

环境宏基因组学技术的主要瓶颈及发展

蒋云霞¹, 艾春香^{2,3*}

(11 厦门大学生命科学学院应用与环境微生物研究所, 厦门 361005; 21 厦门大学海洋与环境学院, 厦门 361005; 31 厦门大学近海海洋环境科学国家重点实验室, 厦门 361005)

摘要: 从环境宏基因组文库的构建及筛选两方面分析了环境宏基因组学技术所存的主要瓶颈, 综述了针对主要瓶颈而发展的最新环境宏基因组学技术, 展望了环境宏基因组学技术在微生物生态学基础研究及生物技术应用研究领域的发展方向。

关键词: 宏基因组; 未培养微生物; 序列筛选; 功能筛选

中图分类号: X17 文献标识码: A 文章编号: 0250-3301(2007)12-2286-06

Main Bottleneck and Developments of Metagenomic Technology

JIANG Yunxia¹, AI Chunxiang^{2,3}

(11 Institute of Applied and Environmental Microbiology, School of Life Science, Xiamen University, Xiamen 361005, China; 21 College of Oceanography & Environmental Sciences, Xiamen University, Xiamen 361005, China; 31 State Key Laboratory of Marine Environmental Science, Xiamen University, Xiamen 361005, China)

Abstract: This review analyzed the main bottleneck of metagenomic technology on construction and screening of environmental library, discussed the recent developments towards overcoming the main bottleneck. Finally, the enormous scope and potential for both fundamental microbial ecology and biotechnological development of metagenomics was highlighted.

Key words: metagenome; uncultured microorganisms; sequence-driven approach; function-driven approach

环境宏基因组学技术的应用拓展了微生物学的研究思路与方法, 为不同自然生境中的未培养微生物研究提供了有力手段。环境宏基因组学研究是当前国际上微生物学研究的前沿领域和热点。

从 Handelsman 等^[1] 正式提出“宏基因组”的概念至今, 环境宏基因组学技术在微生物生态学、生物技术领域均取得了可观的成绩^[2-6]。然而, 作为新型技术, 尚存在若干不足, 需要不断地发展与完善。本文针对环境宏基因组学的关键性技术, 综述了其近期所取得的最新进展。

1 环境宏基因组学基本技术

环境宏基因组学研究的基本思路是直接提取环境中所有微生物的基因组 DNA, 克隆到合适的载体, 通过构建宏基因组文库, 将环境中全部微生物的核酸信息收集在一起, 运用序列筛选或功能筛选方式从文库中获得有用的酶、抗生素等生物活性物质及相关基因^[7]。其前端关键性技术是环境 DNA (environmental DNA, eDNA) 的提取, eDNA 提取方法的精确程度将直接决定文库中微生物信息的精确度; 下游关键性技术是文库的分析与筛选技术。

1.1 环境 DNA 的提取技术

eDNA 的提取方法主要有 2 种, 一种是从环境样品中分离微生物细胞, 再提取 DNA, 称为间接提

取法; 另一种是直接提取环境样品中的微生物 DNA, 称为直接提取法, 也称原位裂解法。间接提取法所得的 eDNA 纯度较高, 片段较大, 但产量及所包含基因组信息的广泛性不及原位裂解法^[7]。迄今为止, 有关 eDNA 的提取与纯化方法的报道较多^[8-11], 然而没有一种提取方法能够适用于所有的环境样品, 应根据研究生境特性与目的, 进行选择和优化。

1.1.2 环境宏基因组文库筛选技术

环境宏基因组文库的筛选策略主要分为序列筛选和功能筛选 2 种。序列筛选策略是先从文库中获得目的基因, 继而对之异源表达, 得到具有生物活性的产物; 功能筛选策略则是先获得具有生物活性的阳性克隆子, 再通过插入片段测序, 得到相应的基因结构。2 种分析策略对于充分利用宏基因组文库资源都是必要的^[12]。

许多学者将微生物 16S rRNA 及 18S rRNA 的 PCR 文库研究归入环境宏基因组学范畴, 但是基于小核糖体基因的分析只能得到环境中微生物群落的组成信息, 并不能阐释其相应微生物的生理功能信

收稿日期: 20070223; 修订日期: 20070815

基金项目: 国家高技术研究发展计划(863)项目(2006AA10A406); 江苏省滩涂生物资源和环境保护实验室项目(JLCBE05007)

作者简介: 蒋云霞(1973~), 女, 博士研究生, 主要研究方向为环境微生物学, E-mail: jiangyuxia@yahoo.com.cn

* 通讯联系人, E-mail: chunxi@sina.com

息. 因此, 此技术领域不在本综述范围内.

