

宏基因组技术在开发未培养环境 微生物基因资源中的应用

李 慧¹, 何晶晶^{1,2}, 张 颖^{1,*}, 徐 慧¹, 陈冠雄¹

(1. 中国科学院沈阳应用生态研究所 陆地生态过程重点实验室, 沈阳 110016; 2. 新疆农业大学, 乌鲁木齐 830052)

摘要:环境微生物宏基因组是一个巨大的基因资源库,但是仅有 0.1% ~ 1% 的微生物在现有技术条件下是可培养的,因此致使未培养微生物基因资源的开发利用受到限制。宏基因组技术直接提取环境样品总 DNA,避开了微生物分离培养的问题,极大扩展了微生物资源的利用空间,增加了获得新生物活性物质的机会。简要介绍了宏基因组的概念及宏基因组克隆技术的基本操作流程和技术要点,重点阐述了目的基因富集、核酸提取、载体和宿主系统选择、宏基因组文库筛选等“瓶颈”技术的研究进展。目的基因富集技术主要包括稳定同位素探针 (SIP)、抑制消减杂交 (SSH) 和差异显示 (DD) 等。基因文库筛选分为序列依赖性筛选和非序列依赖性筛选,其中序列依赖性筛选包括特定基因 PCR、反转录 PCR (RT-PCR)、DNA 微阵、亲和捕获等技术;非序列依赖性筛选主要指基于基因表达活性筛选和基因“陷阱”技术等。此外,介绍了一些近年来通过构建宏基因组文库筛选目的基因的应用实例。

关键词:宏基因组;未培养微生物;核酸提取;宏基因组文库构建;宏基因组文库筛选

文章编号:1000-0933(2008)04-1762-12 中图分类号:Q938.1 文献标识码:A

Application of metagenomic technique in the exploring of uncultured environmental microbial gene resource

LI Hui¹, HE Jing-Jing^{1,2}, ZHANG Ying^{1,*}, XU Hui¹, CHEN Guan-Xiong¹

1 Key Laboratory of Terrestrial Ecological Processes, Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang 110016, China

2 Xinjiang Agricultural University, Wulumuqi 830052, China

Acta Ecologica Sinica, 2008, 28(4): 1762 ~ 1773.

Abstract: Microbial metagenome is the largest gene reservoir in environment. Exploring uncultured microbial gene resource has been restricted, due to the fact that only 0.1%—1% of the microorganisms can be cultivated using current techniques. Metagenome libraries can be constructed by direct extracting DNA from environmental samples to avoid the limitations of culture-dependent methods. Metagenomic technique has greatly enhanced the utilization of microbial resource with more opportunity of obtaining novel bioactive compounds. In this paper, the concept of metagenome, basic operation process and some key technical details are introduced. The introduction is focused on the advancement in some “bottle neck” techniques, including enrichment of interested genes, extraction of nucleic acid, selection of vector and host system, and metagenomic library screening. Gene enrichment involves techniques such as Stable Isotope Probing (SIP), Suppression

基金项目:国家重点基础研究发展计划资助项目(2004CB418505);国家自然科学基金资助项目(30670391)

收稿日期:2006-12-27; **修订日期:**2007-09-17

作者简介:李慧(1977~),女,回族,辽宁锦州人,博士,副研究员,主要从事环境微生物分子生态学研究。E-mail: huili@iae.ac.cn

* 通讯作者 Corresponding author. E-mail: yzhang@iae.ac.cn

Foundation item: The project was financially supported by National Key Basic Research and Development Program of China (No. 2004CB418505); National Natural Sciences Foundation of China (No. 30670391)

Received date: 2006-12-27; **Accepted date:** 2007-09-17

Biography: LI Hui, Ph. D., Associate professor, mainly engaged in molecular ecology for environmental microbiology. E-mail: huili@iae.ac.cn

Subtractive Hybridization (SSH), and Differential Display (DD). Metagenomic library screening includes sequence-dependent and sequence-independent method. The former method includes gene-specific Polymerase Chain Reaction (PCR), reverse transcription PCR (RT-PCR), DNA microarray, affinity capture, and so on. The latter method is mainly referred as activity-based screening and “gene trap” techniques. Some recent examples of gene targeting through metagenome libraries are also introduced.

Key Words: Metagenome; Uncultured microorganisms; Nucleic acid extraction; Metagenomic library construction; Metagenomic library screening

据估计,每克土壤中含有高达 10 000 种不同的微生物,按平均每个基因组大小为 6.1×10^6 bp,每个基因长度为 10^3 bp 计算,则 1g 土壤就包含 6.1×10^7 个基因^[1]。这表明,环境微生物宏基因组是一个巨大的基因资源库。开发环境微生物基因资源较为传统的研究途径是:分离微生物,获得纯培养,基因表达产物活性检测、活性物质的分离纯化直至开发利用。这一途径被证明是十分有效的,也获得了诸多有应用价值的活性物质^[2]。但是,环境中仅有 0.1% ~ 1% 的微生物能够采用现有培养技术进行分离培养^[3],在今天采用传统培养方法已很难发现新的活性物质,极大地限制了未培养微生物基因资源的开发利用。

近年来发展起来的宏基因组克隆技术,直接提取环境样品核酸进行遗传操作,避开了微生物分离培养问题,极大地扩展了微生物资源的利用空间。将重点介绍宏基因组克隆操作过程中的一些“瓶颈”技术,阐明这些技术的基本原理、研究进展、各自的优势和局限性等,并对近年来通过构建宏基因组文库筛选目的基因的应用实例进行了总结。

1 宏基因组概念及筛选活性物质通用流程

“宏基因组(Metagenome)”的概念是 Handelsman 等^[4]于 1998 年提出的,其定义为“生境中全部微小生物遗传物质的总和(the genomes of the total microbiota found in nature)”。宏基因组又常被称做“Collective genome”、“Environmental genome”等,现已普遍采用“Metagenome”的说法,且主要指环境样品中的细菌和真菌的基因组总和。

宏基因组文库既包括了可培养的,又包括了未培养的微生物遗传信息,因此增加了获得新生物活性物质的机会。采用构建宏基因组文库的策略筛选新的基因资源及其表达活性产物的一般流程可以归纳为:样品和基因(组)的富集;提取特定环境中的基因组 DNA 或 mRNA;构建宏基因组 DNA 或 cDNA 文库;筛选目的基因;目的基因活性产物表达(图 1)。

2 环境样品及目的基因/基因组富集

在宏基因组文库筛选过程中,目的基因仅占总 DNA 的一小部分,对环境样品的预富集可以大大提高目的基因的检出几率。预富集技术包括细胞水平富集和基因/基因组水平富集。

2.1 样品和培养物富集

在细胞水平上对宏基因组进行富集已有诸多成功应用的实例。例如,在马尾藻海(Sargasso sea)宏基因组测序项目中,根据细胞大小不同的特性,采用过滤技术有效的去除了真核细胞^[5]。采用不同的离心操作技术对蚜虫巴克纳氏菌(*Buchnera aphidicola*)和海绵共生体(*Cenarchaeum symbiosum*)的共生有机体进行预富集^[6],使其与宿主分开,并对它们的全基因组进行了测序分析。

