

宏基因组学技术在微生物功能酶开发中的应用

金浩 李柏林 潘迎捷 欧杰 陈兰明

(上海海洋大学食品科学与技术学院, 上海 201306)

摘要: 宏基因组学技术以特定环境样品中微生物复杂群落的基因组总和为研究对象,突破了传统微生物纯培养方法的局限,为不可培养微生物中丰富的基因资源的开发和利用提供了强有力的工具,已经取得了令人瞩目的研究进展。对宏基因组学技术及其在微生物功能酶新基因发现中的应用进行综述。

关键词: 宏基因组学 功能酶筛选 环境微生物

Metagenomics Technology Application in Microbial Functional Enzymes

Jin Hao Li Bailin Pan Yingjie Ou Jie Chen Lanming

(College of Food Science and Technology, Shanghai Ocean University, Shanghai 201306)

Abstract: Metagenomic technology is applied in studying the total genomes of complex microbial community in specific environment. It overcomes the limitations of conventional methods of microorganism cultivation, and provides powerful techniques to explore and utilize rich gene resources of unculturable microorganisms. Remarkable progress has been achieved in this research field. This paper reviewed the theory and application of metagenomic technology in screening novel functional enzymes from microorganisms.

Key words: Metagenomics Functional enzyme screening Environmental microorganisms

酶是由活细胞产生的生物催化剂,高效、专一,且需要在适宜的条件下才能发挥相应的作用,酶的这些特性和各种各样的生物催化功能,对人类及其生活都有极大的影响。在工业上,酶催化技术与传统化学过程相比可改变化工行业高污染、高能耗的现状。美国提出到2020年,通过生物催化技术,使化学工业的原料消耗、水资源消耗、能量消耗各降低30%,污染物的排放和污染扩散减少30%^[1]。酶催化技术的研发和应用有利于新型清洁高效化工生产过程的实现和工业可持续发展。

微生物蕴含着极其丰富的基因资源。自然界中的微生物分布在各种环境,甚至高温、高压、高盐碱等极端环境中,据估计数量多达 10^{30} ^[2]。然而,其中99%的微生物在实验室现有条件下不能培养,

它们的基因组编码大量未被开发的新型酶和代谢途径基因^[3]。宏基因组学(metagenomics)绕开环境微生物研究过程中培养技术局限的极大阻碍,发展成为不可培养或未培养微生物研究的重要工具。目前,宏基因组学技术在复杂环境尤其是极端环境中,在微生物代谢途径新基因和新的生物活性物质的发现方面得到了广泛的应用^[4, 5],已取得了令人瞩目的研究进展,为人类开发利用自然环境中的微生物基因资源开辟了新的途径。

1 宏基因组学技术

宏基因组学也称为环境基因组学(environmental genomics)或群落基因组学(community genomics)。最初由Handelsman等^[6]于1998年提出,以生态环境中全部微小生物的基因组(the genomes of the total

收稿日期:2012-05-07

基金项目:上海市科学技术委员会项目(09320503600)

作者简介:金浩,女,硕士,研究方向:生物分子;E-mail: jhkingool@sina.com

通讯作者:陈兰明,女,博士,教授,研究方向:食品科学与工程;E-mail: lmchen@shou.edu.cn

microbiota found in nature) 作为研究对象, 针对特定环境样品中微生物复杂群落的基因组总和进行研究, 而不需要分离、培养微生物。伴随着 PCR 技术、基因克隆技术以及基因组测序技术等逐渐成熟与完善, 以及相关学科的发展与渗透, 宏基因组学技术极大地推动了微生物学的发展。

1.1 宏基因组文库的构建

提取特定环境中微生物群落的总基因组 DNA, 保持环境样品中 DNA 的完整性和纯度是宏基因组文库构建的关键之一。目前采用的方法大体上分为两类^[7], 一类是直接提取法, 即直接裂解环境样品中的微生物细胞, 再抽提总基因组 DNA。裂解方法包括物理法(冻融法、超声法、玻璃珠击打法和液氮研磨法等)、化学法和酶法等; 另一类是间接提取法, 即采用差速离心等方法先将微生物细胞从样品中分离出来, 再用较温和的方法抽提总基因组 DNA。一般前者的提取效率较高, 但提取的 DNA 纯度低, 且由于强烈的机械剪切作用, 抽取的 DNA 片段长度有限(1-50 kb); 而后的纯度好, DNA 片段长度较大(20-500 kb), 但操作繁琐, 成本高, 有些微生物可能会在分离过程中丢失^[8]。

