

宏基因组学技术的研究与挑战

周丹燕^{1,2} 戴世鲲¹ 王广华¹ 李翔^{1*}

(1. 中国科学院南海海洋研究所海洋生物资源可持续利用重点实验室 广东省海洋药物重点实验室 广东
广州 510301)

(2. 中国科学院研究生院 北京 100049)

摘要: 宏基因组学研究作为研究微生物种群生态分布、群体遗传特征和基因相互作用的新兴学科领域, 在很大程度上促进了环境微生物资源, 特别是未培养微生物资源的开发利用, 在土壤、海洋、人体医学、药物等各个领域的应用中取得了突破性的进展, 为发现新的生物活性物质提供了新的有效途径。就宏基因组学研究进展进行综述, 并重点介绍了宏基因组学研究中的机遇及挑战。
关键词: 宏基因组学, 宏基因组文库, 序列筛选, 功能筛选, 新一代测序

The research innovation and challenges in metagenomics

ZHOU Dan-Yan^{1,2} DAI Shi-Kun¹ WANG Guang-Hua¹ LI Xiang^{1*}

(1. Key Laboratory of Marine Bio-resources Sustainable Utilization, Key Laboratory of Marine Drug, South China Sea Institute of Oceanology Chinese Academy of Sciences, Guangzhou, Guangdong 510301, China)
(2. Graduate University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: As the new scientific field of studying the gene function and interactions of microbial community, metagenomics has largely contributed to the development and utilization of microbial resources, and made breakthrough progress in various fields, such as soil, oceans, human medical, pharmaceutical and so on. Metagenomics provided a novel approach for the discovery of new bioactive substances. Here we discuss some recent progress with emphasis on the opportunities and challenges in metagenomics applications.

Keywords: Metagenomics, Metagenomic library, Sequence-based screening, Function-based screening, Next-generation DNA sequencing

宏基因组学提出之前, 长期以来人们主要通过实验室培养的方法认识和研究微生物, 开发利用微生物资源的技术停留在单一微生物物种水平上。然

而, 由于环境中 99.8%的微生物不能用常规的培养方法培养, 人们对于绝大多数难培养微生物物种几乎是一无所知, 对于这些微生物功能的描述受到培

基金项目: 国家 973 计划项目(No. 2010CB833801-03); 大洋协会项目课题(No. DYXM-115-02-2-14); 广东省科技计划项目(No. 2008A030203004)

* 通讯作者: Tel: 86-20-89023013; Fax: 86-20-84451672; ✉ lixiang@scsio.ac.cn © 中国科学院微生物研究所期刊联合编辑部 <http://journals.im.ac.cn>
收稿日期: 2010-11-01; 接受日期: 2011-01-04

养的限制。而从开发微生物资源的角度来看,那些迄今为止还未发掘的难分离培养的微生物才是发掘新化合物的重要来源,传统方法对于这些微生物资源的开发利用是捉襟见肘的。另外,传统生物活性物质筛选法除了存在“培养限制”问题,还会遇到“发酵限制”问题,即在发酵过程中本来产生某活性物质的细菌会丧失该物质的合成能力,造成目的产物的丢失^[1]。

宏基因组学(Metagenomic)技术不依赖培养,从微生物的天然环境中直接提取基因组遗传物质,对微生物群体进行研究分析,丰富了人们对于微生物群体生态和进化的认识,最大限度地挖掘微生物资源,使环境微生物的研究不只集中在“确定它们是什么”,还能绕过培养环节而聚焦于“确定它们能做什么”。宏基因组学技术是寻找新基因、开发新的生物活性物质、研究群落中微生物多样性的新途径,还可以帮助人们发现极端生命空间中存在的特殊生物群体或特殊基因资源,它的兴起填补了未可培养微生物研究的空白,成为国际微生物学研究的热点。

1 宏基因组学技术简介

1.1 宏基因组学技术概述

1991年 Pace 等人^[2]首次提出环境基因组学(Environmental genomics)的概念,他们在对太平洋浮游生物进行系统分类时直接提取样品总 DNA 经酶切连入载体转化大肠杆菌构建了第一个噬菌体文库,经过筛选分析发现了 15 种全新的细菌序列。这是第一次正式发表的利用宏基因组文库发现新物种的报道。1998年 Handelsman 等^[3]首次将“特定小生境中全部微小生物遗传物质的总和”定义为宏基因组(Metagenome),提出针对特定环境样品中遗传物质总和进行研究。利用宏基因组学技术了解环境微生物多样性、开发未可培养微生物的研究逐渐兴起,关于宏基因组学技术的研究报道逐年增多,如图 1 所示。近 10 年来,关于土壤、海洋等环境样品微生物宏基因组学技术的一些研究成果见表 1。

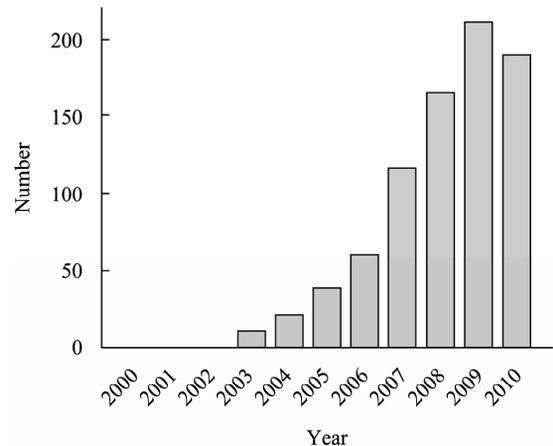


图 1 通过 ISI 科学数据库检索到的有关“Metagenomic”的文献数

Fig. 1 Number of publications retrieved from the ISI Web of Science database

1.2 宏基因组学技术策略

宏基因组学的基本策略是:采用原位裂解法或间接提取法尽可能完全地抽提环境样品中的总 DNA;利用合适的载体,如粘粒载体、Fosmid 载体、细菌人工载体(BAC)等,将宏基因组 DNA 克隆到模式微生物中,如大肠杆菌、链霉菌,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(簇)。