2 环境宏基因组学技术的主要瓶颈

2.1 环境宏基因组文库前端技术的主要瓶颈

eDNA 提取技术尚存在若干瓶颈.¹ 从 DNA 片段大小分析, 高分子量 eDNA 的获得是从文库中捕捉编码药物等物质完整基因簇的先决条件. 土壤环境中, 由于微生物与土壤颗粒紧密结合的特性以及腐殖酸等抑制性物质存在等原因, 从中难以获得适于构建宏基因组文库的高分子量 eDNA^[13]. Bertrand 等^[14] 采用间接提取法, 通过 Nycodenz 梯度离心, 所回收的土壤 eDNA 片段大小已能达到 400 kbp, 但至今基于原位裂解获得 > 100 kbp 土壤 eDNA 的提取技术尚未突破, 运用原位裂解法构建更大片段环境宏基因组文库(现有的土壤宏基因组文库中, 平均插入片段最大为 4415 kbp^[15]) 仍是一个难点;^o 从 DNA 所包含信息的广泛度分析, eDNA 提取技术已经具备区别提取样品中微生物胞外与胞内 DNA^[16]、活细胞与死细胞 DNA 的能力^[17], 然而, 运用现有 DNA 提取技术到底能回收环境中多少类型微生物的核酸信息仍不清楚^[18]. 选用温和的提取方法, 环境中大部分革兰氏阴性细菌能被裂解, 而革兰氏阳性细菌几乎不能被裂解; 采用剧烈的提取方法, 有可能获得环境中革兰氏阳性细菌的信息, 但所得的 eDNA 又将被严重剪切, 无法满足构建大片段环境宏基因组文库的要求^[19]. 不可避免地, 环境宏基因组文库所包含微生物基因组信息的偏差将直接导致/基因遗漏现象发生. 如海洋中普遍存在的微生物固氮基因, 却在测序量高达 11.6 Gbp 的马尾藻海水宏基因组文库^[20] 中被遗漏, 表明仅仅运用宏基因组学技术同样会忽略部分的微生物资源^[21].

2.2 环境宏基因组文库下游技术的主要瓶颈

阳性克隆筛选频率低下是宏基因组学下游技术的主要瓶颈. 运用经典的功能筛选方式, 往往是在数千个, 甚至数百万个重组克隆子中才能检测到有用的活性克隆^[22], 造成此局面一个重要的原因是外源基因的异源表达水平低下. 目前根据核酸序列相似性及基因保守区设计探针或引物的杂交、PCR 筛选方法, 从文库中发现新基因的比率尚不到已知基因的 40%^[12].

2.3 环境宏基因组文库适用范围瓶颈

真菌是很多工业化生产酶制剂的来源菌, 未培养真菌资源的开发将为生物技术领域提供广阔的发展空间. 然而, 由于真核基因存在大量内含子, 现行

的环境宏基因组学技术尚不能应用于真核微生物资源的/生物探矿^o.

3 环境宏基因组学技术的精细化发展

近几年, 针对环境宏基因组学技术的不足, 各种生物信息学方法、新型敏感筛选技术以及与各种先进方法的联盟, 已成功地运用到环境宏基因组学的相关研究中, 环境宏基因组学技术已经步入精细化发展阶段.

