

构建高质量微生物天然产物库研究策略

边疆^{1,2,3**}, 宋福行^{2**}, 张立新^{1,2*}

(¹中国科学院南海海洋研究所, 广州 510301) (²中国科学院微生物研究所, 北京 100101)

(³中国科学院研究生院, 北京 100049)

摘要: 微生物次级代谢产物历来是天然药物的重要来源, 过去曾被称为“生物沙漠”的海洋, 由于从中分离到大量新的微生物、基因及生物活性化合物, 被重新认识逆转而成为一种“生物多样化的热带雨林”。构建一个高质量微生物库及其天然产物库是保证药物和其他筛选成功的前提和关键。但如何高效建立高质量微生物天然产物库仍面临很多瓶颈问题。我们拟从:(1) 扩大可培养微生物的多样性及去重复化;(2) 扩大基因资源多样性及去重复化;(3) 扩大微生物次级代谢产物多样性及去重复化;(4) 寻找崭新次级代谢产物的新技术新方法, 特别是针对多靶位药物的高通量互动筛选方面提出应对的研究策略。利用上述化学微生物学策略分离生物活性化合物不仅在生物技术和制药应用中显现重要性, 也增加了我们对微生物的多样性、生态系统功能和应用生物学的理解。

关键词: 微生物天然产物库; 化学微生物学; 宏基因组技术; 去重复化; 高通量互动筛选

中图分类号: Q935 文献标识码: A 文章编号: 0001-6209 (2008) 08-1132-07

自然界是最好的化学家, 微生物天然化合物 (Microbial Natural Products, MNP) 是自然界赐予我们的宝藏, 至今人们仍然不断从中发现崭新结构化合物,^[1]。1981年1月~2006年6月世界上天然来源的临床药物占总药物的42.5%, 在抗癌和抗感染类药物中更是占到了47%和52.2%^[2]。但20世纪80年代末, 跨国制药公司纷纷减少对MNP研发的投入, 以MNP为基础的药物研发面临着严重考验。主要原因如下:

建立和保存高质量MNP库, 无法实现工业化; MNP往往产量较低, 且混于提取物中, 分离难度大;

缺乏有效的去重复方法, 导致已知化合物复筛率高; 去重复和纯化步骤繁琐耗时, 周期长; MNP往往复杂度高, 难于化学合成。近年来随着高通量制备和筛选、组合化学、计算化学等技术的进步和成熟以及人类基因组计划的完成, 天然药物研发进入复兴

期。在MNP研发过程中, 构建一个化合物结构多样性丰富, 重复率低的MNP库是一切工作成功的基石。

在MNP筛选中新微生物菌种意味着新基因、新代谢途径的存在, 获得新结构次级代谢产物的可能性也大。所以在MNP研究中, 新菌种新基因资源的开发是获得高质量MNP库的重要基础。陆生特殊生境 (高温、高压、高盐等) 以及海洋环境和动植物内生环境, 经过千百万年的自然进化, 微生物生态多样性特殊, 存在新菌种、新基因、新代谢途径的几率大, 产生新的次级代谢产物的可能性也大^[3]。海洋生境不同于陆地, 海洋微生物次级代谢产物的生物合成途径和酶反应系统与陆地生物相比有着巨大的差异, 导致一些化学结构奇特、生物活性多样的化合物的产生^[4]。海洋放线细菌 (*Actinobacteria*) 分布广泛^[5], 新结构和新活性的化合物不断从中发现, 可培养海洋

基金项目: 国家“863计划”(2006AA09Z402, 2007AA09Z443), 中国科学院创新基金(KSCXZ-YW-G-013), 国家“973项目”(2007CB707802)

**对文章具有同等贡献。

*通讯作者。Tel/Fax: +86-10-62566511; Zhanglixin@im.ac.cn

作者简介: 边疆(1978-), 男, 天津人, 博士研究生, 主要从事海洋微生物资源及药用天然产物研究。E-mail: Border78@163.com;