需要注意的是,环境样品中存在诸多不同的因素影响着 DNA 的提取效果^[19],克隆操作过程中基因或基因簇的断裂破坏可能导致基因不能表达,环境中存在的具有核酸酶活性的物质也会使完整 DNA 的获取具有很大的难度。现有的 DNA 提取技术尚不能适用所有的环境样品,不能覆盖环境样品中所有的微生物类群(基因组类群)。牛泽等^[20]以北京地区重金属污染土样为研究对象,对文献报道的各种土壤 DNA 提取纯化方法进行了比较和改进,探索出了一套高效的重金属污染土壤混合基因组的提取、纯化方法,即先用 DNA 提取液直接裂解土壤中微生物,后用蛋白酶 K 和 SDS 处理,再用酚-氯仿-异戊醇(25:24:1, V/V/V)抽提的方法比较适合采集到的土壤,粗提 DNA 的效率可达 150 μg DNA/g 土壤。Liles 等人^[21]证明了用甲酰胺高盐溶液处理的样品

表1 土壤、海洋等环境样品微生物宏基因组学技术研究成果
Table 1 Metagenomics technology in the research of soil, marine and other environmental samples

研究人员 Researchers	研究内容 Research contents	参考文献 References
Rondon MR	利用细菌人工染色体作为载体构建土壤总 DNA 宏基因组文库, 发现了大量的微生物门类, 筛选到了包括抗生素、淀粉酶、脂肪酶和核酸酶等大量生物活性物质	[4]
Gillespie DE	用细菌人工染色体载体 pBeloBACII 构建了 2 个土壤样品的宏基因组 BAC 文库, 获得了 2 个具有广谱抗菌作用的新抗生素 TurbomycinA (橙色化合物)和 TurbomycinB (红色化合物)及其合成酶基因簇	[5]
Courtois S	构建了一个可以在大肠杆菌和链霉菌间进行转移的穿梭 Cosmid 载体, 并以此构建了一个土壤宏基因组文库, 通过序列驱动筛选获得了新的抗肿瘤活性物质聚酮的合成酶基因簇	[6]
Hildebrand M	构建了 2 个不同地域苔藓虫的富含“ <i>E. sertula</i> ”菌基因组的总 DNA 文库, 通过锚定 KS 结构域保守区的引物筛选、序列分析和拼接, 获得了全长约为 80 kb 的 Bryostatin 生物合成基因簇	[7]
Lim HK	构建了森林土壤宏基因组文库, 从中筛选出了产色素 Indirubin 和 Indigo, 且具有对 <i>Bacillus subtilis</i> 的抗菌活性的克隆, 并为目前还了解甚少的 <i>Acidobacteria</i> 的研究提供宝贵的资料	[8]
吴杰 WU Jie	构建了澳大利亚厚皮海绵的 Fosmid 宏基因组文库, 这是我国首次尝试构建海绵宏基因组文库, 对于今后开发利用海绵丰富的基因资源具有重要的意义	[9]
CHEN J	构建了南海海域水深 5 m 处的海绵 <i>Gelliodes gracilis</i> 共附生生物的宏基因组文库, 从中筛选得到具有抗菌活性的克隆, 这是利用宏基因组 DNA 文库筛选海绵未培养共附生微生物抗菌活性的首次报道	[10]
陆伟 LU Wei	从草甘膦极端污染土壤中直接分离提取细菌 DNA 构建宏基因组文库, 筛选草甘膦不敏感的 5-烯醇式丙莽草酰-3-磷酸合成酶(EPSPS)基因, 并利用缺少芳香族氨基酸的 MOPS 限制性培养基筛选功能互补的 <i>aroA</i> 基因, 获得了 2 个对草甘膦具有天然耐受能力的克隆	[11]
张金伟 ZHANG Jin-Wei	从南极普利兹湾深海 900 m 深的表层沉积物中提取获得宏基因组 DNA, 获得了低温脂肪酶 <i>lip3</i> 的开放阅读框完整序列及该酶的表达	[12]
Lee MH	构建了第一个朝鲜西海岸潮间带海区的环境基因组文库	[13]
蒋云霞 JIANG Yun-Xia	构建了天然红树林土壤 Cosmid 宏基因组文库, 该文库包含了 4 个季节潮间带微生物的基因组信息, 为开发红树林湿地生态系统的特色微生物资源提供了可能	[14]
XU M	构建了太平洋深海沉淀物宏基因组文库, 并成功获得携带独特的烷烃羟化酶基因 <i>alkB</i>	[15]
Matrin-Cuadrad AB	构建了地中海 3 000 m 水深处浮游生物的宏基因组 Fosmid 文库, 发现了生物发光(Bioluminescence)基因	[16]
Mori T	以活化污泥为 DNA 来源构建宏基因组文库, 获得 3 个抗争光霉素的克隆	[17]
Zehr JP	利用 GS LLX Titanium 高通量测序系统对海洋样品的宏基因组进行测序, 发现了一种基因型全新的蓝藻物种, 该种生物只专注于固氮, 从基因看出无法进行产生氧气的光合作用	[18]

能够提取大片段的 DNA, 并能有效地提高宏基因组文库的构建效率, 从而为我们获取完整的大片段 DNA 提供了可借鉴的方法。欧阳永长^[22]利用琼脂糖凝胶包埋法从中国南海的多种海绵中分离得到超过 400 kb 大小的细菌 DNA 片段, 并用获得的大片段 DNA 构建了海绵 *Halichondria* sp. 的 BAC 文库, 提供了一个有效的异源表达海绵未培养基因资源来进行药物开发的途径。