不同载体承载外源 DNA 片段大小的能力不同, 克隆片段的大小决定了构建生物基因组文库所需要的克隆子的数目。合适载体的选择取决于生物基因组的大小、欲插入目的片段的大小、宿主菌及筛选方法等因素。目前, 宏基因组文库构建中常用的载体主要包括质粒, 黏粒和 BAC 载体^[9]。宿主菌株的选择则取决于宿主细胞的转化率、重组载体在宿主中的稳定性、DNA 插入片段的表达, 以及目标性状的筛选方法等因素。目前大肠杆菌(*Escherichia coli*) 或其改造菌株是最为常用的宿主菌; 其次, 链霉菌(*Streptomycetaceae* spp.) 嗜热栖热菌(*Thermus thermophilus*) 等也已作为构建文库的宿主菌^[10]。

1.2 宏基因组文库的筛选

由于取样环境的复杂性和样品中微生物种类的多样性, 宏基因组文库容量一般较大, 因此, 从几万或几十万文库克隆子中筛选所需基因需要高效的筛选方法。目前, 宏基因组文库的高通量筛选方法主要有功能驱动的筛选和序列驱动的筛选^[11]。

功能驱动的筛选(Function-driven screening) 是以目标基因表达产物的活性测定为基础的筛选法^[12]。Healy 等^[13] 最早采用功能驱动的筛选方法从宏基因组文库中成功筛选到纤维素酶和木糖苷酶基因。此后, 从不同环境样品的宏基因组文库中相继筛选到了产生脂酶、蛋白酶、淀粉酶、多糖修饰酶、脱氢酶及抗菌活性的克隆子^[14]。该筛选方法可直接筛选到应用于工农业和医药业的酶等蛋白质和抗菌素等天然产物。但其缺陷是必须依赖于功能基因在外源宿主中的表达, 往往会因许多基因或基因产物在外源宿主中的不表达或表达后没有活性而降低了筛选效率^[15]。只有优化筛选方法, 选择合适的底物及筛选条件, 并建立合适的宿主载体系统, 才能加快对新型的生物催化剂的挖掘和利用。

序列驱动的筛选(sequence-driven screening) 是基于基因序列本身而进行的筛选方法, 以序列相似性为基础, 根据已知功能的基因序列设计探针或 PCR 引物, 通过核酸杂交或 PCR 扩增来筛选具有目标序列的克隆子^[16]。此筛选策略不依赖克隆基因在宿主内的表达, 但必须对相关基因序列有一定的了解。因此, 较难在宏基因组文库中筛选到与已知基因序列完全不同的全新功能基因。Venter^[17] 最早采用以序列为基础的筛选方法分析马尾藻海域的微生物, 鉴定出大量的新基因。利用基因的同源序列设计引物, 与宏基因组文库进行杂交已获得聚酮合成酶^[18]和脱水酶基因^[19], 有望开发应用于工业生产。此外, 鉴于第二代 DNA 测序技术的快速发展^[20], 直接对所构建的宏基因组文库的克隆子进行 DNA 序列的测定, 可以获得丰富的基因信息。利用基因芯片技术大大提高筛选效率, 有可能筛选到某一类结构或功能的新基因, 但同样需要了解相关基因序列, 较难发现全新的生物活性物质。

2 宏基因组学技术的应用——功能酶基因发现的新策略

2.1 极端环境中功能酶新基因的发现

微生物对环境的极强适应力和微生物本身的多样性及极高的变异率为发现大量新基因提供了可能^[21], 同时, 特殊的自然环境对微生物起到筛选和富集的作用, 因此以特殊环境样品中微生物复杂群

落作为宏基因组研究对象,为功能酶新基因的发现提供了更高的可能性。杜丽琴等^[22]利用宏基因组学技术从蔗糖富集的土壤样品中获得了两种蔗糖水解酶基因,与 GenBank 数据库中已知蔗糖水解酶的氨基酸序列相似性分别为 38% 和 68%。在极寒、酷热、高压、高渗透压等极端环境中生物多样性虽然低,但极端环境微生物是发现新型生物催化剂的重要来源^[21]。例如,在盐碱地、硫质温泉、冰河土壤、冰川冰南极冻土等极端环境样品的宏基因组中,已发现了具有生物学研究价值的大量的功能酶新基因^[23]。我国幅员辽阔,地质环境多样,通过宏基因组学技术已在温泉、沼气池等环境中发现醛脱氢酶基因^[24]、木聚糖酶基因^[25]等。西藏高原土壤中富含独特的微生物资源,芦晓飞等^[26]成功构建了一个包含 30 624 个克隆且稳定性好的宏基因组 Fosmid 文库,为从西藏高原土壤中挖掘和利用新的功能基因研究奠定了基础。