宋福行(1976-), 男, 助研, 河北饶阳人, 主要从事天然药物研究。E-mail: Songfuhang@163.com

收稿日期: 2008-01-26; 修回日期: 2008-04-06

放线细菌已经逐步成为了海洋微生物药物的研究热点。

我们已经建立了一个以海洋微生物菌种分类及基因信息(X轴),诱导次级代谢产物条件(Y轴)和通过诱变/基因工程改造菌株获得非天然的MNP及活性信息(Z轴)的三维MNP库及信息库^[6]。当发现新活性或新结构化合物时,可以通过信息库回到同一微生物类群,测试其他菌株,进一步找到产生同类化合物的高产或具有其他优良性状的菌株。

本文重点从扩大可培养微生物的多样性、扩大基因资源多样性、扩大微生物次级代谢产物多样性、寻找崭新结构及活性化合物和同步去重复化等方面,针对打破天然药物研发过程中的瓶颈提出应对策略。

1 扩大可培养微生物的多样性及去重复化

新菌种-新基因-新化合物的创新思路是以微生物分类学为指导,利用选择性培养条件,通过传统和创新两条分离途径复苏和纯化那些珍贵的被认为是难以培养的微生物。

1.1 通过传统培养和創新培养方法从海洋中分离微生物新种

海洋微生物分离和培养的难点在于:海洋微生物的耐氧能力不同;代谢底物对细胞有不同程度抑制;饥饿引起细胞的吞噬作用;实验室培养缺乏细胞间的信号交流等。为了增加对难培养微生物的分离,稀释分离样品,采用寡营养的培养方式和加入细胞间信号分子都是可行的解决方案^[4, 6]。基于对样品的高度稀释和寡营养培养技术,Connon和Gontang分离得到了很多新菌种^[7]。

随着环境宏基因组技术的不断提高,利用16S rDNA克隆文库研究环境微生物分类的方法越来越成熟^[5]。Rappe通过DNA检测,发现SAR11类细菌在海洋中大量存在,设计该菌种的特异荧光探针,通过特异的培养方法,使该菌得到增殖达到分离并纯培养^[8]。Stevenson设计了针对分离目标*Acidobacteria*和*Verrucomicrobia*的PCR引物,确定培养目标,然后进行培养筛选,获得了目标微生物的纯培养。极大地减少了培养过程中的盲目性,提高了分离效率^[9]。

Zengler建立的一种将微生物分散包裹到微滴中的高通量培养方法,可以保证微生物产生代谢产物/信号因子在模拟的自然环境中相互交换,经过长时

间培养,获得多种以前认为是不可培养的海洋新种细菌,有些甚至是新目级别的菌种^[10]。Toledo利用包裹无菌的海绵和海底沉积物浸出液提高了分离效率,经过5周培养,分离得到的细菌中42%是新种,远高于用包裹海水微滴的分离效率^[11]。

1.2 微生物菌株去重复化及化学微生物学

MNP的筛选和相应活性化合物的纯化分离是一个高消耗、低产出的工作。前期做好菌株和产物的去重复化,才能保证后期工作中避免分离得到已知化合物。传统的菌种多相分类鉴定的方法可以很准确的进行菌株去重复化,但是分类鉴定周期长,损耗大量人力物力。我们利用分类学家的经验和种属特异性引物(见表1)可以快速鉴定一些常见微生物,然后根据初步分类结果,对同种的微生物通过后续研究[例如,分析限制性片段长度多态性(Restriction Fragment Length Polymorphism, RFLP)等]进一步鉴定至亚种级别。针对那些需要深入研究的菌株,再采用多相分类的手段进行细致分类鉴定。

表1 部分用于快速鉴定放线细菌类群的16S rDNA特异性引物

Table 1 Part of the 16S rDNA Actinobacteria genus-specific primers for quick identification