在常用的载体中, BAC 载体可以克隆大片段 DNA, 增加了获得完整的指导分子如抗生素合成途

径的编码基因或基因簇的机会, 在宏基因组克隆表达中具有相当优势。Martinez 等^[23]利用 BAC 载体建立了包含大片段 DNA 的大肠杆菌宏基因组文库, 通过高通量接合转移实现大分子量载体分子从大肠杆菌向变铅青链霉菌 *S. lividans* 和恶臭假单胞菌 *P. putida* MBD1 的转移, 使宏基因组文库得到了更有效的表达。

环境样品中的微生物种类繁多, 环境宏基因组文库容量一般较大, 如 Henne 等^[24]筛选一个小片段 (<10 kb) 的文库每次都需要筛选 10^5 – 10^6 个克隆。从

宏基因组文库中筛选出带有目的基因克隆的方法主要有基于序列的筛选法和基于功能的筛选法。

序列筛选(Sequence-based screening)依赖目的基因保守的DNA序列,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,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(Function-based screening)以生物活性为线索,主要是根据一些酶蛋白的功能,在功能平板上检测宿主菌出现的一些变化,如阳性克隆周围出现透明圈或酶降解产物显现某种颜色,先获得具有生物活性的阳性克隆子,再进行相关测序分析得到相应的基因结构。通过功能筛选能够得到完整的功能基因或带有目的基因的基因簇,而且这些功能基因很可能是全新的。

2 序列筛选法和功能筛选法的特点

2.1 序列筛选法

实践证明,海绵宏基因组编码许多指导不同途径的聚酮合成酶(Polyketide synthase, PKS)^[25-27],来自海绵的具有药用价值的天然产物大多数是复杂的聚酮类化合物。PKS包括一个或多个模块,每个模块含有3个关键催化功能位点:酮体合成酶(Ketosynthase, KS)、酰基转移酶(Acyl transferase, AT)、酰基载体蛋白(Acyl carrier protein, ACP)。其中,KS编码区同源性最高,因此,可以根据KS结构域的保守序列设计PCR引物^[25-26],从海绵宏基因组文库中筛选携带PKS基因的克隆。这一序列筛选法同样适用于海洋中其他生物,如*Bugula nertina*幼虫宏基因组文库^[28]以及土壤宏基因组文库^[29-30]的筛选。

与功能筛选法相比,基于序列的筛选策略在很大程度上并不能保证全基因(簇)的获得,其局限性主要在于需要对功能产物已知的编码基因序列进行分析,文库中那些和现有基因序列完全不一样的基因无法被筛选,较难发现全新的活性物质。不过,Gupta等^[31]根据他们的经验认为,当宏基因组文库基因的保守序列与数据库中已知基因的保守序列相似度低至40%时,基于序列的筛选策略也可能筛选

到新的功能产物。

2.2 功能筛选法

功能筛选法是生物活性水平的筛选,或是对具有特殊功能的克隆子进行直接检测,这种方法具有较高的灵敏度,产色素的阳性克隆因其在平板上呈现的颜色可以被快速地筛选出来,如Yali Huang等^[32]从构建的中国南海深海沉积物Fosmid宏基因组文库中筛选到了产黑色素(Melanin)的克隆;或是基于异源基因的宿主菌株与其突变体在选择性条件下功能互补生长的特性进行筛选^[33]。功能筛选方法适合那些不存在同源序列而无法设计通用探针或PCR引物的编码酶或蛋白质基因的筛选,如在脂肪酶的筛选上目前常用功能筛选法^[34],环境基因组中N-酰基酪氨酸合成酶(NAS)基因序列之间的相似性很低,杂交或基于PCR的策略对于快速鉴定新的NAS序列是不可行的,基于表达的功能筛选策略是唯一可行的途径^[35]。通过功能性筛选可以快速地多个克隆子中鉴定出全长基因,并由此获得这些功能基因的产物,为工业、医药和农业提供一些新的天然活性物质如各种酶类及一些次生代谢产物等,包括4-羟基丁酸脱氢酶^[24]、N-酰基酪氨酸合成酶^[35]、几丁质酶^[36]、淀粉酶^[37]、蛋白酶^[38]、脂肪酶^[39]、酯酶^[40-41]等。

然而,功能筛选法往往需要目的基因(簇)在宿主中能良好地表达,表达产物能够正常溶解,并且产生的蛋白质要能够正确地折叠。然而,很多克隆的外源基因并不能在新的遗传背景中实现表达。例如,宏基因组DNA在使用的宿主菌株中不能适当转录,mRNA不能有效翻译,或者在很多情况下翻译产生的蛋白质由于缺乏分子伴侣不能正确折叠而不具有活性。另外,不同菌株细胞的密码子使用偏好性不同,不同的密码子对应的tRNA的数量也不同,从而影响外源蛋白的翻译。这些局限性造成了外源基因的异源表达水平低下或外源蛋白的活性低,进而使功能筛选法的阳性克隆率很低,通常数千个甚至数百万个克隆中只能检测到少于10个的阳性克隆,比如:Henne等^[39]从构建的宏基因组文库中筛选具有分解脂肪活性的克隆,阳性克隆率分别

为 1/730 000 和 3/286 000; Lee 等^[13]以潮间带沉积物为样品构建文库, 获得的阳性克隆率为 4/386 400; Lee 等^[38]构建韩国西海岸深海沉积物宏基因组文库, 在 30 000 个重组子中具有金属蛋白酶活性的阳性克隆也仅为 1 个。另外, Courtois 等^[6]在对文库抗菌活性等方面进一步分析时还发现, 具有在大肠杆菌中稳定表达的卡那霉素抗性基因的粘粒 a8E12 转化链霉菌 TK24 时卡那霉素抗性丢失。这些现象的发生都不利于功能筛选的顺利实现。