利用选择培养基对目的微生物进行富集培养,也是有效的富集手段之一,其中最常用的方法是底物选择,此外还包括营养选择以及物理化学指标选择。例如,通过在含有羧甲基纤维素的培养基上富集培养,使基因文库中的纤维素酶基因富集了 4 倍^[7]。但是,由于富集培养选择性地富集了具有快速生长特性的菌群,因此导致大部分物种多样性信息丢失。通过先在严格胁迫条件下短期处理,然后改为较温和的处理条件,可以从一定程度上克服这种方法的局限性。

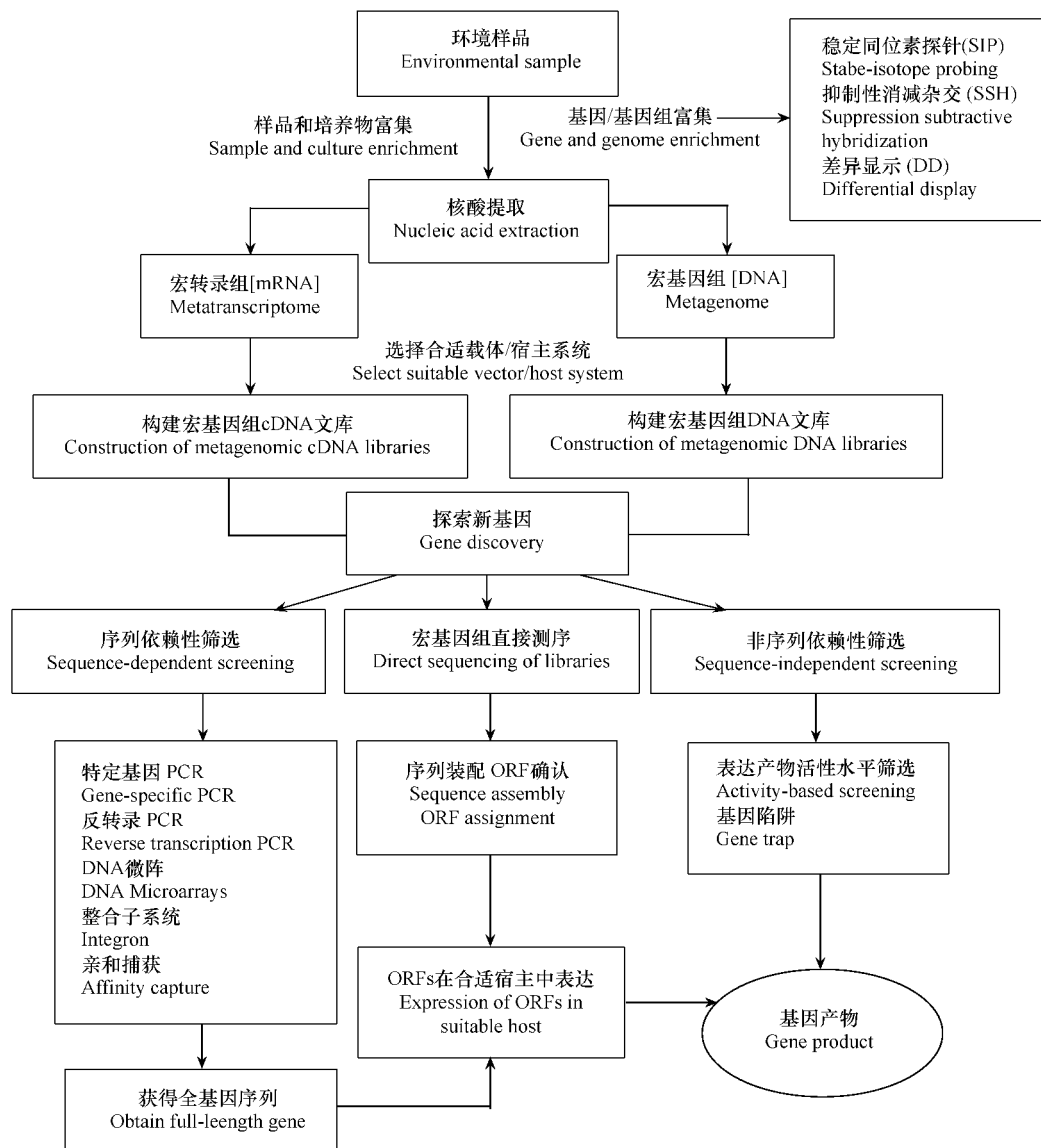


图1 采用构建宏基因组文库的策略筛选新基因资源的通用步骤

Fig. 1 General process of metagenomic strategies for novel genes discovery

2.2 基因组和基因的富集

2.2.1 稳定同位素探针 (Stable- Isotope Probing, SIP) 技术

通过基因组富集可以富集微生物种群中最活跃的部分。SIP 技术采用稳定同位素标记底物,其中的“重”原子掺入到具有代谢活性的微生物核酸中,采用密度梯度离心的方法将“重”的 DNA 或 RNA 与“轻”组分分离,被标记的“重”核酸可以作为 PCR 或者 RT-PCR 的模板,用来构建宏基因组文库^[8]。SIP 技术的主要优势在于可以直接检测微生物群落中具有代谢活性的菌群,而不需要预先知道其特性及在样品中的存在状态。采用特殊培养条件与 SIP 富集策略相结合的手段,从煤焦油污染位点的微生物群落中分离了 1 株新的变形细菌 β -亚纲 (β -proteobacteria) 菌株 CJ2,采用 ^{13}C 标记的蔡对该菌株进行喂饲,结果表明 CJ2 的 16S rRNA 基因在 ^{13}C DNA 部分所占的丰度最大^[9]。SIP 技术另一个较为典型的应用实例,是采用 ^{13}C 标记的甲醇研究森林土壤宏基因组 DNA,结果显示能够鉴定出变形细菌 α -亚纲 (α -proteobacteria) 中所有已知的嗜甲基菌,并在一种嗜酸菌中发现了新的甲醇还原酶基因 (*mxoF*)^[10]。

RNA-SIP 技术即采用标记底物喂饲后,直接提取环境样品的 16S rRNA 或 mRNA,可以不依赖于 DNA 复

制而得到活性菌群的相关信息。RNA-SIP 技术的另一个特点是加入标记的底物后培养时间较短,与 ^{13}C 掺入 DNA 的过程不同, RNA 合成较快并且不依赖于细胞分裂。用 ^{13}C 对酚进行标记,采用 RNA-SIP 技术研究厌氧生物反应器中酚的生物降解菌群,结果表明,酚的主要降解菌是索氏菌(*Thauera*)^[11]。种群中的活跃微生物还可以用 5-溴-尿嘧啶(BrdU)进行标记,标记的 DNA 或 RNA 可以通过免疫捕获或密度梯度离心进行分离^[12]。

SIP 技术本身也存在一些局限性,使其应用受到一定程度的限制。首先,标记底物的量要足够大,以保证具有活性的微生物 DNA 被分离出来。研究表明,从高 G + C mol 含量(35% ~ 70%)的 DNA 中分离“重”的 DNA,至少要有 50% 的 DNA 被 ^{13}C 标记^[13]。其次,如果培养时间较长,会出现交叉喂饲(Cross-feeding)现象,从而导致特异性富集失败^[6]。第三,生长缓慢的物种由于在“重”的组分中所占比例较小,再加上 PCR 扩增的偏倚性,所以很难检测出来。此外, ^{13}C 标记的底物不象 ^{12}C 那样容易被代谢,可能影响菌体的生长。尽管存在诸多影响因素,环境 mRNA-SIP 技术仍然是探索某些新的未知酶的有效方法,这些酶在特殊 C 源生物降解和生物转化中发挥着重要作用^[14]。