3.1 环境宏基因组文库构建技术的改进

3.1.1 建立新的载体2宿主系统

宿主选择范围的扩大及克隆载体的改良对于提高宏基因组克隆子的异源表达水平非常关键. 可以推测, 高 GC 含量 eDNA 片段中的生物合成基因簇有可能在链霉菌类的宿主中才能实现有效地异源表达^[23]. 因此, 除 *Escherichia coli* 之外, 应用 *Pseudomonas putida* 及 *Streptomyces lividans* 和 *Bacillus subtilis* 等革兰氏阳性宿主菌可显著提高异源表达的成功性^[24]. 新型穿梭 BAC^[25]、穿梭 Cosmid^[25]、广宿主 Cosmid^[26, 27] 等载体的发明使得不同宿主的平行使用成为可能. 经序列鉴定后的目的 DNA, 可通过穿梭载体从第 1 个贮存宿主转移到表达宿主中, 最终产生相关活性物质. Wang 等^[25] 首次运用穿梭 Cosmid, 以链霉菌为宿主, 筛选到新型抗生素 Terragene, 充分证明了新型载体2宿主系统的有效性.

针对质粒型小插入片段环境宏基因组文库的异源表达问题, 最新的技术突破是双向表达质粒载体的构建与应用. 该新型载体的应用显著增强了质粒型小文库中目的 DNA 的表达水平、加快了有用分子的发现速率^[28].

3.1.2 构建文库过程中连接反应的改善

通过剧烈的原位裂解方式所获得的 eDNA 片段很小(0.5~ 5 kbp), 为避免已被严重剪切的 eDNA 携带信息的进一步损失, 不宜选择限制性酶切反应产生粘性末端与载体连接, 采用末端补平或 T2A 连接的方式能有效解决这一难题^[29].

3.2 环境宏基因组文库分析技术的提高

3.2.1 高通量测序技术在宏基因组文库中的应用及相关生物信息学分析方法的发展

自动高通量测序技术在环境宏基因组文库中的运用产生了海量的生态学数据. 如, 通过随机鸟枪法测序从马尾藻海水宏基因组文库获得了至少包含该海域 1 800 种微生物基因组信息, 其中 148 种为新的微生物物种; 并检测到 120 万个新基因, 极大程度丰

富了海洋微生物的多样性资料^[20]。

应该重视的是,环境宏基因组文库的高通量测序结果是混合在一起的,来源于多种不同微生物的序列,必须结合先进的生物信息学方法、新型计算工具,才能将各个序列片段精确组装,以准确地归类于它所对应的微生物物种,得到无偏差的群落结构分析结果。尤其对微生物物种多样性高度丰富的土壤宏基因组文库进行大量序列分析时,颇具难度^[30]。近几年,综合微生物基因组(integrated microbial genomes, IMG)数据管理系统的更新^[31];新型数学运算法则^[32]及新型MEGAN软件的发明^[33];先进测序技术^[34]、多位点序列分型(multilocus sequence typing)^[35]、多基因表达谱(metagene)预测^[36]、DNA改组(metagenomic DNA shuffling)^[37]、SSAKE(Short Sequence Assembly by progressive Kmer search and 3-read Extension program)^[38]等技术的发明;以及Visualization技术^[39]、统计学方法在环境宏基因组学领域的应用^[40],为文库所包含海量信息的精确分析奠定了坚实的理论基础。

基于基因芯片的高通量测序技术,即在对宏基因组克隆子测序之前,运用基因芯片在文库中预选基因,以缩小克隆子测序的范围,简化后续生物信息分析过程,同时能减轻大量序列筛选工作的负荷,使得快速筛选大量克隆子成为可能^[41]。

31212 抑制性消减杂交技术(suppressive subtractive hybridization, SSH)

SSH是用于比较分析文库中微生物群落结构的新技术^[42]。

31213 突变转座子(transposon MuExpress)诱导表达技术

构建新的突变转座子MuExpress,体外随机整合于环境BAC或Cosmid文库克隆子中,可诱导它侧面序列的双向表达,从而增强目的基因的表达能力^[43]。

31214 底物诱导基因表达(substrate induced gene expression screening, SIGEX)技术的发展

SIGEX是基于功能分析的新型筛选方法。基本原理是通过operon2trap gfp表达载体构建宏基因组文库,用添加了底物的液体培养基培养克隆子,继而通过荧光激活流式细胞分选器寻找含有目的基因的细胞。该技术的运用能增强宏基因组文库中外源基因的表达能力,有助于揭示全新的蛋白质序列^[44]。SIGEX技术已成功地应用于地下水样品环境基因组文库研究,用安息香酸盐诱导得到33个阳性克隆,