Genus	Primer sequences	Size /bp	Ref.
<i>Actinobacteria</i>	CGCGGCCTATCAGCTTGTGTTG CCGTACTCCCCAGGCGGGG	640	Stach <i>et al.</i> ,2003
<i>Amycolatopsis</i> spp.	GGTGTGGGCGACATCCAGTTGT GCTGGTACAGAGGGCTGCGATA	450	Tan <i>et al.</i> ,2006
<i>Micromonosporaceae</i>	SAGAAGAAGCGCCGGCC CCAGCCCCACCTTCGAC	1000	Pisano <i>et al.</i> ,1989
<i>Pseudonocardia</i> spp.	GTGGAAAGTTTTTCGGCTGGGG GCGGCACAGAGACCGTGAAT	640	Moron <i>et al.</i> ,1999
<i>Streptomycetaceae</i>	GGTGGCGAAGCGGA GAACTGAGACCGCTTTTTGA	600	Monciardini <i>et al.</i> ,2002
<i>Streptosporangiaceae</i>	GACGAARNTGACGTGTA CGTTGCGTCTAATTAAGCAA	500	Hayakawa <i>et al.</i> ,1989
<i>Thermononosporaceae</i>	GGGAGAATGGAATCCC CCCCACCTTCGACC	800	Athalye <i>et al.</i> ,1981
<i>Streptomonospora</i>	TCTGTGCGGTTGACGGTAG CGAAGGCACCCECCGTCTC	565	Zhi <i>et al.</i> ,2007
<i>Nesterenkonia</i>	CGCATAGGTGCTGGTGAAG GAGGTCGGGTTGCAGACTTCG	1120	
<i>Actionopolyspora</i>	CGGTGAGGGCGTACCAAGGCGATG AGCCCTGCCGTCAATTCGTCA	209	Zhi <i>et al.</i> ,2007
<i>Saccharomonospora</i>	ACGGCACGGGACACGTGCACAGC CGTCTGCCGTGAAAACCTGCGGC	266	Salazar <i>et al.</i> ,2000
<i>Nocardiosis</i>	TAAATGACCTCACATCTCT TCTCTGGGGTTGACGGTAG	575	Salazar <i>et al.</i> ,2002
<i>Saccharothrix</i>	TCGACCGCAGGCTCCACG AAGGCCCTTCGGGGTACACGAG	545	Salazar <i>et al.</i> ,2002
<i>Saccharopolyspora</i>	GAGGAACCCATCCCCACACC GGTGACGGTAGGTGTAGAAG	371	Moron <i>et al.</i> ,1999

除了通过采用传统方式从菌种分类角度去重复化,我们还应用化学微生物学,从化学角度解析微生物的规律和进化关系。相同种属微生物往往存在相似

的化学特征。究其根本是由于分类学上近缘微生物具有某些保守的基因造成的。在菌种库和产物库的高通量构建中,很难做到对每个入库的菌种进行分类鉴定。而从微生物次级代谢产物入手,比对 TLC、HPLC、MS、NMR 上的指纹图谱,寻找出一个属或更大的分类类群中的代表性化合物,快速确定其分类学地位并且排除重复。最近,我们在研究一株海洋来源康氏木霉(*Trichoderma koningii*)时,将 Koninginin 类化合物作为该种属的特征产物,快速指导我们分类鉴定并发现新的该类化合物(另文发表)。

2 扩大基因资源多样性并去重复化

通过外显荧光显微镜和 rDNA 测序策略对微生物群落分析揭示了能被培养的只代表了微生物的一小部分,高达 99% 以上的微生物是尚未培养或难培养的,宏基因组(Metagenome)绕过了很多微生物难培养或不可培养这个瓶颈,利用宏基因组技术,可直接从环境中提取具有丰富多样性的微生物基因组;然后将大片段环境 DNA 克隆到适当的载体上,转入宿主,构建成宏基因组文库;或利用分析保守序列设计探针或简并引物,通过序列分析找到合成功能基因簇(PKS, NRPS 等),预测活性产物的结构等;抑或直接通过一定的活性筛选模型得到化合物^[12]。Venter 等利用高通量 DNA 测序和生物信息学的方法从北大西洋百慕大群岛附近马尾藻海域 1500 L 海水中鉴定了 120 多万个微生物新基因^[13]。在世界上最贫瘠的水体之一收集到的如此小的样品中获得如此大量的新基因,对海洋微生物分子生态学和进化生物学的新兴领域提出了重大挑战。