2.2.2 抑制性消减杂交(Suppression subtractive hybridization, SSH)技术

抑制性消减杂交是以抑制性 PCR 反应为基础的 cDNA 消减杂交技术,可以选择性扩增差异表达 cDNA 片段,同时抑制非目的 cDNA 的扩增。该方法是一种分离并鉴定组织细胞中选择性表达基因的技术,常被用于区分具有微小遗传差异的 2 个亲缘关系较近的物种,例如用于区分 2 种不同反刍动物瘤胃中复杂 DNA 样品的微小差别^[15]。SSH 技术也可以作为富集宏基因组中特定目的基因的有效技术,若试图鉴定一个参与某环境污染物降解的基因,可以将“胁迫”基因组(即特定污染物胁迫处理)与对照基因组进行比较。但这种方法相对较粗略,仅能区分两个微生物种群总的遗传差异,而不能检测到目的基因或者在污染物胁迫下表达上调的基因。

2.2.3 差异显示(Differential display, DD)技术

DD 技术广泛应用于普通微生物学的研究,但在环境样品基因表达的研究中,还存在一定的问题。因为细菌不同于真核生物,其 mRNA 没有 polyA“尾巴”,很难用 poly(dT)进行反转录。但是,作为一种目的基因富集手段, mRNA 差异显示技术仍然是一种较为实用的方法,通常用于检测某个基因在特定条件诱导下的差异表达^[16]。例如,提取某环境污染物处理前后的细胞 mRNA,进行反转录 PCR,经过表达图谱的比较,就可以检测出上调表达的特定功能基因。目前,已报道了诸多 DD 技术的应用实例。例如, Bowler 等^[17]成功检测了 *Neisseria meningitidis* 在缺铁条件下的上调表达基因; Walters 等^[18]从 *Rhodococcus erythropolis* HL PM-1 中发现 1 个 2, 4-二硝基苯酚降解操纵子; Brzostowicz 等^[19]从混合培养物中鉴别了环己酮单加氧酶基因。此外,应用 DD 技术还从环境样品中发现了 1 个由甲苯诱导产生的新基因^[20]。虽然 DD 技术需要进行大量的 RT-PCR 反应和阳性克隆鉴定工作,但因其不依赖于已知序列信息和传统培养方法,因此仍然是研究环境样品基因表达的一种强有力手段。

2.2.4 其他富集方法

采用噬菌体展示(Phage-display)^[21]、亲和捕获(Affinity capture)^[22]及 DNA 微阵(Microarrays)^[23]等技术也可以对基因(组)进行预富集。噬菌体展示技术可以通过表面展示蛋白与免疫受体的多轮结合,使宏基因组中数量较少的 DNA 序列得到富集。亲和捕获可以用于变性 cDNA 或基因组 DNA 片段的富集,在固定载体上连接一段寡核苷酸片断作为接头,便能够富集用该寡核苷酸片断作为特异性引物扩增出来的 DNA 片段。DNA 微阵技术可以在构建宏基因组之前通过基因标签对特定基因进行预富集;还可以用于宏基因组鸟枪测序前的预筛,以减轻测序任务并减少进行序列分析时不匹配序列的比例。

3 环境样品中核酸的提取

获得高质量的环境样品总 DNA 是宏基因组文库构建的关键之一^[24]。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。目前,已有许多学者对各种环境样品核酸

提取方法进行了综述^[25, 26]。提取方案大致上可以分为 2 类,一类是直接提取法,将环境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、重复性好,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。另一类是间接提取法,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 也不容易抽提出来。本实验室对土壤、废水、污泥等环境样品的总 DNA 提取方法进行了摸索,研究工作表明,直接提取法提取效率较高,不容易丢失物种信息,能够获得 25 kb 左右的 DNA 片段。加入聚乙烯吡咯烷酮(PVPP)能够有效去处腐殖酸类物质的影响,但 DNA 收率略有降低。间接提取法 DNA 提取效率较低,容易丢失物种信息。因此,建议采用直接提取法获得环境样品的总 DNA,以便分析环境样品微生物群落多样性信息^[27]。

环境样品中 RNA 的提取技术与 DNA 提取技术类似,不同的是要尽量减少单链多核苷酸的降解以提高完整 mRNA 的产量^[28]。物理降解和 RNase 酶是导致 mRNA 产量低的主要因素,可以通过低温(-80℃)保存样品、RNA 与蛋白共沉淀、合成 RNA“帽子”等措施减少 mRNA 的降解^[29]。mRNA 提取的相关技术为构建宏基因组 cDNA 文库和进一步开发真核生物功能基因提供了可行性。

4 宏基因组 DNA / cDNA 文库的构建

构建宏基因组文库的关键是要根据具体环境样品的特点和建库目的,选择合适的载体/宿主系统,并采取一些特殊的步骤和对策,如选用多宿主系统、构建穿梭载体等。

4.1 载体的选择

载体选择的原则是有利于目的基因的扩增、表达及在筛选细胞毒类物质时表达量的调控等等。细菌人工染色体(Bacterial artificial chromosome, BAC)和柯斯质粒(Cosmid)是目前构建宏基因组文库常用的载体,前者插入的片段大(可达 350 kb),但克隆效率低,后者插入的片段较小(约为 20~40 kb),但克隆效率较高。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,因此需要尽量插入大片段 DNA 以获得完整的代谢途径。Fosmid 载体的插入片段与 Cosmid 相当,但 Fosmid 插入片段在 *E. coli* 中的克隆效率和稳定性更高^[30]。BAC 和 Cosmid 载体在宿主细胞中的拷贝数低,基因扩大表达存在一定困难,为了提高宏基因的表达,便于重组克隆子活性检测,可以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,适合于筛选单基因或小操纵子产物。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将有利于外源基因表达的提高。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 11 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并调控其表达量^[31]。

此外, λ -噬菌体和其他病毒也可以作为构建宏基因组克隆文库的载体。噬菌体展示库是一种十分有效的高通量筛选技术,该技术通过对表面展示表达产物的亲和选择分离相应的 DNA 序列,能够从宏基因组中富集稀有 DNA 序列^[21],但其局限性在于只能表达分子量 <50kDa 的蛋白。

4.2 宿主的选择

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(如抗菌性)及缺陷型等因素。*E. coli* 是最为常用的宿主,此外,链霉菌和假单胞菌也可以作为构建文库的宿主^[32~34]。研究经验表明,不同微生物种类所产生的活性物质有明显差异,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,而筛选新的酶则选用 *E. coli* 为宜。但应用 *E. coli* 作为宿主具有一定的局限性,仅通过一轮筛选所获得的阳性克隆数极少(<0.01%)。最近的生物信息学研究表明,对于一个插入子 <10 kb 的文库,为寻找一个目的基因,需要筛选