2.3 功能筛选法的改进

要使功能筛选法更好地服务于宏基因组学技术, 需要建立合适的宿主载体系统, 不断改进表达系统, 进一步解决宏基因组基因在多宿主系统中的表达问题, 如采用穿梭载体, 以利于其在不同宿主中进行目的基因的表达, 或是对 *E. coli* 进行改造以扩大基因表达范围。目前宏基因组文库的构建常采用的表达系统是大肠杆菌和链霉菌, 也可以通过综合考虑宿主的天然生物合成能力和它对外源启动子的倾向性拓展新的宿主菌用于环境 DNA 的研究。Craig 等人^[42]利用 *Ralstonia metallidurans* 作为克隆宿主成功获得了土壤 eDNA 的编码产物, 其中有 2 个是新天然产物, 他们还进一步验证了获得的产物在大肠杆菌中不能表达或无法被检测到表达, 进而证明了 *Ralstonia metallidurans* 可以作为宏基因组学文库的新的模式系统, 用于发现新产物。

另外, 还可以结合一些已发展较成熟的技术来改进功能筛选法, 如在克隆前运用超速离心法、Nycodenz 密度梯度离心法、稳定同位素示踪技术、Brd U 标记的 DNA 免疫捕捉法^[43-44]、抑制性消减杂交技术(Suppressive subtractive hybridization, SSH)、荧光活化细胞分拣法(Fluorescence-activated cell sorting, FACS)、微射流技术等方法对样品中感兴趣的基因或基因组进行富集, 可以提高阳性克隆的占有率; 将亲和筛选技术和噬菌体表面展示表达产物结合起来分离 DNA 是一个高效易控制的高通量筛选方法, 在宏基因组文库中即使是稀少的 DNA 序列也可以得到富集, 有利于文库中稀有 DNA 序列的表达^[45]。

2005 年 Uchiyama 等人^[46]提出了底物诱导基因表达检出法(SIGEX), 并利用这一方法从地下水宏基因组文库中筛选到了由芳香族化合物诱导表达的基因。底物诱导基因表达检出法(Substrate-induced gene-expression screening, SIGEX)是基于功能分析的新型筛选方法, 主要用于筛检那些需要利用各种底物诱导才能表达的分解代谢型基因。该方法的原理是: 编码分解代谢途径的基因通常是由底物或代谢产物诱导表达的, 并受其附近的调控元件调控, 先用荧光活化细胞分拣法(FACS)筛去 IPTG 诱导表达绿色荧光蛋白(GFP)的细胞, 再用添加了底物的液体培养基培养克隆子, 结合 FACS 分离出由于携带了目的基因而表达 GFP 的克隆^[47]。SIGEX 筛选技术快速经济, 半自动化操作省时省力, 适用于分离序列未知的新基因, 在克隆宿主中难以检测到酶活性的新型分解代谢型基因用这一技术也可以有效地获得。SIGEX 技术的运用能增强宏基因组文库外源基因的表达能力。另外, Uchiyama 等人^[48]发现, SIGEX 技术趋向于选择那些原始宿主与 SIGEX 克隆宿主相关的宏基因片段, 他们认为使用不同的细菌作为克隆宿主可能有利于获取尚未得到开发的基因。

然而, SIGEX 技术也有其局限性: 不能筛选组成型基因; 对基因的结构与表达方向敏感, 当目的基因与转录调控元件在基因组上相距较远时, 将不能被筛选到^[49]; 当活性克隆在目的基因和 GFP 基因之间有一个转录终止子时因不表达 GFP 而无法被 FACS 筛选到^[46]。

3 宏基因组学技术的弊端、机遇与挑战

3.1 宏基因组学技术的弊端

宏基因组学技术绕过纯培养技术, 丰富了人们对环境微生物的认识, 并提供了有效的更全面开发利用微生物资源的途径。然而, 宏基因组学研究不可避免地面临着诸多挑战。

序列筛选法和功能筛选法各有不足之处, 在发现产物上都存在局限, 把这两种方法结合起来对宏基因组文库进行筛选, 可能会有利于新产物的发

现。宏基因组文库更多的是未可培养微生物的基因信息,而不是菌体本身,对这些微生物的遗传背景几乎一无所知,无从研究其基因表达的调控机制,因而没有足够的信息使基因置于有效的表达调控下,外源宏基因在宿主细胞中不表达或检测不到表达都是极其可能的,这样就使得宏基因组文库的筛选漏掉了许多环境微生物的天然产物。另外,目前尚没有标准化的环境 DNA 提取方法适用于所有的环境样品,获取环境大片段 DNA (如合成抗生素需要 100 多个基因)仍然比较困难,这些都会影响宏基因组研究的结果。

3.2 宏基因组学技术的机遇与挑战

3.2.1 大规模 DNA 测序技术的发展: 过去的几年里,随着大规模 DNA 测序平台的广泛使用, DNA 测序成本比传统测序方法减少了 2 个数量级,极大地促进了宏基因组学的发展。基于高通量测序(如 Solexa 技术能够获得海量的数据)的宏基因组学研究是环境相关微生物研究的新方法。新一代测序技术备受关注的的一个原因是它的通量持续增长,使得基因组的高覆盖率测序成为可能,有助于人们以更低廉的价格更全面、深入地分析基因组、转录组、蛋白质组的各项数据^[49]。目前,大多数大规模的宏基因组学研究依赖焦磷酸测序技术,如罗氏的 454 技术^[18,50]。鉴于微生物群体的宏基因组测序只允许重构环境中高丰度的微生物基因组, Woyke 等人^[51]利用单细胞测序方法获得了 2 个未可培养的重要的海洋微生物的高质量基因组信息。越来越多的生物全基因组序列和各种生态环境中宏基因组序列被测定并公开,与基因序列相关的数据库(如 GOLD 数据库^[52], CAMERA 数据库^[53])逐渐丰富,有效的生物信息学算法和分析软件(如发现新基因的算法^[54],宏基因组数据分析和功能预测软件^[55])的开发与应用,对宏基因组学的发展都具有积极的影响。