采用蔡诱导得到2个阳性克隆^[45]。

31215 提高稀有基因在宏基因组文库中呈现频率的技术

因稀有基因在自然生境中所占比例低下,在文库的构建与分析过程往往被高丰度的DNA片段蒙蔽。目前,已经报道了3种提高稀有基因在宏基因组文库中呈现频率的方法。

(1) Normalisation Normalisation具体操作方式有2种。一种是使用嵌合剂类化合物,如bis2benzimidazole,根据G+C含量不同,运用氯化铯密度梯度离心从混合的环境微生物DNA中分离特定的基因型^[46]。另一种是根据严格退火条件下,已变性的混合的基因组片段中,丰度高的单链DNA复性速度快于稀有DNA复性速度的特征分离得到稀有基因型^[47]。

(2) 噬菌体表达文库(phage display expression library) 噬菌体表达文库技术提供了一种通过亲和选择筛选法从噬菌体表达产物中分离DNA序列的手段^[48]。此方法在宏基因组文库中的运用不仅提供了从文库中富集稀有DNA序列的潜能,而且有利于提高文库的高通量筛选效率^[47]。

(3) 改良的反向PCR技术 2006年,Uchiyama等^[49]报道了一种从环境基因组文库中获得稀有基因的反向PCR技术,主要操作步骤为:首先用含亲和标记物的引物进行反向PCR,然后用亲和纯化法除去背景宏基因组DNA,最后通过巢式PCR回收目的DNA。

31216 宏基因组学方法与其它方法联用

(1) 与荧光原位杂交方法的联用 目前该技术的主要应用范围是对文库中所包含的微生物进行系统进化分析^[50]。

(2) 与稳定同位素标记方法的联用 与稳定同位素标记方法的联用能直接富集目的基因,以提高目的基因的筛选频率。Schwarz等^[51]首次将基于DNA的同位素标记方法运用于环境宏基因组文库研究,证明了该方法的有效性。

(3) 与传统的富集培养方法的联用 根据研究目的,在构建文库之前对样品进行预富集,提供了一个从整体细胞富集到目的基因富集的范围。环境宏基因组学技术与传统的富集培养方法的联用能明显提高阳性克隆子的比率,是从复杂样品中分离大量新型目的基因的高效技术^[4,5]。

(4) 与新型微生物培养方法的联用 根据宏基因组文库提供的线索,通过新的培养方法,获得具有重要生态功能的未培养微生物,将进一步推进微生

物地球化学循环的研究, 丰富微生物生态学内容^[52]. 与新型微生物培养方法的联用一定程度上克服了宏基因组文库技术的固有瓶颈, 突破未培养微生物/ 可知不可得0的被动局面.

分析环境宏基因组学技术的发展进程可知, 宏基因组文库筛选策略的优化是该领域的发展重点.

新方法的发明与运用均以提高筛选频率及更充分发掘文库资源为目的. 每种研究方法各有优缺点, 实际工作中, 不同方法的选用与组合应根据研究目的等因素全面考虑. 表 1 列出了现阶段环境宏基因组文库筛选方法的主要优缺点, 图 1 显示了环境宏基因组学的发展历程.