$10^5 \sim 10^6$ 个克隆^[35],也就是说,如果不经预富集,从复杂的宏基因组中筛选目的基因在技术上是十分困难的。因此,宿主系统的进一步探索是更有效的开发宏基因组资源的关键技术之一。

4.3 宏基因组 cDNA 文库(转录库)构建

一直以来,对 cDNA 克隆文库的大规模测序是发现新真核基因的一种快速方法,但由于真核生物基因组存在内含子序列,因此宏基因组表达库技术往往并不适合开发真核基因资源^[36]。此外,在操作上也存在一些需要解决的问题,首先,完整宏基因组 mRNA 的提取十分困难;其次,由于宏基因组 cDNA 文库不包括非表达基因,因此所包含的信息不如基因组文库全面;此外,RT-PCR 扩增限制了插入子的大小,并且使文库具有较强的序列偏倚性^[37]。

5 宏基因组文库测序和拼接

纯培养微生物基因组 DNA 的测序分析技术已比较成熟,为完整宏基因组测序和最终开发新的基因资源奠定了坚实的基础^[37, 38]。近年来,随着自动化高通量测序设备和强大序列分析软件的开发,使完整宏基因组测序具备了技术可行性。但是,完整宏基因组测序任务仍十分繁杂,1g 土壤或者 1L 海水中约含有上千个病毒和原核细胞基因组,上百个低等真核物种和高等真核细胞 DNA^[4, 38],按基因组平均大小保守估计,土壤宏基因组包含 20 ~ 2000Gb 的 DNA 序列。因此,昂贵的费用和繁冗的工作成为限制宏基因组测序技术应用的主要原因。

完整宏基因组测序的首例报道是对酸性矿排水装置生物膜宏基因组 76 Mb 的 DNA 进行序列分析,结果表明,该环境样品生物多样性极低,通过鸟枪测序法拼接了 2 个完整的基因组,鉴别了 4000 多个可能的基因,并初步探索了生物膜群落的代谢途径^[39]。而对马尾藻海(Sargasso sea)宏基因组 >1Gb 的 DNA 序列进行测定则是一个技术上的挑战^[5],共鉴别了约 1.2×10^6 个可能的基因,该研究表明这种技术在开发新的基因资源方面具有巨大的潜力。由于马尾藻海微生物具有高度多样性,序列重叠较差,因此仅拼接出 2 个接近完整的基因组。通过同时构建不同大小插入子文库的方法可以获得丰富的序列信息和大量重叠序列^[30, 40],同时提高全基因组拼接效率。但是,核酸多态性、基因重排、基因复制和水平转移等因素都会影响基因组拼接。由于真核细胞宏基因组较大,且包括内含子和“垃圾”DNA(Junk DNA),使得真核细胞宏基因组的拼接更为复杂,利用宏转录组和构建 cDNA 文库可以从一定程度上解决这一问题。

6 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,活性克隆子的筛选成为开发新活性物质资源的瓶颈。近年来,根据研究目的不同,相继报道了各种宏基因组文库筛选方法。根据是否依赖已知序列信息,大致可以分为两大类,即序列依赖性(Sequence-dependent)筛选和非序列依赖性(Sequence-independent)筛选。

6.1 DNA 序列依赖性筛选

6.1.1 已知功能基因的 PCR 扩增(Gene-specific PCR)

酶活性中心的保守氨基酸序列是检测相似功能蛋白和基因的基础,根据已知相关功能基因的序列设计 PCR 引物,可以通过 PCR 扩增筛选阳性克隆子。例如,利用简并引物从三氯乙烯(TCE)降解菌中发现的多组分酚羟化酶/2-单加氧酶(LmPH)基因可以编码一种新的酶,与已知同种类酶相比具有更高的活性^[41]。但是,该方法存在 2 个主要的缺陷。其一,引物的设计依赖于已有的序列信息,共同进化的 2 个功能相似的基因难于用“族特异性(Family-specific)”引物区分开来,因此很难发现新的功能基因。其二,通过 PCR 扩增功能基因,一般只能获得结构基因的一个片段,而不能获得完整的功能基因。可以将扩增产物进行标记,作为探针筛选宏基因组文库,以获得完整的功能基因。例如,Knietsch 等^[42]根据所有已知脱水酶基因的保守序列设计了一对简并引物,对宏基因组 DNA 进行 PCR 扩增,用扩增产物制作探针与宏基因组文库进行杂交,筛选出两个具有高甘油脱水酶和 1, 3-丙二醇脱水酶活性的克隆子。还可以采用基于 PCR 技术的方法获得扩增产物上游或下游的基因序列,例如,快速步移法(Universal fast walking)^[43]、锅柄 PCR(Panhandle PCR)^[44]、随机引物

PCR(Random primed PCR)^[45]、反向 PCR(Inverse PCR)和接头连接 PCR(Adaptor ligation PCR)^[46]等。应用这些技术从环境样品中发现了一些新的基因,例如 2,5-二酮基-D-葡萄糖酸还原酶基因^[47]。但是,由于宏基因组 DNA 样品十分复杂,使得这些技术在宏基因组水平上的应用存在一定困难,并且操作起来费时费力。因此开发了一种新的替代技术——片匣 PCR(Cassette PCR),应用该技术从酚、原油降解菌群的宏基因组中分离了 2,3-邻苯二酚双加氧酶基因的中心片段^[48]。

6.1.2 反转录 PCR(Reverse transcription PCR, RT-PCR)

在研究功能微生物群落图谱时,由于 mRNA 在菌体代谢过程中更新速度快,因此可以作为比 DNA 更为敏感的生物标记物^[11]。提取环境样品 mRNA,经反转录 PCR(RT-PCR),可以从环境样品中分离相关功能基因。例如,提取煤焦油废弃物的微生物群落 mRNA,经反转录后,再根据萘双加氧酶大亚基基因 *nahAc* 设计引物进行 PCR 扩增,并对该微生物群落中萘双加氧酶的多样性进行了指纹图谱分析^[49],测序结果表明来源于不同微生物的酶具有不同的底物范围和特异活性;再如,根据 II 型芳环羟基化双加氧酶(Ring-hydroxylating dioxygenases, RHDs)的保守序列设计引物,对土壤样品 mRNA 进行 RT-PCR,扩增产物的测序结果表明土壤中存在很多未知的烃类降解菌^[50]。尽管该技术在回收环境样品 mRNA 时存在一定困难,但是 RT-PCR 能够广泛用于筛选包括低等原核细胞和真核细胞的结构基因,并且适用于筛选在环境条件胁迫下发生变化的功能基因。

6.1.3 DNA 微阵技术(DNA microarrays)

DNA 微阵技术是一种高通量的基因分析技术,常用于细菌差异基因表达的研究,但在环境微生物学中的应用才刚刚起步^[23, 51]。目前已建立了 3 种环境样品 DNA 微阵:功能基因微阵(Functional gene arrays, FGAs)、群落基因组微阵(Community genome arrays, CGAs)和系统发育基因组微阵(Phylogenetic genome arrays, PGAs)^[52]。CGAs 是由环境样品或纯培养物的基因组 DNA 制作的,而 PGAs 由 rRNA 基因制作的,两者都不能直接给出环境样品中有关功能基因的详细信息。FGAs 的探针是根据编码功能酶的基因设计的,因此可以检测微生物群落的功能多样性,并对功能酶活性进行定量。FGAs 技术已应用于检测与污染物降解、碳固定、氮循环和硫还原等过程相关的基因序列。目前,最全面的环境样品 DNA 检测微阵包含了数据库中 2402 种降解基因和金属抗性基因,其探针为 50-mer 的寡核苷酸^[53]。通过寡核苷酸设计和杂交条件的优化,可以最大限度的区分相似基因。一般来说,DNA 微阵的质量可以从 3 个方面进行评价:对目标基因的特异性;样品中目标基因的检测限;信号是否与目标基因的数量成正比。