人体肠道中绝大多数种类的微生物是难以培养的,只有运用宏基因组学技术才能研究人类肠道中的所有微生物群落,进而了解人类肠道中细菌的物种分布,最后为后续研究肠道微生物与人类肠炎、糖尿病等疾病的关系提供理论依据(人类肠道宏基

因组计划, MetaHIT)。中国深圳华大基因关于人类肠道宏基因组计划的研究成果发表于 2010 年 3 月份的《Nature》上^[56],他们收集了 124 例欧洲人肠道菌群基因组样本,采用新一代大规模高通量测序技术进行深度测序,产出近 6 千亿的碱基序列,经过序列组装和基因注释分析,从中获得 330 万个非冗余的人体肠道宏基因组的参考基因,约是人自身基因的 150 倍,这些基因主要是目前未知的微生物基因。由这些基因信息可以估计人肠道中存在约 1 000–1 150 种细菌,平均每个体内约含有 160 种优势菌种,并且这些细菌是绝大部分个体所共有的。MetaHIT 计划的完成,将极大地推动宏基因组学研究的快速进展,使得大规模的宏基因组学研究相继展开,大量新的微生物种群和新的基因将得以发现。

然而,新一代测序技术并不是十全十美的,几乎所有的新一代的测序长度都要明显短于传统测序技术的测序长度,且测序准确率要比传统测序技术低 10 倍,我们没有一个关于环境样品各种微生物基因组 DNA 组成的确定模型,也不知道测序的样品中含有多少种不同的微生物,当宏基因组学技术应用用于环境微生物生态学研究时,根据测序结果反映的各种微生物在环境中的丰度也是有偏差的。即使 2 个不同的物种在环境中具有相同的数量,基因组大的物种在序列信息上也将表现出优势,因为测序时它产生了更多的片段^[57]。

另外,在“1 000 美元测一个基因组”的时代真正到来之前,基于高通量测序宏基因组学计划仍然是一项高成本的庞大工程。

3.2.2 宏转录组学、宏蛋白质组学的发展: 由于重复基因的存在、基因表达的时空特异性和蛋白质修饰作用等原因,复杂环境条件下环境微生物基因特异性表达及其功能,并不能通过基于 DNA 的宏基因组学的研究得到揭示,而这部分信息往往是生态环境中最重要的部分。环境转录组学方法研究环境中微生物的生物、物理、化学活性则不需要已知微生物群体中基因表达的信息,还能同时获得样品中的生物群体的结构和功能信息^[58–59]。另外, Frias-Lopez 等人^[59]利用 RNA 扩增技术研究海洋表层水的微生

物群落基因表达时发现, RNA 反转录形成的 cDNA 没有被包含在现有的数据库中, 这暗示着我们所知的海洋微生物代谢多样性仅是皮毛而已。

与宏基因组学相似, 宏蛋白质组是指环境混合微生物群落中所有生物的蛋白质总和, 宏蛋白质组学是以整个微生物群落为研究对象, 绕过了传统的纯培养技术, 了解环境中不同物种之间的相互作用及微生物如何应对环境条件变化的理想途径。宏蛋白质组学已经在少数极端生态环境的功能研究中取得了成功, 提供了一种对环境样品总蛋白质进行研究的有效方法, 在新活性物质和重要蛋白质产品开发中展现了其研究潜力^[60]。

宏基因组学提供环境中总 DNA 的信息, 宏转录组学提供实时的环境基因表达信息, 宏蛋白质组学可以提供实时状况下环境功能的信息, 把这些“组学”联系起来, 形成从基因到蛋白质的研究联系, 将极大地丰富人类对自然界的认识。

然而, 单就宏蛋白质组学技术而言, 有不少的挑战摆在我们面前, 如: 目前还没有通用的应用于环境样品蛋白质提取的方法, 特别是从海水、土壤等复杂环境样品中提取还有很大困难, 应用目前的蛋白质分离技术仅能对复杂环境样品不到 1% 的蛋白质组分进行研究; 如何分离具有生物活性的蛋白质又是一个重要问题, 已有的蛋白质数据库中的蛋白质大多来源于已培养微生物, 数据库中包含的未培养微生物蛋白质的信息非常少, 蛋白质组数据库急需丰富完善。

3.2.3 组合生物合成技术、纳米技术的发展: 组合生物合成根据天然产物的生物合成途径, 通过将不同来源的天然产物合成基因进行重组, 在微生物体内建立组合的新型代谢途径, 已被运用到非天然天然化合物开发中。纳米技术在发现新药方面也开始成熟, 通过 ISI Web of Knowledge 以“Nanotechnology”为检索词, 仅 2010 年在生物技术和应用微生物学领域就有 56 篇相关文章。

通过宏基因组学方法获得的抗生素合成基因簇可能不完整, 因而将导致无法获得完整的生物活性化合物, 将由宏基因组学技术获得的不同基因通过

组合生物合成技术重组起来, 将能有效地促进环境未可培养微生物的开发和有价值的天然产物的发现和发展。现阶段宏基因组学技术与纳米技术并没有被结合起来用于新药开发^[61], 将这两项技术相结合, 对于作为天然活性产物开发主要来源的微生物及其产物的挖掘, 可能有强大的促进作用。

4 总结

宏基因组技术是目前为止最有效的开发和利用未培养微生物资源的工具, 而其本身也在不断的成熟丰富中, 在环境微生物多样性研究和基因结构分析方面的应用已经取得重大进展, 宏基因组文库中克隆的片段越大, 其包含完整基因簇的可能性就越大, 从而为构建高质量的微生物天然产物库提供有效的途径。改进和完善宏基因组方法, 并结合更新换代的高通量测序技术, 构建特定环境样品的微生物基因图书馆, 开发和利用未培养微生物资源, 可以极大地丰富我们对微生物的认识, 促进微生物天然产物的开发。