原核生物中 MNP 合成的酶基因和调控基因一般成簇分布在一条染色体上,通过宏基因组技术可以获得活性 MNP 的整个操纵子(Operon)。Kosan 公司成功的建立了一套利用组合各类 PKS (polyketide synthase) 基因的方法,构建聚酮类化合物库^[14],并被 NIH 用于药物筛选。宏基因组作为一项新兴技术,具有广泛的应用前景,但其本身同时也有不少技术难题需要去突破。当前存在的技术瓶颈主要有: 如何获得大片段且无污染的基因组 DNA; 异源基因的表达效率以及宿主菌的选择; 文库的去重复化; 如何针对文库进行筛选?

宏基因组文库中克隆片段越大,所包含整条合成基因簇的可能性就越大。当前,eDNA 的提取方法主

要可归为两大类:直接法和间接法,直接提取法就是直接从未处理样本(如土壤、粪便、沉积物、水体)中裂解细胞,抽提 eDNA,这一方法的好处在于 eDNA 的恢复率高,通常高于间接法 100 倍,但是其缺点在于 DNA 剪切严重,通常得到的 DNA 片段较小。间接提取法即将细胞从环境样本中分离出来后,再裂解,缺点在于 eDNA 抽提不完全,需要进行富集,通常样本消耗量较大。但其优点在于样本量充足时,可以得到较大的 DNA 片段,甚至高达 400 kb^[15]。

杂质的存在如:多酚类化合物(polyphenols)和腐殖酸(humic acids),会阻碍 Taq DNA 聚合酶的作用,干扰限制性酶的作用,降低转化效率,造成实验失败。Zhou 等人于 1995 年建立利用 CTAB(十六烷基三甲基溴化铵)和 PVPP(聚乙烯吡咯烷酮)去除腐殖酸等干扰物的方法被大家广泛应用^[16]。我们发现,对于海底沉积物来说,用琼脂糖凝胶电泳辅以电洗脱的方法组合取得了不错的效果(另文发表)。

异源基因的表达效率和宿主的选择也是一个棘手的问题。目前常用的宏基因组宿主菌有大肠杆菌(*Escherichia coli*)、链霉菌(*Streptomyces* spp.)、恶臭假单胞菌(*Pseudomonas putida*)。其中大肠杆菌作为表达受体适合低 GC 含量启动子小片段基因的表达;链霉菌有适合外源基因簇表达的分子基础^[17],很多链霉菌本身都有 PKS 和 NRPS(Non-ribosomal peptide synthase)的合成基因簇,有外源 PKS 和 NRPS 的表达调控基础,又可能发生宿主和外源的功能基因簇杂合共表达,产生自然界从未有过的新型结构^[18]。

对于宏基因组水平的去重复化,RFLP 和 DGGE(Denaturing gradient gel electrophoresis)的手段都是很好的方法。我们推荐 DGGE 的方法,因为即使 DNA 序列上只出现一个碱基的变动,其变性解链温度都会不同,迁移距离也会不同,可以清晰的分辨出不同的序列,同时有利于节省测序费用和提高文库质量^[19]。

在建立宏基因组文库后,还存在筛选策略的问题。基于功能的筛选(functional-based screening)利用一些特异性的指示培养基或带有报告基因的载体,挑选阳性克隆测序,分析并找出目的基因。基于序列的筛选(sequence-based screening)利用简并引物或特异性引物扩增 eDNA 以获得高丰度的目的基因,转入宿主进行进一步的筛选^[20]。2006 年,Roh 和 Kim 等人利用基于序列的筛选从宏基因组文库中找出了新颖的细胞色素 b5 基因和 P450 单加氧酶基因^[21,22]。此

外,还有一些衍生出来的新策略,如 Uchiyama 等人建立的底物诱导基因表达筛选(substrate-induced gene expression)^[23]。