表 1 宏基因组文库不同筛选方法比较

Table 1 Comparison of the screening methods of metagenome library

方法	主要优点	主要缺点
杂交	快速、直接	只能检测已知家族基因
PCR	快速、直接	只能检测已知家族基因; 不能获得合成目的产物的全长基因
基于序列筛选的相关技术 鸟枪法全基因组测序	能获得重组克隆子的完整序列; 能获得大量数据, 揭示全新基因, 推断未培养微生物的代谢途径	测序费用昂贵, 劳动强度极大; 对数据的正确分析及精细组装难度大
基于生物芯片的高通量筛选技术	能迅速识别、鉴定大量的克隆; 降低全基因组测序负荷	不能获得合成目的产物的全部基因簇; 得到的序列不能被运用于生物化学反应
与荧光原位杂交法联用	允许对大片段宏基因组文库的克隆子进行快速系统进化分析	从经固定的细胞中分离重组 DNA 有一定难度
经典的功能筛选	能产生全新的基因或化合物	频率低下, 工作量大
基于功能筛选的相关技术 底物诱导基因表达	允许半自动操作, 节省时间; 提高阳性克隆比率	对于预表达基因的结构与方向敏感; 易漏检克隆; 不适用于大片段宏基因组文库的筛选
噬菌体表达文库 与传统的富集培养方法联用	利于文库中稀有 DNA 序列的表达 提高阳性克隆比率	能表达蛋白质大小的最高极限是 50 kDa 导致文库中微生物多样性的大量丧失
与稳定同位素标记方法联用	提高阳性克隆比率	底物昂贵, 复杂的底物不易获得; 底物消耗之后, 非目标微生物的营养共生作用将导致重密度和中间密度 DNA 带型分离不清

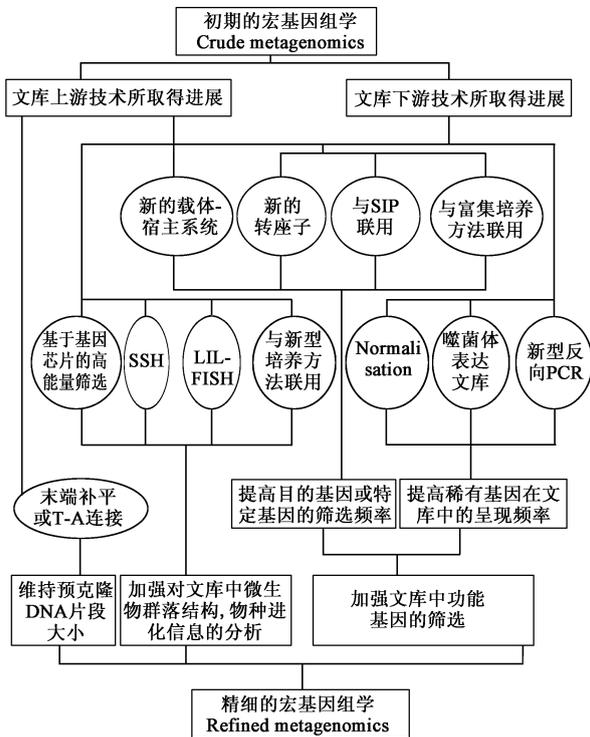


图 1 环境宏基因组文库技术发展进程

4 展望

环境宏基因组学技术是与多种先进技术交叉的综合性现代前沿技术, 对于微生物学的发展发挥了极大的推动作用; 在世界范围内的/ 生物探矿热0中, 其本身也在不停地发展与完善, 新型技术的发明与应用已经取得了明显的效果.

今后的发展中, 在微生物生态学基础研究领域, 应加强不同方法的科学结合, 以减少单一研究方法造成的偏差. 对特定生境进行微生物物种多样性研究时, 宏基因组文库技术可避免 PCR 偏好性所导致的误差, 而 PCR 文库可矫正构建宏基因组文库的基因遗漏现象. 因此, 宏基因组文库与 PCR 文库结合分析更能准确地反映生境微生物的多样性. 此外, 应注重与宏蛋白质组学技术的联用, 只有两者的高度结合, 才能探寻复杂微生物群落中的感应信号分子及其基因调节机制, 探知微生物物种在群落中的功能.

研究范围的进一步拓宽在今后的发展中也不容忽视. 目前, 运用宏基因组技术对海洋病毒已有较深入的研究^[53, 54], 而对土壤病毒所开展的研究还很少.