相关研究表明, 海洋是一个巨大的新生物物种的重要来源, 新的物种意味着新的基因、新的活性产物, 世界各国和组织越来越重视海洋生物资源的开发利用。2010 年 10 月 4 日公布的“海洋生物普查”报告结果显示, 海洋生物物种总计可能约 100 万种, 其中 75 万种海洋物种是人类不甚了解、知之甚少的, 这份报告让我们看到了开发海洋生物资源的无限潜力和希望。但由于实验室无法完全模拟海洋微生物的天然生长环境, 企图通过培养的方法研究海洋微生物恐怕只能获取其皮毛而已。宏基因组学技术弥补了这一缺陷, 随着宏基因组学技术的发展与完善, 海洋这个浩瀚的有别于陆地的环境将被广泛充分地开发利用, 为人类创造出宝贵的财富。

参考文献

- [1] 杨官品, 茅云翔. 环境细菌宏基因组研究及海洋细菌生物活性物质 BAC 文库筛选[J]. 青岛海洋大学学报: 自然科学版, 2001, 31(5): 718-722.
- [2] Schmidt TM, DeLong EF, Pace NR. Analysis of a marine

- picoplankton community by 16S rRNA gene cloning and sequencing[J]. *Journal of Bacteriology*, 1991, 173(14): 4371-4378.
- [3] Handelsman J, Rondon MR, Brady SF, et al. Molecular biological accesses to the chemistry of unknown soil microbes: a new frontier for natural products[J]. *Chemistry Biology*, 1998, 5(10): 245-249.
- [4] Rondon MR, August PR, Bettermann AD, et al. Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms[J]. *Applied and Environment Microbiology*, 2000, 66(6): 2541-2547.
- [5] Gillespie DE, Brady SF, Bettermann AD, et al. Isolation of antibiotics turbomycin A and B from a metagenomic library of soil microbial DNA[J]. *Environmental Microbiology*, 2002, 68(9): 4301-4306.
- [6] Courtois S, Cappellano CM, Ball M, et al. Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products[J]. *Environmental Microbiology*, 2003, 69(1): 49-55.
- [7] Hildebrand M, Waggoner LE, Liu HB, et al. *BryA*: an unusual modular polyketide synthase gene from the uncultivated bacterial symbiont of the marine bryozoan *Bugula neritina*[J]. *Chemistry & Biology*, 2004, 11(11): 1543-1552.
- [8] Lim HK, Chung EJ, Kim JC, et al. Characterization of a forest soil metagenome clone that confers indirubin and indigo production on *Escherichia coli*[J]. *Applied and Environmental Microbiology*, 2005, 71(12): 7768-7777.
- [9] 吴杰, 李志勇, 张戎生. 海绵宏基因组文库构建及抗菌肽功能基因的初步筛选[J]. *生物技术通报*, 2006(3): 95-103.
- [10] Chen J, Zhu TJ, Li DH, et al. Construction of a metagenomic DNA library of sponge symbionts and screening of antibacterial metabolites[J]. *Journal of Ocean University of China*, 2006, 5(2): 119-122.
- [11] 陆伟, 梁爱敏, 孟秀萍, 等. 利用宏基因组文库筛选草甘膦不敏感的5-烯醇式丙酮莽草酰-3-磷酸合酶(EPSPS)基因[J]. *高技术通讯*, 2006, 16(12): 1284-1288.
- [12] 张金伟, 曾润颖. 南极深海沉积物宏基因组 DNA 中低温脂肪酶基因的克隆、表达及性质分析[J]. *生物化学与生物物理进展*, 2006, 33(12): 1207-1214.
- [13] Lee MH, Lee CH, Oh TK, et al. Isolation and characterization of a novel lipase from a metagenomic library of tidal flat sediments: evidence for a new family of bacterial lipases[J]. *Applied and Environmental Microbiology*, 2006, 72(11): 7406-7409.
- [14] 蒋云霞, 郑天凌. 天然红树林土壤微生物大片段宏基因组文库的构建[J]. *环境科学*, 2007, 28(11): 2609-2614.
- [15] Xu MX, Xiao X, Wang FP. Isolation and characterization of alkane hydroxylases from a metagenomic library of Pacific deep-sea sediment[J]. *Extremophiles*, 2008, 12(2): 255-262.
- [16] Martín-Cuadrado AB, López-García P, Alba JC, et al. Metagenomics of the deep Mediterranean, a warm bathypelagic habitat[J]. *PLoS ONE*, 2007, 2(9): e914.
- [17] Mori T, Mizuta S, Suenaga H, et al. Metagenomic screening for bleomycin resistance genes[J]. *Applied and Environmental Microbiology*, 2008, 74(21): 6803-6805.
- [18] Zehr JP, Bench SR, Carter BJ, et al. Globally distributed uncultivated oceanic N₂-fixing cyanobacteria lack oxygenic photosystem II[J]. *Science*, 2008, 322(5904): 1110-1112.
- [19] Miller DN, Bryant JE, Madsen EL, et al. Evaluation and optimization of DNA extraction and purification procedures for soil and sediment samples[J]. *Applied and Environmental Microbiology*, 1999, 65(11): 4715-4724.
- [20] 牛泽, 曾艳, 王敏, 等. 北京污灌区重金属污染土壤 DNA 提取及宏基因组文库构建[C]. 第十次全国环境微生物学术研讨会论文摘要集, 2007.
- [21] Liles MR, Williamson LL, Rodbumer J, et al. Recovery, purification, and cloning of high-molecular-weight DNA from soil microorganisms[J]. *Applied and Environmental Microbiology*, 2008, 74(10): 3302-3305.
- [22] Ouyang YC, Dai SK, Xie LW, et al. Isolation of high molecular weight DNA from marine sponge bacteria for BAC library construction[J]. *Mar Biotechnol*, 2010, 12(3): 318-325.
- [23] Martinez A, Kolvek SJ, Yip CLT, 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]. *Applied and Environmental Microbiology*, 2004, 70(4): 2452-2463.
- [24] Henne A, Daniel R, Schmitz RA, et al. Construction of environmental DNA libraries in *Escherichia coli* and screening for the presence of genes conferring utilization of 4-hydroxybutyrate[J]. *Applied and Environmental Microbiology*, 1999, 65(9): 3901-3907.
- [25] Schirmer A, Gadkari R, Reeves CD, et al. Metagenomic analysis reveals diverse polyketide synthase gene clusters in microorganisms associated with the marine sponge *Discodermia dissoluta*[J]. *Applied and Environmental Microbiology*, 2005, 71(8): 4840-4849.
- [26] Kim TK, Fuerst JA. Diversity of polyketide synthase