3 扩大微生物次级代谢产物多样性及去重复化

通过分析微生物基因组数据发现,参与次级代谢的基因簇的数量远远超过现在得到的次级代谢产物的数量^[24]。为了扩大微生物次级代谢产物的多样性,传统的方法是根据一株菌多种次级代谢产物(One strain many compounds, OSMAC)的原则,优化培养条件刺激微生物产生更广泛的次级代谢产物。然而这种方法具有很大盲目性,并且可能只得到部分基因簇的代谢产物。另一方面我们可以直接从基因入手, MNP 合成的相关基因聚合成簇并且相对保守^[25,26],通过筛选、测序、定位基因组资源来预测和表达 MNP 的合成基因簇,可以针对性地获得结构和活性新颖的次级代谢产物。

通过 genome mining 和基因组扫描(Genome scanning)能揭示新的基因簇进而获得新的 MNP。Genome mining 的方法体现了一种先基因后产物的研究思路,通过 genome mining 我们可以得到新的 MNP 基因簇并根据基因簇的保守性,预测其最终产物的结构从而有利于其分离纯化。Lautru 等人利用 genome mining 在 *Streptomyces coelicolor* M145 中发现了一个新基因簇 CCH,在缺铁的培养基中能合成螯合铁离子的多肽 Coelichelin^[27]。

与 genome mining 需要全基因组序列相比,基因组扫描(Genome Scanning)则只需要一系列的 GST 序列就能探测到目的基因簇。因为基因簇大小一般都在 20~100 kb 之间,只要 GST 足够多就能将所有的代谢物基因簇全部覆盖到,这样通过某基因簇保守的 GST 作为探针筛选含有这些保守序列的微生物,进而获得其完整的基因簇。基于这种原理,Zazopoulos 等人成功阐明了 dynemicin 和 macromomycin 合成基因簇;同时以这些基因探测到 70 株放线菌中有 11 株含有沉默的 enediyne 合成基因簇,通过优化培养基等条件,最终诱导产生了结构未知的 enediyne 物质^[28]。

如何激活沉默基因簇,打破表达瓶颈也是一个亟待解决的问题。Hertweck 等人通过 genome mining 的方法在 *Apergillus nidulans* 中发现了一个新的

PKS-NRPS 杂合基因簇 ApdA,通过将该基因簇的激活基因 ApdR 过量表达使整条沉默的 PKS-NRPS 杂合聚合基因簇特异表达,最终得到了该基因簇的两个新产物,分别命名为 aspyridones A 和 aspyridones B^[29]。这是首次在真核生物中通过过量表达调节基因而使整条沉默基因簇表达,从而扩大了微生物次级代谢产物的多样性。

4 寻找崭新次级代谢产物的新技术新方法

在进行 MNP 研究时,获得新结构或新活性的化合物是研究的目的。结构排重主要指将代谢产物结构信息与相关数据库(如 Dictionary of Natural Products, Antibase, MarinLit)比较,去除已经发现的化合物,达到获得全新次级代谢产物的目的;活性追踪是指对构建的微生物 MNP 库分别进行抗癌症,糖尿病,感染性疾病和心脑血管疾病等高通量生物活性筛选,迅速追踪活性成分,借助现代波谱和数据库比对技术,进一步获得作用浓度低、活性强、毒性低的活性化合物。LCMS-IT-TOF 的出现,实现了离子阱质谱和飞行时间质谱的完美结合,实现多级质谱分析,提高了化合物结构分析的精度和速度。我们期待着 LC-DAD-NMR-IT-TOF 的联用,这一崭新的分析手段必将在 MNP 去重复化和发现崭新结构方面发挥优越作用。

在化学家们一味寻找崭新结构的同时,也不应忘记我们已经从自然界中获得了数百万种化合物,这些化合物的功效并未开发完全,目前越来越多的已知化合物被赋予新的活性。随着分子靶位研究的深入,以靶位为基础的药物研发模式已经建立^[30]。BCR-ABL 酶抑制剂 Gleevec (imatinib mesylate)、COX-2 抑制剂 Vioxx (rofecoxib)和 Celebrex (celecoxib)都是上述药物研发模式的成功例子。