同已知功能基因 PCR 扩增一样,DNA 微阵技术同样受到目的基因已有序列信息的限制。探针设计特异性越强,所筛选基因的功能范围就越窄,通过这种功能相似基因的区分,可以获得更为精确的基因图谱信息。

6.1.4 整合子(Integron)系统

近年来,Integron 被认为是除抗药性质粒和转座子外,细菌抗药基因传播的另一个重要机制,它们在自然界中广泛分布并在细菌基因组进化过程中起着重要的作用^[54]。一个整合子的主要结构包括一个基因片匣(gene cassette)整合位点(*attI*)、一个编码整合酶的基因(*intI*)和两个启动子,这两个启动子分别启动整合酶和插入基因片匣的表达。该结构中的可变区为插入的基因片匣,通常包括一个或多个开放阅读框架(ORFs)和染色体结合位点(*attC*,也称为 59 碱基单元)。整合酶催化外源基因片匣插入结合位点,这个过程是由强启动子控制的位点特异性重组。因此,整合子可以看作是编码基因产物 ORFs 的储备库,为发现新的基因提供了丰富的资源。根据 59 碱基单元保守序列设计引物,成功筛选了与 DNA 糖苷酶、磷酸转移酶、甲基转移酶和巯基转移酶同源的新基因^[55]。还可以通过优化引物设计方案提高目的基因靶定的特异性,即根据 59 碱基单元保守序列设计其中一条引物,再根据目的基因特异性序列设计另一条引物。

6.1.5 亲和捕获(Affinity capture)

根据特定基因片段或者报告基因设计一段寡核苷酸,将寡核苷酸共价连接在固定载体上,便可以对目的基因进行亲和结合。这种方法通常用于回收 cDNA 文库中带 polyA “尾巴”的 RNA^[56]。该技术的局限性在于

杂交速率较慢,因此目前尚未有用于目的基因筛选的报道。

6.2 非 DNA 序列依赖性筛选

6.2.1 基于表达产物活性水平的筛选(Activity-based screening)

基于表达产物活性水平的筛选方法又称为功能驱动筛选(Function-driven screening),不依赖于任何已知序列信息,仅根据文库克隆子产生的酶活性进行筛选。成功筛选宏基因组文库取决于几个重要因素:文库中目的基因的丰度;插入片段大小与基因(或操纵子)长度的关系;表达系统含有本身带强启动子的载体,并能够对外源序列进行融合转录或翻译;插入片段的序列具有核糖体结合位点并与宿主的表达机制相容;诱导剂、分子伴侣、辅助调节因子、翻译后修饰和转录因子等因素的影响;选用缺陷型菌株进行筛选方案优化^[57]。比较理想的情况是,编码目的酶的基因能够在替代宿主中表达,并且其活性可以通过颜色反应等可见性状进行检测。通过这种方法已筛选出了产生脂酶、淀粉酶、七叶苷水解酶、甘油脱水酶及抗菌活性的克隆子^[35, 58, 59]。也可以采用高通量模型,对每个克隆的产物都进行化学结构分析,这种方法显然很耗时费力^[34]。基于生物活性水平的筛选能够发现全新的活性物质或基因,但工作量大、效率低,并且生物转化的产物仅在少数情况下具有可见性状,酶活性的检测受到研究方法的限制。

6.2.2 基因陷阱(Gene trap)技术

由于受到酶活检测方法的限制,大多数宝贵的酶资源不能通过上述基于活性的方法发掘出来,许多学者开始关注一些新的筛选方法,其中对基因的遗传操作是一种更为有效的手段。以下总结了几种分子陷阱技术,使编码特定生物转化过程的基因被“选择”出来,而不是被“筛选”出来。方法一,选择能利用产物但不能利用底物的宿主包装宏基因组文库,以底物为唯一碳(氮或硫)源筛选文库,就可以选择出携带编码特定生物转化酶 DNA 片段的克隆。方法二的原理是,细菌代谢某特定营养物质的操纵子通常与该营养物质诱导的启动子和同源转录调控子毗邻^[60]。将环境 DNA 随机克隆到无启动子的绿色荧光蛋白基因(Green fluorescent protein, GFP)前面,然后通过荧光激活细胞分离(Fluorescence activated cell sorting, FACS)技术,在添加特定底物的条件下对表达库进行富集,就能够选择出对该底物具有代谢活性的克隆^[61]。但这一技术也存在一些问题,第一,代谢特定底物的调控子和操纵子在位置上不一定都相邻;第二,某些代谢基因的启动子可能根本不接受调控进而表达该基因;最大的局限性在于转录调控子可能被不是底物的另一种效应物激活。第三种遗传操作方法是转录陷阱技术,其原理为转录调控子与反应产物专一结合而不识别底物,该调控子能够顺序激活与报告基因融合的启动子,报告基因的表达产物可进行表型筛选。选择带有这种调控子/启动子/报告基因系统的宿主包装宏基因组文库,便可以用含有底物的选择培养基筛选携带编码特定反应 DNA 片段的阳性克隆^[62]。上述基因陷阱方法的关键技术在于建立高度耐受表达系统,以表达任意来源的编码蛋白和酶,目前“超相容性”表达载体/宿主系统的研究已取得了一定进展。

7 宏基因组技术在筛选生物活性物质研究中的应用

自从 1991 年 Pace 等首次构建了海洋微小浮游生物环境 DNA 文库以来,目前已构建了土壤、海洋、人唾液、堆肥样品等环境样品的宏基因组文库(表 1)。所采用的载体种类十分广泛,包括 Cosmid、Fosmid、BAC、λ-

表 1 已构建的宏基因组文库特征及其目的基因筛选应用实例

Table 1 Characteristics of metagenomic libraries and application examples for gene targeting

环境样品 Environmental sample	目的基因 Target gene	宿主 Host	载体 Vector	克隆子数 Number of clones	插入片段大小(kb) Insert size	参考文献 References
农田土壤 Agricultural soil	核酸酶、淀粉酶、抗菌活性、脂肪酶 Nuclease, Amylase, Antibacterial, Lipase	大肠杆菌 DH10B <i>E. coli</i> DH10B	pBeloBAC 11	24 576	44.5	[58]
马粪、各种土壤 Horse excrement, various soil	生物素合成 Biotin biosynthesis	大肠杆菌 VCS257 <i>E. coli</i> VCS257	pWE15	35 000	30 ~ 40	[63]