- genes from bacteria associated with the marine sponge *Pseudoceratina clavata*: culture-dependent and culture-independent approaches[J]. *Environmental Microbiology*, 2006, 8(8): 1460–1470.
- [27] Fieseler L, Hentschel U, Grozdanov L, et al. Widespread occurrence and genomic context of unusually small polyketide synthase genes in microbial consortia associated with marine sponges[J]. *Appl Environ Microbiol*, 2007, 73(7): 2144–2155.
- [28] Lopanik NB, Targett NM, Lindquist N. Isolation of two polyketide synthase gene fragments from the uncultured microbial symbiont of the marine bryozoan *Bugula neritina*[J]. *Applied and Environmental Microbiology*, 2006, 72(12): 7941–7944.
- [29] Ginolhac A, Jarrin C, Gillet B, et al. Phylogenetic analysis of polyketide synthase I domains from soil Metagenomic libraries allows selection of promising clones[J]. *Applied and Environmental Microbiology*, 2004, 70(9): 5522–5527.
- [30] Wawrik B, Kerkhof L, Zylstra GJ, et al. Identification of unique type II polyketide synthase genes in soil[J]. *Applied and Environmental Microbiology*, 2005, 71(5): 2232–2238.
- [31] Gupta R, Beg QK, Lorenz P. Bacterial alkaline proteases: molecular approaches and industrial applications[J]. *Applied and Microbiology Biotechnology*, 2002, 59(1): 15–32.
- [32] Huang YL, Lai XT, He XC, et al. Characterization of a deep-sea sediment metagenomic clone that produces water-soluble melanin in *Escherichia coli*[J]. *Mar Biotechnol*, 2009, 11(1): 124–131.
- [33] 贺纪正, 张丽梅, 沈菊培, 等. 宏基因组学(Metagenomics)的研究现状和发展趋势[J]. *环境科学学报*, 2008, 28(2): 209–218.
- [34] 祝心舟, 周世宁. 基于宏基因组学的脂肪酶筛选[J]. *生命的化学*, 2008, 28(5): 654–657.
- [35] Brady SF, Chao CJ, Clardy J. Long-chain *N*-acetyltyrosine synthases from environmental DNA[J]. *Applied and Environmental Microbiology*, 2004, 70(11): 6865–6870.
- [36] Cottrell MT, Moore JA, Kirchman DL. Chitinases from uncultured marine microorganisms[J]. *Applied and Environmental Microbiology*, 1999, 65(6): 2553–2557.
- [37] Yun J, Kang S, Park S, et al. Characterization of a novel amylolytic enzyme encoded by a gene from a soil-derived metagenomic library[J]. *Applied and Environmental Microbiology*, 2004, 70(12): 7229–7235.
- [38] Lee DG, Jeon JH, Jang MK, et al. Screening and characterization of a novel fibrinolytic metalloprotease from a metagenomic library[J]. *Biotechnol Lett*, 2007, 29(3): 465–472.
- [39] Henne A, Schmitz RA, Bömeke M, et al. Screening of environmental DNA libraries for the presence of genes conferring lipolytic activity on *Escherichia coli*[J]. *Applied and Environmental Microbiology*, 2000, 66(7): 3113–3116.
- [40] Steele HL, Streit WR. Metagenomics: advances in ecology and biotechnology[J]. *FEMS Microbiology Letters*, 2005, 247(2): 105–111.
- [41] 徐士庆, 胡永飞, 袁爱花, 等. 深海沉积物微生物元基因组文库来源的新的酯酶基因的克隆、表达及酶学性质[J]. *微生物学报*, 2010, 50(7): 891–896.
- [42] Craig JW, Chang FY, Brady SF. Natural products from environmental DNA hosted in *Ralstonia metallidurans*[J]. *American Chemical Society*, 2009, 4(1): 23–28.
- [43] Urbach E, Vergin KL, Giovannoni SJ. Immunochemical detection and isolation of DNA from metabolically active bacteria[J]. *Applied and Environmental Microbiology*, 1999, 65(3): 1207–1213.
- [44] Borneman J. Culture-independent identification of microorganisms that respond to specified stimuli[J]. *Applied and Environmental Microbiology*, 1999, 65(8): 3398–3400.
- [45] Cowan D, Meyer Q, Stafford W, et al. Metagenomic gene discovery: past, present and future[J]. *Trends in Biotechnology*, 2005, 23(6): 321–329.
- [46] Uchiyama T, Abe T, Ikemura T, et al. Substrate-induced gene-expression screening of environmental metagenome libraries for isolation of catabolic genes[J]. *Nature Biotechnology*, 2005, 23(1): 88–93.
- [47] Yun J, Ryu S. Screening for novel enzymes from metagenome and SIGEX, as a way to improve it[J]. *Microbial Cell Factories*, 2005, 4(1): 8–13.
- [48] Uchiyama T, Watanabe K. Substrate-induced gene expression (SIGEX) screening of metagenome libraries[J]. *Nature Protocols*, 2008, 3(7): 1202–1212.
- [49] Shendure J, Ji H. Next-generation DNA sequencing[J]. *Nature Biotechnology*, 2008, 26(10): 1135–1145.
- [50] Palenik B, Ren Q, Tai V, et al. Coastal *Synechococcus* metagenome reveals major roles for horizontal gene transfer and plasmids in population diversity[J]. *Environmental Microbiology*, 2009, 11(2): 349–359.
- [51] Woyke T, Xie G, Copeland A, et al. Assembling the marine metagenome, one cell at a time[J]. *PLoS ONE*, 2009, 4(4): e5299.
- [52] Liolios K, Mavromatis K, Tavernarakis N, et al. The Genomes On Line Database (GOLD) in 2007: status of genomic and metagenomic projects and their associated metadata[J]. *Nucleic Acids Research*, 2008, 36: D475–D479.