一种以系统生物学为基础的多靶位药物研发模式正成为新趋势。Parsons 等研究了酵母的化学-遗传相互作用(chemical-genetic interaction)体系^[31]。利用这种体系,我们可以直接从活性粗提物水平确定其活性靶位,并通过与已知作用于该靶位化合物比较,进行快速排重。

药物相互之间作用效果被区分为互动(synergy)相加(additivism)和拮抗(antagonism)^[32]。Alexis 选用 1000 种已知化合物,组成 5000 种组合,通过构建的联合筛选模型(Combination High-Throughput

Screening) 筛选得到很多有药用潜力的组合物^[33]。我们提出了多靶位药物的互动筛选模式, 针对那些专利过期、有一定毒副作用、生产利润低但有一定竞争力的药物(药物 A), 通过降低药物 A 的使用浓度(大大降低毒副作用), 从多种药物(传统中药、天然药物、合成药物以及成药等)中筛选出能与药物 A 互动增效的、作用于不同的靶位点的新型药物, 使其作用远远大于单个药品作用的加和, 产生互动效应^[34,35]。这种互动药物增强了药物的疗效作用, 从而可大大降低药物的剂量, 减小了药物的毒副作用, 经过审批流程可以重新进入市场。利用通过互动筛选理念构建的高通量互动筛选模型(High-throughput synergy screening), 我们从构建的微生物 MNP 库中筛选得到了多种有互动效果的先导化合物。

5 展望

过去的几年经历了活性 MNP 领域的信息爆炸, 因此回顾 MNP 成为药物的成功, 考察将传统方法和新技术结合进一步发现更多微生物药物是非常及时的。但是, 未来的成功不是新与旧的对决, 而是依赖于学习如何运用已经存在的方法学, 包括基因组学、蛋白质组学、组合化学、DNA 重组、组合生物合成、生物多样性、生物信息学和高通量筛选, 来快速评估来自微生物的提取物和纯化合物的活性。基因组扫描技术使现代 MNP 研究制备逐步脱离了盲目性, 结合计算机模拟技术已经可以初步模拟代谢产物结构, 用于指导 MNP 分离和筛选^[36]; 比较基因组学(Comparative Genomics)的发展会在不久的将来解析出更多丰富多彩的微生物代谢机制; MNP 高通量制备系统的不断完善大大缩短了产物制备的时间, 提高了分离效率; 各种分析检测仪器不断突破极限, 使科学家推陈出新的眼光更加锐利。截至 2007 年, 67 种微生物来源的新药进入到了临床的各个阶段^[37]。我们期待着更多的 MNP 药物为人类健康造福。同时, 建立一个来源广泛的微生物菌种库和高质量的 MNP 库, 也是在为新的药物研发、生物安全、生物再生能源开发、环境保护以及生物食品等方面提供丰富的资源储备。微生物代谢物的生物多样性、创造性和多功能性正在通过依赖或不依赖培养基的方法进行研究。我们开始探索许多重要的问题, 这在几年前还远远超过我们的认知范围, 其中包括自然环境中微生物及其代谢物的功能作用和自然界中如此丰富微生物的生

态和进化过程。

致谢 感谢中国科学院微生物研究所刘志恒教授, 中国科学院沈阳生态研究所王书锦教授, 美国 Emory 大学付海安教授, 首都医科大学刘宏伟教授等对文章的审阅及指导。感谢张立新团队成员对文章的贡献和支持。