续表

环境样品 Environmental sample	目的基因 Target gene	宿主 Host	载体 Vector	克隆子数 Number of clones	插入片段大小 (kb) Insert size	参考文献 References
甜菜地、底泥样品 Sugar beet field、sediment	乙醇氧化还原酶 Alcohol oxidoreductase	大肠杆菌 DH5 α <i>E. coli</i> DH5 α	pSK +	400 000	3.0 ~ 5.6	[59]
撂荒土壤 Unplanted field	淀粉酶、琼脂糖酶、酰胺酶、纤维素酶、葡聚糖分支酶 Amylase、Agarase、Aminidase、Cellulases、Glucan branching enzyme	大肠杆菌 VCS257 <i>E. coli</i> VCS257	pWE15	1 532	25 ~ 40	[64, 65]
草场、河谷、甜菜地 Meadow、River valley、Sugar beet field	4-羟基丁酸代谢酶系 4-hydroxybutyrate metabolism system	大肠杆菌 DH5 α <i>E. coli</i> DH5 α	pSK +	930 000	5 ~ 8	[35]
地热沉积物 Geothermal sediments	亚铁血红素合成酶、磷酸二脂酶 Heme biosynthesis、Phosphodiesterase	大肠杆菌 TOPO10 <i>E. coli</i> TOPO10	pCR-TOPO-XL	37 000	1 ~ 10	[66]
耕作土壤 Arable soil	聚酮体合成酶 Polyketide synthase	变铅青链霉菌 <i>Streptomyces lividans</i>	大肠杆菌-变铅青链霉菌穿梭粘粒 <i>E. coli</i> - <i>S. lividans</i> shuttle cosmid	5000	50	[32]
人唾液 Human saliva	四环素抗性 Tetracycline Resistance	大肠杆菌 TOPO10 <i>E. coli</i> TOPO10	TOPO-XL	450	0.8 ~ 3.0	[67]
沿海海水 Coastal seawater	几丁质酶 Chitinase	Gigapack III	λ -Zap II	825 000	2 ~ 10	[68]
甜菜地、湖水、河水 Sugar beet field、River、Lake	甘油脱水酶、二醇脱水酶 Glycerol dehydratase、Diol dehydratase	大肠杆菌 DH5 α <i>E. coli</i> DH5 α	pSK +	560 000	3.5 ~ 5.0	[42]
土壤 Soil	抗菌和抗真菌活性 antibacterial and antifungal	变铅青链霉菌、恶臭假单胞菌 <i>Streptomyces lividans</i> 、 <i>Pseudomonas putida</i>	BAC 穿梭载体 BAC shuttle vectors		85 ~ 120	[33]
泥地、海岸、森林 Mud flat、Beach、Forest	酯酶 Esterase	大肠杆菌 <i>E. coli</i>	Fosmid	60 000	30 ~ 40	[69]
池塘水 Pond water	酯酶 Esterase	大肠杆菌 DH10B <i>E. coli</i> DH10B	pUC19	30 000	2 ~ 12	[70]
溪水 Stream water	淀粉酶 Amylase	大肠杆菌 DH5 α <i>E. coli</i> DH5 α	pUC19	30 000	3 ~ 7	[71]
活性污泥、土壤 Activated sludge、soil	聚-3-羟基丁酸代谢酶系 Poly-3-hydroxybutyrate metabolism system	大肠杆菌 HB101 <i>E. coli</i> HB101	pRK7813	45 630	25 ~ 45	[72]
堆肥样品 Compost	解脂酶、淀粉酶、磷酸酶、双加氧酶 lipolytic enzyme、amylase、phosphatase、dioxygenase	大肠杆菌 DH5 α <i>E. coli</i> DH5 α	pJOE930	560 000	3 ~ 8	[73]
堆肥样品 Compost	木聚糖酶 Xylanase	大肠杆菌 EPI100 <i>E. coli</i> EPI100	pWEB:TNC	50 000	35	[74]
牛瘤胃 Bovine rumen	β -葡萄糖苷酶 β -glucosidase	大肠杆菌 EPI100 <i>E. coli</i> EPI100	pWEB:TNC	12 000	31.5 ~ 45.5	[75]

噬菌体以及各种穿梭载体,所采用的宿主系统为常用的大肠杆菌、链霉菌和假单胞菌。尽管这些文库所

采用的载体/宿主系统不同,但实验方案和技术操作基本一致。通过对文库的筛选,发现了许多新的基因及其编码的活性物质,涉及各种酶类、抗菌抗肿瘤活性物质和色素等一些小分子物质。

8 结语

宏基因组克隆充分运用前沿分子生物学技术,不依赖于传统培养方法发掘丰富的微生物资源,同时也是一种基因资源的保存形式。但是,正如文中所述,宏基因组技术还存在诸多问题亟待解决,因此现在说“宏基因组时代”已经到来可能为时过早。不管怎样,随着研究方法的不断进步和创新,宏基因组技术在微生物资源的开发利用上无疑能够发挥巨大的潜力。我国地域辽阔,蕴藏着丰富的微生物资源,但是应用宏基因组技术开发环境微生物资源的研究才刚刚起步^[74, 75],许多新的技术方法还有待建立和不断完善。

References:

- [1] Streit W R, Daniel R, Jaeger K E. Prospecting for biocatalysts and drugs in the genomes of non-cultured microorganisms. *Curr. Opin. Biotechnol.*, 2004, 15: 285 — 290.
- [2] Li Y Z, Chen Q. The sources of marine microorganisms and their bioactive metabolites. *China Biotechnology*, 2000, 20(5): 28 — 31.
- [3] Ellis R J, Morgan P, Weightman A J, *et al.* Cultivation-dependent and -independent approaches for determining bacterial diversity in heavy-metal-contaminated soil. *Appl. Environ. Microbiol.*, 2003, 69: 3223 — 3230.
- [4] Handelsman Jo, Rondon M R, Brady S F, *et al.* Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol.*, 1998, 5(10): 245 — 249.
- [5] Venter J C, Remington K, Heidelberg J F, *et al.* Environmental genome shotgun sequencing of the Sargasso Sea. *Science*, 2004, 304: 66 — 74.
- [6] Schloss P D, Handelsman J. Biotechnological prospects from metagenomics. *Curr. Opin. Biotechnol.*, 2003, 14: 303 — 310.
- [7] Rees H C, Grant W D, Jones B E, *et al.* Diversity of Kenyan soda lake alkaliphiles assessed by molecular methods. *Extremophiles*, 2004, 8: 63 — 71.
- [8] Radajewski S, Ineson P, Parekh N R, *et al.* Stable-isotope probing as a tool in microbial ecology. *Nature*, 2000, 403: 646 — 649.
- [9] Jeon C O, Park W, Padmanabhan P, *et al.* Discovery of a bacterium, with distinctive dioxygenase, that is responsible for in situ biodegradation in contaminated sediment. *Proc. Natl. Acad. Sci. USA*, 2003, 100: 13591 — 13596.
- [10] Radajewski S, Webster G, Reay D S, *et al.* Identification of active methylophilic populations in an acidic forest soil by stable-isotope probing. *Microbiology*, 2002, 148: 2331 — 2342.
- [11] Manefield M, Whiteley A S, Griffiths R I, *et al.* RNA stable isotope probing, a novel means of linking microbial community function to phylogeny. *Appl. Environ. Microbiol.*, 2002, 68: 5367 — 5373.
- [12] Urbach E, Vergin K L, Giovannoni S J. Immunochemical detection and isolation of DNA from metabolically active bacteria. *Appl. Environ. Microbiol.*, 1999, 65: 1207 — 1213.
- [13] Radajewski S, McDonald I R, Murrell J C. Stable-isotope probing of nucleic acids: a window to the function of uncultured microorganisms. *Curr. Opin. Biotechnol.*, 2003, 14: 296 — 302.
- [14] Wackett L P. Stable isotope probing in biodegradation research. *Trends Biotechnol.*, 2004, 22: 153 — 154.
- [15] Galbraith E A, Antonopoulos D A, White B A. Suppressive subtractive hybridisation as a tool for identifying genetic diversity in an environmental metagenome: the rumen as a model. *Environ. Microbiol.*, 2004, 6: 928 — 937.
- [16] Liang P. A decade of differential display. *Biotechniques*, 2002, 33: 338 — 346.
- [17] Bowler L D, Hubank M, Spratt B G. Representational difference analysis of cDNA for the detection of differential gene expression in bacteria: development using a model of iron-regulated gene expression in *Neisseria meningitidis*. *Microbiology*, 1999, 145: 3529 — 3537.
- [18] Walters D M, Russ R, Knackmuss H, *et al.* High-density sampling of a bacterial operon using mRNA differential display. *Gene*, 2001, 273: 305 — 315.
- [19] Brzostowicz P C, Walters D M, Thomas S M, *et al.* mRNA differential display in a microbial enrichment culture: simultaneous identification of three cyclohexanone monooxygenases from three species. *Appl. Environ. Microbiol.*, 2003, 69: 334 — 342.
- [20] Fleming J T, Yao W H, Saylor G S. Optimization of differential display of prokaryotic mRNA: application to pure culture and soil microcosms. *Appl. Environ. Microbiol.*, 1998, 64: 3698 — 3706.
- [21] Cramer R, Suter M. Display of biologically active proteins on the surface of filamentous phages: a cDNA cloning system for the selection of functional gene products linked to the genetic information responsible for their production. *Gene*, 1993, 137: 69 — 75.
- [22] Demidov V V, Bukanov N O, Frank-Kamenetskii M D. Duplex DNA capture. *Curr. Issues Mol. Biol.*, 2000, 2: 31 — 35.
- [23] Wu L, Thompson D K, Li G S, *et al.* Development and evaluation of functional gene arrays for the election of selected genes in the environment. *Appl. Environ. Microbiol.*, 2001, 67: 5780 — 5790.
- [24] Wilkinson D E, Jeanicke T, Cowan D A. Efficient molecular cloning of environmental DNA from geothermal sediments. *Biotechnol. Lett.*, 2002,

- 24: 155 — 161.
- [25] Krsek M, Wellington E M H. Comparison of different methods for the isolation and purification of total community DNA from soil. *J. Microbiol. Methods*, 1999, 39: 1 — 16.
- [26] Schneegurt M A, Dore S Y, Kulpa C F Jr. Direct extraction of DNA from soils for studies in microbial ecology. *Curr. Issues Mol. Biol.*, 2003, 5: 1 — 8.
- [27] Li H, Zhang Y, Zhang C G, *et al.* Effect of petroleum-containing wastewater irrigation on bacterial diversities and enzymatic activities in paddy soil irrigation area. *Journal of Environmental Quality*, 2005, 34: 1073 — 1080.
- [28] Alm E W, Zheng D, Raskin L. The presence of humic substances and DNA in RNA extracts affects hybridization results. *Appl. Environ. Microbiol.*, 2000, 66: 4547 — 4554.
- [29] Edery I, Chu L L, Sonenberg N, *et al.* An efficient strategy to isolate full-length cDNAs based on an mRNA cap retention procedure (CAPture). *Mol. Cell. Biol.*, 1995, 15: 3363 — 3371.
- [30] Beja O. To BAC or not to BAC: marine ecogenomics. *Curr. Opin. Biotechnol.*, 2004, 15: 187 — 190.
- [31] Handelsman Jo, Liles M, Mann D, *et al.* Cloning the metagenome: culture-independent access to the diversity and functions of the uncultivated microbial world. *Methods in Microbiology*, 2002, 33: 241 — 255.
- [32] Courtois S, Cappellano C M, Ball M, *et al.* Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. *Appl. Environ. Microbiol.*, 2003, 69: 49 — 55.
- [33] Martinez A, Kolvek S J, Yip C L T, *et al.* Genetically modified bacterial strains and novel bacterial artificial chromosome shuttle vectors for constructing environmental libraries and detecting heterologous natural products in multiple expression hosts. *Appl. Environ. Microbiol.*, 2004, 70: 2452 — 2463.
- [34] Wang G Y, Graziani E, Waters B, *et al.* Novel natural products from soil DNA libraries in a streptomycete host. *Org. Lett.*, 2000, 2: 2401 — 2404.
- [35] Henne A, Daniel R, Schmitzet R A, *et al.* Construction of environmental DNA libraries in *Escherichia coli* and screening for the presence of genes conferring utilization of 4-hydroxybutyrate. *Appl. Environ. Microbiol.*, 1999, 65: 3901 — 3907.
- [36] Starkey M P, Umrana Y, Mundy C R, *et al.* Reference cDNA library facilities available from European sources. *Mol. Biotechnol.*, 1998, 9: 35 — 57.
- [37] B  j   O, Suzuki M T, Koonin E V, *et al.* Construction and analysis of bacterial artificial chromosome libraries from a marine microbial assemblage. *Environ. Microbiol.*, 2000, 2: 516 — 529.
- [38] Torsvik V, Ovreas L. Microbial diversity and function in soil: from genes to ecosystems. *Curr. Opin. Microbiol.*, 2002, 5: 240 — 245.
- [39] Tyson G W, Chapman J, Hugenholtz P, *et al.* Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature*, 2004, 428: 37 — 43.
- [40] Schmidt T M, Long D E F, Pace N R. Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. *J. Bacteriol.*, 1991, 173: 4371 — 4378.
- [41] Futamata H, Harayama S, Watanabe K. Group-specific monitoring of phenol hydroxylase genes for a functional assessment of phenol-stimulated trichloroethylene bioremediation. *Appl. Environ. Microbiol.*, 2001, 67: 4671 — 4677.
- [42] Knietzsch A, Bowien S, Whited G, *et al.* Identification and characterization of coenzyme B12-dependent glycerol dehydratase- and diol dehydratase-encoding genes from metagenomic DNA libraries derived from enrichment cultures. *Appl. Environ. Microbiol.*, 2003, 69: 3048 — 3060.
- [43] Myrick K V, Gelbart W M. Universal fast walking for direct and versatile determination of flanking sequence. *Gene*, 2002, 284: 125 — 131.
- [44] Megonigal M D, Rappaport E F, Wilsonet R B, *et al.* Panhandle PCR for cDNA: a rapid method for isolation of MLL fusion transcripts involving unknown partner genes. *Proc. Natl. Acad. Sci. USA*, 2000, 97: 9597 — 9602.
- [45] Liu Y, Whittier R F. Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from PI and YAC clones for chromosome walking. *Genomics*, 1995, 25: 674 — 681.
- [46] Ochman H, Ayala F J, Hartl D L. Use of polymerase chain reaction to amplify segments outside boundaries of known sequences. *Methods Enzymol.*, 1993, 218: 309 — 321.
- [47] Eschenfeldt W H, Stols L, Rosenbaumet H, *et al.* DNA from uncultured organisms as a source of 2,5-diketo-D-gluconic acid reductases. *Appl. Environ. Microbiol.*, 2001, 67: 4206 — 4214.
- [48] Okuta A, Ohnishi K, Harayama S. PCR isolation of catechol 2,3-dioxygenase gene fragments from environmental samples and their assembly into functional genes. *Gene*, 1998, 212: 221 — 228.
- [49] Wilson M S, Bakermans C, Madsen E L. In situ, real-time catabolic gene expression: extraction and characterization of naphthalene dioxygenase mRNA transcripts from groundwater. *Appl. Environ. Microbiol.*, 1999, 65: 80 — 87.
- [50] Taylor P M, Medd J M, Schoenborn L, *et al.* Detection of known and novel genes encoding aromatic ring-hydroxylating dioxygenases in soils and in aromatic hydrocarbon-degrading bacteria. *FEMS Microbiol. Lett.*, 2002, 216: 61 — 66.
- [51] Sebat J L, Colwell F S, Crawford R L. Metagenomic profiling: microarray analysis of a metagenomic DNA library. *Appl. Environ. Microbiol.*,

