,获得了五种全新的小分子抗菌化合物<sup>[2]</sup>。这一策略可直接筛选到新结构化合物,但不一定有生物活性,且工作量大,成本高。

**4.3 DNA 序列水平的筛选** 又称为序列驱动筛选(sequence-driven screening),根据已知相关功能基因的序列设计探针或 PCR 引物,通过杂交或 PCR 扩增筛选阳性克隆子。如 Knietsch (2003) 等根据所有已知脱水酶基因的保守序列设计了一对简并 PCR 引物对宏基因组 DNA 扩增,扩增产物制作探针与宏基因组文库进行杂交,筛到两个高甘油脱水酶和 1,3-丙二醇脱水酶活性的克隆子,有望开发应用于工业生产<sup>[7]</sup>。这一策略有可能筛选到某一类物质中的新分子,而且基于 DNA 的操作有可能利用基因芯片技术大大提高筛选效率,但必须对相关基因序列有一定的了解,较难发现全新的活性物质。

## 5 结语

微生物蕴藏着巨大的基因多样性资源,采用传统的培养途径筛选微生物活性物质虽然有效但受培养技术的限制。宏基因组克隆充分利用了生物学和生物技术手段,使人们有可能挖掘和利用占微生物种类 99% 以上的传统途径触及不到资源。虽然分子克隆技术已足以完成宏基因组文库的构建及筛选,且已筛选到了一些新的生物活性物质,但宏基因在宿主细胞中不表达或表达量低、活性克隆子的频率极低等制约了它的实用性。研究者们正在探索构建新的载体、扩大宿主范围等提高宏基因的表达;通过富集特定类群的基因组提高目标基因在文库中的频率,如在拟采样环境中添加纤维素和稳定同位素<sup>13</sup>C,与纤维素降解有关的微生物代谢活跃,<sup>13</sup>C 就进入到相关的基因组中,提取带<sup>13</sup>C 标记的 DNA 片段建库有望获得较多的纤维素降解活性克隆;开发 DNA 芯片和蛋白质芯片提高筛选效率。相信随着技术的不断进步,宏基因组克隆必将成为微生物活性物质筛选的重要手段并发挥巨大作用。

## 参 考 文 献

- [1] Handelsman J, Rondon M R, Brady S F, *et al.* Chem Biol, 1998, 5: 245 ~ 249.
- [2] Wang G Y, Graziani E, Waters B, *et al.* Org Lett, 2000, 2: 2401 ~ 2404.
- [3] Henne A, Daniel R, Schmitz RA, *et al.* Appl Environ Microbiol, 1999, 65: 3901 ~ 3907.
- [4] Entcheva P, Liebl W, Johann A, *et al.* Appl Environ Microbiol, 2001, 67: 89 ~ 99.
- [5] Voget S, Leggewie C, Uesbeck A, *et al.* Appl Environ Microbiol, 2003, 69: 6235 ~ 6242.
- [6] Knietsch A, Waschowitz T, Bowien S, *et al.* Appl Environ Microbiol, 2003, 69: 1408 ~ 1416.
- [7] Knietsch A, Bowien S, Whited G, *et al.* Appl Environ Microbiol, 2003, 69: 3048 ~ 3060.
- [8] Rondon M R, August P R, Bettermann A D, *et al.* Appl Environ Microbiol, 2000, 66: 2541 ~ 2547.
- [9] Gillespie D E, Brady S F, Bettermann A D, *et al.* Appl Environ Microbiol, 2002, 68: 4301 ~ 4306.
- [10] Brady SF, Chao C J, Clardy J. J Am Chem Soc, 2002, 124: 9968 ~ 9969.
- [11] Courtois S, Cappellano C M, Ball M, *et al.* Appl Environ Microbiol, 2003, 69: 49 ~ 55.
- [12] Cottrell M T, Moore J A, Kirchman D L. Appl Environ Microbiol, 1999, 65: 2553 ~ 2557.
- [13] Diaz-Torres M L, McNab R, Spratt D A, *et al.* Antimicrobial Agents and Chemotherapy, 2003, 47 (4): 1430 ~ 1432.
- [14] 徐平,李文均,徐丽华,等.微生物学通报,2003,4: 84 ~ 85.
- [15] Wren B, Dorrell N. Methods in Microbiology, 2003, 33: 241 ~ 255.