Fig. 1 Diagram illustrating the development of metagenomics

构建环境 cDNA 文库,促进真菌资源的开发应用研究,也将是宏基因组学的主要研究内容之一。相应地,有效裂解土壤真菌以保证 poly(A) mRNA 基因完整性的 eDNA 提取方法的发展十分重要。

研究生境类型的范围还应进一步拓展。2005 年,美国联合基因组研究所构建了来源于 4 万年前洞穴熊(已经灭绝的种类)的骨骼宏基因组文库,得到 26 861 bpp 洞穴熊基因组序列,通过与现代熊基因序列比较,揭示了其进化关系^[5]。运用宏基因组学方法揭示已灭绝物种的基因组信息,探知古老基因的秘密,是一个非常有价值的发展方向。针对与人类健康密切相关的胃肠道环境的微生态平衡调节机制及失衡条件下发病机制的研究应当在环境宏基因组学研究的范围之内。对处于/海洋/陆地/界面潮间带的微生物资源也应重视。目前世界范围内,采用宏基因组学方法从此生境中寻找新资源的报道极少。笔者针对这一空白,克服了运用原位裂解法从高粘质性、高有机质含量的土壤构建大片段宏基因组文库的难点,构建了 4 个季节红树林土壤大片段 Cosmid 宏基因组文库,随机挑取的克隆子经 BamHI 酶切验证,平均插入片段均大于 35 kbp,且片段类型均不相同,为潮间带特色微生物资源的开发应用提供了良好素材。

生物技术领域应注重加快天然药物的发现进程。因为编码天然药物的基因簇主要存在于放线菌中,笔者认为在提取 eDNA 时,可根据 G+ C 含量的不同将放线菌 DNA 与其它类型微生物 DNA 分离,单独构建环境样品的放线菌宏基因组文库,以减轻从文库中筛选药物的负荷,提高筛选效率。

参考文献:

- [1] Handelsman J, Rondon M R, Brady S F, et al. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products [J]. *Chem Biol*, 1998, 5(10): 245~ 249.
- [2] Schmeisser C, Steele H, Streit W R. Metagenomics, biotechnology with nonculturable microbes [J]. *Appl Microbiol Biotechnol*, 2007, 75(5): 955~ 962.
- [3] Howard E C, Henriksen J R, Buchan A, et al. Bacterial taxa that limit sulfur flux from the ocean [J]. *Science*, 2006, 314(5799): 649~ 652.
- [4] Lorenz P, Eck J. Metagenomics and industrial applications [J]. *Nat Rev Microbiol*, 2005, 3(6): 510~ 516.
- [5] Daniel R. The metagenomics of soil [J]. *Nat Rev Microbiol*, 2005, 3(6): 470~ 478.
- [6] Steele H L, Streit W R. Metagenomics: Advances in ecology and biotechnology [J]. *FEMS Microbiol Lett*, 2005, 247(2): 105~ 111.
- [7] Miller D N, Bryant J E, Madsen E L, et al. Evaluation and optimization of DNA extraction and purification procedures for soil and sediment samples [J]. *Appl Environ Microbiol*, 1999, 65(10): 4715~ 4724.
- [8] Zhou J, Bruns M A, Tiedje J M. DNA recovery from soils of diverse composition [J]. *Appl Environ Microbiol*, 1996, 62(2): 316~ 322.
- [9] Bergmann H, Pesaro M, Widmer F, et al. A strategy for optimizing quality and quantity of DNA extracted from soil [J]. *J Microbiol Methods*, 2001, 45(1): 7~ 20.
- [10] Krsek M, Wellington E M H. Comparison of different methods for the isolation and purification of total community DNA from soil [J]. *J Microbiol Methods*, 1999, 39(1): 1~ 16.
- [11] Roese-Amsaleg C L, Gamie-Sillam E, Harry M. Extraction and purification of microbial DNA from soil and sediment samples [J]. *Appl Soil Ecol*, 2001, 18(1): 47~ 60.
- [12] Gupta R, Beg Q K, Lorenz P. Bacterial alkaline proteases: molecular approaches and industrial applications [J]. *Appl Microbiol Biotechnol*, 2002, 59(1): 15~ 32.
- [13] Pettit R K. Soil DNA libraries for anticancer drug discovery [J]. *Cancer Chemother Pharmacol*, 2004, 54(1): 1~ 6.
- [14] Bertrand H, Poly F, Van V T, et al. High molecular weight DNA recovery from soils prerequisite for biotechnological metagenomic library construction [J]. *J Microbiol Methods*, 2005, 62(1): 1~ 11.
- [15] 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 [J]. *Appl Environ Microbiol*, 2000, 66(6): 2541~ 2547.
- [16] Lee S Y, Bollinger J, Bezdecke D, et al. Estimation of the abundance of an uncultured soil bacterial strain by a competitive quantitative PCR method [J]. *Appl Environ Microbiol*, 1996, 62(10): 3787~ 3793.
- [17] Nocker A, Camper A K. Selective removal of DNA from dead cells of mixed bacterial communities by use of ethidium monoazide [J]. *Appl Environ Microbiol*, 2006, 72(3): 1997~ 2004.
- [18] Green B D, Keller M. Capturing the uncultivated majority [J]. *Current Opinion in Biotechnology*, 2006, 17(3): 236~ 240.
- [19] Wintzingerode F, Ebel U B, Stackebrandt E. Determination of microbial diversity in environmental samples: pitfalls of PCR-based analysis [J]. *FEMS Microbiol Rev*, 1997, 21(3): 213~ 229.
- [20] Venter J C, Remington K, Heidelberg J F, et al. Environmental genome shotgun sequencing of the Sargasso Sea [J]. *Science*, 2004, 304(5667): 66~ 74.
- [21] Johnston A W B, Li Y G, Ogilvie L. Metagenomic marine nitrogen fixation: feast or famine? [J]. *Trends Microbiol*, 2005, 13(9): 416~ 420.
- [22] Riesenfeld C S, Goodman R M, Handelsman J. Uncultured soil bacteria are a reservoir of new antibiotic resistance genes [J]. *Environ Microbiol*, 2004, 6(9): 981~ 989.
- [23] Seow K T, Meurer G, Gerlitz M, et al. A study of iterative type II polyketide synthases, using bacterial genes cloned from soil DNA: a means to access and use genes from uncultured microorganisms [J]. *J Bacteriol*, 1997, 179(23): 7360~ 7368.
- [24] 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 [J]. *Appl Environ Microbiol*, 2004, 70(4): 2452~ 2463.
- [25] Wang G Y, Graziani E, Waters B, et al. Novel natural products from soil DNA libraries in a streptomycete host [J]. *Org Lett*, 2000, 2(16): 2401~ 2404.
- [26] Wexler M, Bond P L, Richardson D J, et al. A wide host range metagenomic library from a waste water treatment plant yields a novel alcohol aldehyde dehydrogenase [J]. *Environ Microbiol*, 2005, 7(12): 1917~ 1926.
- [27] Li Y, Wexler M, Richardson D J, et al. Screening a wide host range, waste water metagenomic library in tryptophan auxotrophs of *Rhizobium leguminosarum* and of *Escherichia coli* reveals different classes of cloned *trp* genes [J]. *Environ Microbiol*, 2005, 7(12): 1927~ 1936.