- 2003, 69: 4927—4934.
- [52] Zhou J. Microarrays for bacterial detection and microbial community analysis. *Curr. Opin. Microbiol.*, 2003, 6: 288—294.
- [53] Rhee S K, Liu X, Wu L, *et al.* Detection of genes involved in biodegradation and biotransformation in microbial communities by using 50-mer oligonucleotide microarrays. *Appl. Environ. Microbiol.*, 2004, 70: 4303—4317.
- [54] Rowe-Magnus D A, Mazel D. Integrons: natural tools for bacterial genome evolution. *Curr. Opin. Microbiol.*, 2001, 4: 565—569.
- [55] Stokes H W, Holmes A J, Nield B S, *et al.* Gene cassette PCR: sequence-independent recovery of entire genes from environmental DNA. *Appl. Environ. Microbiol.*, 2001, 67: 5240—5246.
- [56] Stull D, Pisano J M. Purely RNA: New innovations enhance the quality, speed, and efficiency of RNA isolation techniques. *Scientist*, 2001, 15: 29—31.
- [57] Galvão T C, Mohn W W, de Lorenzo V. Exploring the microbial biodegradation and biotransformation gene pool. *Trends Biotechnol.*, 2005, 23: 497—506.
- [58] Rondon M R, August P R, Bettermann A D, *et al.* Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl. Environ. Microbiol.*, 2000, 66: 2541—2547.
- [59] Knietzsch A, Waschowitz T, Bowien S, *et al.* Construction and screening of metagenomic libraries derived from enrichment cultures: generation of a gene bank for genes conferring alcohol oxidoreductase activity on *Escherichia coli*. *Appl. Environ. Microbiol.*, 2003, 69: 1408—1416.
- [60] Warren P B, ten Wolde P R. Statistical analysis of the spatial distribution of operons in the transcriptional regulation network of *Escherichia coli*. *J. Mol. Biol.*, 2004, 342: 1379—1390.
- [61] Uchiyama T, Abe T, Ikemura T, *et al.* Substrate-induced gene-expression screening of environmental metagenome libraries for isolation of catabolic genes. *Nat. Biotechnol.*, 2005, 23: 88—93.
- [62] Mohn W W, Garmendia J, Galvão T C, *et al.* Surveying biotransformations with à la carte genetic traps: translating dehydrochlorination of lindane (γ -hexachlorocyclohexane) into *lacZ*-based phenotypes. *Environmental Microbiology*, 2006, 8: 546—555.
- [63] Entcheva P, Liebl W, Johann A, *et al.* Direct cloning from enrichment cultures, a reliable strategy for isolation of complete operons and genes from microbial consortia. *Appl. Environ. Microbiol.*, 2001, 67: 89—99.
- [64] Voget S, Leggewie C, Uesbeck A, *et al.* Prospecting for novel biocatalysts in a soil metagenome. *Appl. Environ. Microbiol.*, 2003, 69: 6235—6242.
- [65] Voget S, Steele H L, Streit W R. Characterization of a metagenome-derived halotolerant cellulose. *Journal of Biotechnology*, 2006, 126: 26—36.
- [66] Wilkinson D E, Jeanicke T, Cowan D A. Efficient molecular cloning of environmental DNA from geothermal sediments. *Biotechnol. Lett.*, 24: 155—161.
- [67] Diaz-Torres M L, McNab R, Spratt D A, *et al.* Novel tetracycline resistance determinant from the oral metagenome. *Antimicrob. Agents Chemother.*, 2003, 47: 1430—1432.
- [68] Cottrell M T, Moore T A, Kirchman D L. Chitinases from uncultured marine microorganisms. *Appl. Environ. Microbiol.*, 1999, 65: 2553—2557.
- [69] Kim Y J, Choi G S, Kim S B, *et al.* Screening and characterization of a novel esterase from a metagenomic library. *Protein Expression and Purification*, 2006, 45: 315—323.
- [70] Ranjan R, Grover A, Kapardar R K, *et al.* Isolation of novel lipolytic genes from uncultured bacteria of pond water. *Biochem. Biophys. Res. Commun.*, 2005, 335: 57—65.
- [71] Yun J, Kang S, Park S, *et al.* Characterization of a novel amylolytic enzyme encoded by a gene from a soil-derived metagenomic library. *Appl. Environ. Microbiol.*, 2004, 70: 7229—7235.
- [72] Wang C, Meek D J, Panchal P, *et al.* Isolation of poly-3-hydroxybutyrate metabolism genes from complex microbial communities by phenotypic complementation of bacterial mutants. *Appl. Environ. Microbiol.*, 2006, 72: 384—391.
- [73] Lämmle K, Zipper H, Breuer M, *et al.* Identification of novel enzymes with different hydrolytic activities by metagenome expression cloning, *Journal of Biotechnology*, 2007, 127: 575—592.
- [74] Zhang P, Duan C G, Pang H, *et al.* The construction of a metagenomic library of uncultured bacteria from compost and the cloning and identification of novel xylanase genes. *Guangxi Sciences*, 2005, 12: 343—346.
- [75] Zhao G C, Duan C G, Pang H, *et al.* Cloning and identification a β -glucosidase gene *umbgl 3A* of uncultured bacteria from bovine rumen. *Southwest China Journal of Agricultural Sciences*, 2005, 18(4): 472—476.

参考文献:

- [2] 李越中, 陈琦. 海洋微生物资源及其产生生物活性代谢产物的研究. *中国生物工程杂志(生物工程进展)*, 2000, 20: 28~31.
- [74] 张鹏, 段承杰, 庞浩, 等. 堆肥未培养细菌的宏基因组文库构建及新的木聚糖酶基因的克隆和鉴定. *广西科学*, 2005, 12: 343~346.
- [75] 赵广存, 段承杰, 庞浩, 等. 牛瘤胃未培养细菌中一个 β -葡萄糖苷酶基因 *umbgl3A* 的克隆及鉴定. *西南农业学报*, 2005, 18: 472~476.