参 考 文 献

- [1] Demain AL, Zhang L. Natural products and drug discovery. In: Zhang L and Demain AL. eds. *Natural Products: Drug Discovery and Therapeutics Medicines*. 1st. USA: Humana Press, 2005, 3–32.
- [2] Newman DJ, Cragg GM. Natural products as sources of new drugs over the last 25 years. *J Nat Prod*, 2007, 70(3): 461–477.
- [3] Pedrós-Alió C. Genomics and marine microbial ecology. *Int Microbiol*, 2006, 9(3): 191–197.
- [4] Zhang L, An R, Wang J, et al. Exploring novel bioactive compounds from marine microbes. *Curr Opin Microbiol*, 2005, 8(3): 276–281.
- [5] Ward AC, Bora N. Diversity and biogeography of marine actinobacteria. *Curr Opin Microbiol*, 2006 9(3): 279–286.
- [6] Knight V, Sanglier JJ, DiTullio D, et al. Diversifying microbial natural products for drug discovery. *Applied Microbiology and Biotechnology*, 2003, 62(5–6): 446–458.
- [7] Gontang EA, William F, Jensen PR. Phylogenetic diversity of Gram-positive bacteria cultured from marine sediments. *Appl Environ Microbiol*, 2007, 73(10): 3272–3282.
- [8] Michael SR, Stephanie AC, Kevin LV, et al. Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. *Nature*, 2002, 418(6898): 630–633.
- [9] Stevenson BS, Eichorst SA, Wertz JT, et al. New strategies for cultivation and detection of previously uncultured microbes. *Applied and Environmental Microbiology*, 2004, 70(8): 4748–4755.
- [10] Zengler K, Toldeo G, Rappe M, et al. Cultivating the uncultured. *Proc Natl Acad Sci USA*, 2002, 99(24): 15681–15686.
- [11] Toledo G. High Throughput cultivation for isolation of novel marine microorganisms. *Oceanography*, 2006, 19(2): 120–124.
- [12] Elend C, Schmeisser C, Hoebenreich H, et al. Isolation and characterization of a metagenome-derived and cold-active lipase with high stereospecificity for (R)-ibuprofen esters. *J Biotechnol*, 2007, 130(4): 370–377.
- [13] Venter JC, Remington K, Heidelberg JF, et al. Environmental genome shotgun sequencing of the Sargasso Sea. *Science*. 2004, 304(5667): 66–74.
- [14] Chaitan K, Ashley G, Hong F et al. Combinatorial polyketide libraries produced using a modular PKS gene clusters as scaffold. USA: US 6558942. 2003.
- [15] Roh C, Francois V, Schmid RD, et al. Comparative study of methods for extraction and purification of environmental DNA from soil and sludge samples. *Appl Biochem Biotechnol*, 2006, 134(2): 97–112.
- [16] Zhou J, Bruns MA, Tiedje JM. DNA recovery from soils of diverse composition. *Appl Environ Microbiol*, 1996, 62(2): 316–322.
- [17] Martin JF. Phosphate control of the biosynthesis of antibiotics and other secondary metabolites is mediated by the PhoR-PhoP system: an unfinished story. *J Bacteriol*, 2004, 186(16): 5197–5201.
- [18] Brady SF, Clardy J. Cloning and heterologous expression of isocyanide biosynthetic genes from environmental DNA. *Angew Chem Int Ed Engl*. 2005, 44(43): 7063–7065.
- [19] Babcock DA, Wawrik B, Paul JH, et al. Rapid screening of a large insert BAC library for specific 16S rRNA genes using