- [28] Lammle K, Zipper H, Breuer M, et al. Identification of novel enzymes with different hydrolytic activities by metagenome expression cloning [J]. *J Biotechnol*, 2007, 127(4): 575~ 592.
- [29] Wilkinson D E, Jeanicke T, Cowan D A, et al. Efficient molecular cloning of environmental DNA from geothermal sediments [J]. *Biotechnol Lett*, 2002, 24(2): 155~ 161.
- [30] Tringe S G, von Mering C, Kobayashi A, et al. Comparative metagenomics of microbial communities [J]. *Science*, 2005, 308(5721): 554~ 557.
- [31] Markowitz V M, Ivanova N, Palaniappan K, et al. An experimental metagenome data management and analysis system [J]. *Bioinformatics*, 2006, 22(20): 2580~ 2580.
- [32] Krause L, Diaz N N, Bartels D, et al. Finding novel genes in bacterial communities isolated from the environment [J]. *Bioinformatics*, 2006, 22(14): 281~ 289.
- [33] Huson D H, Auch A F, Qi J, et al. MEGAN analysis of metagenomic data [J]. *Genome Res*, 2007, 17(3): 377~ 386.
- [34] Hall N. Advanced sequencing technologies and their wider impact in microbiology [J]. *J Exp Biol*, 2007, 210(9): 1518~ 1525.
- [35] Mahenthiralingam E, Baldwin A, Drevinek P, et al. Multilocus sequence typing breathes life into a microbial metagenome [J]. *PLoS ONE*, 2006, 1(1): 1~ 8.
- [36] Noguchi H, Park J, Takagi T. MetaGene: prokaryotic gene finding from environmental genome shotgun sequences [J]. *Nucleic Acids Res*, 2006, 34(19): 5623~ 5630.
- [37] Boubakri H, Beuf M, Simonet P, et al. Development of metagenomic DNA shuffling for the construction of a xenobiotic gene [J]. *Gene*, 2006, 375(6): 87~ 94.
- [38] Warren R L, Sutton Granger G, Jones Steven J M, et al. Assembling millions of short DNA sequences using SSAKE [J]. *Bioinformatics*, 2007, 23(4): 500~ 501.
- [39] Have S L, Webb Robertson B J, Shah A, et al. Bioinformatic insights from metagenomics through visualization [A]. In: *Proc IEEE Comput Syst Bioinform Conf [C]*. Los Alamitos, CA: IEEE Computer Society, 2005. 341~ 350.
- [40] Rodriguez Brito B, Rohwer F, Edwards R A. An application of statistics to comparative metagenomics [J]. *BMC Bioinformatics*, 2006, 7: 162.
- [41] Sebat J L, Colwell F S, Crawford R L. Metagenomic profiling: microarray analysis of an environmental genomic library [J]. *Appl Environ Microbiol*, 2003, 69(8): 4927~ 4934.
- [42] Galbraith E A, Antonopoulos D A, White B A. Suppressive subtractive hybridization as a tool for identifying genetic diversity in an environmental metagenome: the rumen as a model [J]. *Environ Microbiol*, 2004, 6(9): 928~ 937.
- [43] Leggovic C, Henning H, Schmeisser C, et al. A novel transposon for functional expression of DNA libraries [J]. *J Biotechnol*, 2006, 123(3): 281~ 287.
- [44] Yun J, Ryu S. Screening for novel enzymes from metagenome and SIGEX, as a way to improve it [J]. *Microb Cell Fact*, 2005, 4(8): 1~ 5.
- [45] Uchiyama T, Abe T, Ikemura T, et al. Substrate-induced gene expression screening of environmental metagenomic libraries for isolation of catabolic genes [J]. *Nat Biotechnol*, 2005, 23(1): 88~ 93.
- [46] Li X, Qin L. Metagenome-based drug discovery and marine microbial diversity [J]. *Trends Biotechnol*, 2005, 23(11): 539~ 543.
- [47] Wenzel S C, Muller R. Recent developments towards the heterologous expression of complex bacterial natural product biosynthetic pathways [J]. *Curr Opin Biotechnol*, 2005, 16(6): 594~ 606.
- [48] Gramer 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 [J]. *Gene*, 1993, 137(1): 69~ 75.
- [49] Uchiyama T, Watanabe K. Improved inverse PCR scheme for metagenome walking [J]. *Biotechniques*, 2006, 41(2): 183~ 188.
- [50] Leveau J H J, Gerards S, de Boer W, et al. Phylogenetic function analysis of (meta)genomic libraries: screening for expression of ribosomal RNA genes by large insert library fluorescent in situ hybridisation (LILFISH) [J]. *Environ Microbiol*, 2004, 6(9): 990~ 998.
- [51] Schwarz S, Washkowitz T, Daniel R. Enhancement of gene detection frequencies by combining DNA-based stable isotope probing with the construction of metagenomic DNA libraries [J]. *World J Microbiol Biotechnol*, 2005, 22(4): 363~ 368.
- [52] Rappe M S, Connon S A, Vergin K L, et al. Cultivation of the ubiquitous SAR11 marine bacterioplankton clade [J]. *Nature*, 2002, 418(6898): 630~ 633.
- [53] Cullley A I, Lang A S, Suttle C A. Metagenomic Analysis of Coastal RNA Virus Communities [J]. *Science*, 2006, 312(5781): 1795~ 1798.
- [54] Delwart E L. Viral metagenomics [J]. *Rev Med Virol*, 2007, 17(2): 115~ 131.
- [55] Noonan J P, Hofreiter M, Smith D, et al. Genomic Sequencing of Pleistocene Cave Bears [J]. *Science*, 2005, 309(5734): 597~ 599.