- [53] Seshadri R, Kravitz SA, Smarr L, et al. CAMERA: a community resource for metagenomics[J]. PLoS Biology, 2007, 5(3): e75.
- [54] Krause L, Diaz NN, Bartels D, et al. Finding novel genes in bacterial communities isolated from the environment[J]. Bioinformatics, 2006, 22(14): e281–e289.
- [55] Wooley JC, Ye YZ. Metagenomics: facts and artifacts, and computational challenges[J]. Journal of Computer Science and Technology, 2010, 25(1): 71–81.
- [56] Qin JJ, Li RQ, Raes J, et al. A human gut microbial gene catalogue established by metagenomic sequencing[J]. Nature, 2010, 464(7285): 59–65.
- [57] Gitig D. Marine metagenomics[J]. BioTechniques, 2010, 48(5): 361–365.
- [58] Poretsky RS, Bano N, Buchan A, et al. Analysis of microbial gene transcripts in environmental samples[J]. Applied and Environmental Microbiology, 2005, 71(7): 4121–4126.
- [59] Frias-Lopez J, Shi YM, Tyson GW, et al. Microbial community gene expression in ocean surface waters[J]. Proc Natl Acad Sci USA, 2008, 105(10): 3805–3810.
- [60] 郝纯, 刘庆华, 杨俊仕, 等. 宏蛋白质组学: 探索环境微生态系统的功能[J]. 应用与环境生物学报, 2008, 14(2): 270–275.
- [61] Singh BK, Macdonald CA. Drug discovery from uncultivable microorganisms[J]. Drug Discovery Today, 2010, 15(17/18): 792–799.

征 稿 简 则

1 刊物简介与栏目设置

《微生物学通报》是由中国科学院微生物研究所和中国微生物学会主办的, 以微生物学应用基础研究及技术创新与应用为主的综合性学术期刊。刊登内容包括: 工业微生物学、海洋微生物学、环境微生物学、基础微生物学、农业微生物学、食品微生物学、兽医微生物学、药物微生物学、医学微生物学、病毒学、酶工程、发酵工程、代谢工程等领域的最新研究成果, 产业化新技术和新进展, 以及微生物学教学研究和改革等。设置的栏目有: 研究报告、专论与综述、生物实验室、高校教改纵横、名课讲堂、教学与科研成果展示、显微世界、专题专栏、专家论坛、书讯、会讯等。

2 投稿方式

投稿时请登陆我刊主页 <http://journals.im.ac.cn/wswxtbcn>, 点击作者投稿区, 第一次投稿请先注册, 获得用户名和密码, 然后依照提示提交稿件, 详见主页“投稿、征稿须知”。

作者必须在网站投.doc 格式的电子稿, 图与文字编好页码、图号后合成一个文件上传。凡不符合(投稿须知)要求的文稿, 本部恕不受理。

3 写作要求

来稿要求论点明确, 数据可靠, 简明通顺, 重点突出。

3.1 篇幅

以 A4 纸 5 号字计算, 综述、教学和方法类文章最好在 6 页以内, 研究报告 4–8 页(以上均包括图表)。

3.2 图表

文中的图表须清晰简明, 文字叙述应避免与图表重复。所有小图的宽度应小于 8 cm (占半栏), 大图的宽度应小于 17 cm (通栏)。

3.3 参考文献及脚注

参考文献按文内引用的先后顺序排序编码, 未公开发表的资料请勿引用。我刊的参考文献需要注明著者(文献作者不超过 3 人时全部列出, 多于 3 人时列出前 3 人, 后加“等”或“et al.”, 作者姓前、名后, 名字之间用逗号隔开)、文献名、刊名、年卷期及页码。国外期刊名可以缩写, 但必须标准, 不加缩写点, 不用斜体。参考文献数量不限。

参考文献格式举例:

期刊: [1] 刘杰, 成子强, 史宣玲. SARS 冠状病毒 *nsp14* 基因的克隆和表达[J]. 微生物学通报, 2007, 34(2): 1–3.

[2] Kajiura H, Mori K, Tobimatsu T, et al. Characterization and mechanism of action of a reactivating factor for adenosylcobalamin-dependent glycerol dehydratase[J]. J Biol Chem, 2001, 276(39): 36514–36519.

图书: [3] 钱存柔, 黄仪秀. 微生物实验教程[M]. 北京: 北京大学出版社, 2000: 4.

[4] 董志扬, 张树政, 方宣钧, 等. 海藻的生物合成及抗逆机理//华珞等. 核农学进展[M]. 北京: 中国农业出版社, 1996: 115–120.

脚注(正文首页下方):

基金项目: 基金资助(No.)

*通讯作者: Tel: ; Fax: ; E-mail:

收稿日期: 2011-00-00; 接受日期: 2011-00-00

(下转 p.622)