- TRFLP. *J Microbiol Methods*, 2007, 71(2): 156–161.
- [20] Lorenz P, Liebeton K, Niehaus F, *et al.* Screening for novel enzymes for biocatalytic processes: accessing the metagenome as a resource of novel functional sequence space. *Curr Opin Biotechnol*, 2002, 13(6): 572–577.
- [21] Kim BS, Kim SY, Park J, *et al.* Sequence-based screening for self-sufficient P450 monooxygenase from a metagenome library. *J Appl Microbiol*, 2007, 102(5): 1392–1400.
- [22] Roh C, Villatte F, Kim BG, *et al.* Screening and purification for novel cytochrome b5 from uncultured environmental microorganisms. *Lett Appl Microbiol*, 2007, 44(5): 475–480.
- [23] Yun J, Ryu S. Screening for novel enzymes from metagenome and SIGEX, as a way to improve it. *Microb Cell Fact*, 2005, 4(1): 8.
- [24] Farnet CM, Zazopoulos. Improving drug discovery from microorganisms. In: Zhang L and Demain AL. eds. *Natural Products: Drug Discovery and Therapeutics Medicines*. 1st. USA: Humana Press, 2005, 95–106.
- [25] 任立成, 鲍时翔. 多聚酮的微生物合成及其多样性研究进展. *微生物学通报(Microbiology)*, 2007, 34(1): 127–131.
- [26] Zheng HM, Pan JW, Zhu MY. The progress of the nonribosomal polypeptide synthases. *Prog Biochem Biophys*, 2002, 29(5): 667–669.
- [27] Lautru S, Deeth RJ, Bailey LM, *et al.* Discovery of a new peptide natural product by *Streptomyces coelicolor* genome mining. *Nat Chem Biol*, 2005, 1(5): 265–269.
- [28] Zazopoulos E, Huang K, Staffa A, *et al.* A genomics-guided approach for discovering and expressing cryptic metabolic pathways. *Nat Biotechnol*, 2003, 21(2): 187–190.
- [29] Bergmann S, Schumann J, Scherlach K, *et al.* Genomics-driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. *Nat Chem Biol*, 2007, 3(4): 213–217.
- [30] Gibbs JB. Mechanism-based target identification and drug discovery in cancer research. *Science*, 2000, 287(5460): 1969–1973.
- [31] Parsons AB, Lopez A, Givoni IE, *et al.* Bioactive compounds by chemical-genetic profiling in yeast. *Cell*, 2006, 126(3): 611–625.
- [32] Berenbaum MC. What is synergy? *Pharmacol Rev*, 1989, 41(2): 93–141.
- [33] Borisy AA, Elliott PJ, Hurst NW, *et al.* Systematic discovery of multicomponent therapeutics. *Proc Natl Acad Sci U S A*, 2003, 100(13): 7977–82.
- [34] Zhang L. Screening for synergistic compounds. PCT patent WO 2005/051303. 2005.
- [35] Zhang L, Yan K, Zhang Y, *et al.* High-throughput synergy screening identifies microbial metabolites as combination agents for the treatment of fungal infections. *Proc Natl Acad Sci U S A*, 2007, 104(11): 4606–4611.
- [36] Wilkinson B, Micklefield J. Mining and engineering natural-product biosynthetic pathways. *Nat Chem Biol*, 2007, 3(7): 379–386.
- [37] Butler MS. Natural products to drugs: natural product derived compounds in clinical trials. *Nat Prod Rep*, 2005, 22(2): 162–195.

Strategies on the construction of high-quality microbial natural product library—A review

Jiang Bian^{1,2,3**}, Fuhang Song^{2**}, Lixin Zhang^{1,2*}

¹South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou 510301, China)

²Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China)

³Graduate University of Chinese Academy of Sciences, Beijing 10049, China)

Abstract: The microbial secondary metabolites are always the main source of the natural drugs. The historical paradigm of the deep ocean as a biological ‘desert’ has shifted to one of a ‘rainforest’ owing to the isolation of many novel microbes and their associated bioactive compounds. A high quality microbial and its natural product library are crucial for successful drug and other screenings. However, how to build up the library efficiently is still faced with many bottlenecks. To overcome the difficulties and limitations, we reviewed the following strategies: (1) diversifying microbial sources and dereplication; (2) diversifying gene sources and dereplication; (3) diversifying microbial metabolite sources and dereplication; (4) novel methods and technologies for bioactive secondary metabolites, especially the high-throughput synergy screening for multi-target drugs. Bioactive compounds isolated using the above chemical microbiology strategies have not only shown importance in biotechnological and pharmaceutical applications but also increased our understanding of the diversity of microbe, ecosystem functions and the exploitable biology.

Keywords: microbial natural product library; chemical microbiology; metagenome; dereplication; high-throughput synergy screening

Supported by the National Programs for High Technology Research and Development of China (2006AA09Z402, 2007AA09Z443), the Chinese Academy of Sciences Innovation Projects (KSCXZ-YW-G-013) and the Major Project of Chinese National Programs for Fundamental Research and Development (2007CB707802)

**These authors contributed equally to this work.

*Corresponding author. Tel/Fax: +86-10-62566511; E-mail: Zhanglixin@im.ac.cn

Received: 26 January 2008/ Revised: 6 April 2008