此外,环境微生物也是新药物发现的一个重要来源。据报道,一些从土壤中分离获得的微生物活性物质有望用于预防和治疗有机磷杀虫剂及神经毒剂对哺乳动物的作用^[27],从蚯蚓生存相关的微生物群落的宏基因组中发现的一种新型的血小板活化因子乙酰水解酶可能被应用于生物医药业^[28]。虽然微生物来源的酶和生物活性物质在医药业的应用有一定的难度,但宏基因组学技术对于从未培养微生物的基因资源中挖掘具有生物医药应用潜力的新基因有极大帮助^[29]。

2.2 共生菌中功能酶新基因的发现

共生菌是长期与宿主共同生活,呈相互依赖、互惠互利的共生关系的一大类群特殊微生物,蕴含着丰富的基因资源及生态资源,过去长期被忽略,相关研究也很薄弱。共生菌的生理生化特点因长期与宿主共同进化更具特色。不少从宿主中分离获得的生物活性物质的真正生产者是与共生或附生的未培养微生物,宏基因组学技术在共生菌基因资源的开发上具有独特的优势。

植物共生菌通常对植物生长有益,如固氮菌。通过对植物共生菌,地衣共生菌藻,尤其是农作物共生菌的宏基因组学研究,可以分析共生菌与植物

之间的作用机制。Sessitsch 等^[30]通过对水稻根部中植物共生细菌或内生真菌的宏基因组序列的分析,发现内生真菌细胞具有植物聚合物降解酶及蛋白质分泌系统等,内生真菌可能涉及整个水稻的氮循环。高质量的内生真菌群体对于植物生长的促进,植物抗逆性的提高,以及病原菌的生物防治和生物修复具有重要意义。

动物体表和体内也存在大量的微生物,利用宏基因组学技术可以推测微生物的功能酶对动物的影响。例如,从山羊皮肤表层中发现了碱性蛋白酶^[31],白蚁的后肠中发现了木聚糖酶和纤维素水解酶等微生物功能酶基因^[32]。研究较为深入的是反刍动物的瘤胃,瘤胃里存在着一个较复杂的瘤胃微生物生态系统,由于瘤胃微生物的存在,反刍动物可以连续、高效地利用低质粗饲料。但是通过纯培养技术仅分离到占其总量 11% 左右的瘤胃微生物^[33],而宏基因组学技术绕开了微生物纯培养,研究对象是瘤胃环境中全部微小生物的遗传物质总和,已经从瘤胃宏基因组中发现了木聚糖酶、葡萄糖苷酶、复合糖基水解酶,纤维素酶等瘤胃微生物的功能酶基因,受到养殖业、饲料业、食品、造纸和纺织业的关注^[34]。

2.3 人类宏基因组中功能酶新基因的发现

人体内部或体表有数以万亿的微生物,包括细菌、古细菌、真菌、寄生虫和病毒等,据估计有 90% 未在实验室条件下被纯培养^[35]。“人体微生物工程”(the human microbiome project, HMP),旨在描绘与人体内脏、口腔、皮肤、阴道相关的微生物群落的基因图谱,研究人体宏基因组结构和功能、相互之间关系、作用规律和与疾病关系^[36]。2004 年,美国国立卫生研究院专门设立“利用宏基因组学研究口腔微生物”的研究项目,2006 年,正式成立人类宏基因组计划国际联盟,启动人类宏基因组计划。在过去的五年里,在微生物与人类健康和疾病相关性研究方面已经开展了大量的研究^[37]。大多数与人体共生的微生物具有许多独特的生物学功能,这对于许多疾病发生机理的阐明、新药的研发等都将产生重大影响^[38]。

3 展望

宏基因组学突破了物种界限,将有助于揭示生

命活动的基本规律,拓展对生命本质的认识。例如,微生物极端环境的耐受机制、共生菌与宿主互相作用的特性、生物进化的规律等。随着科技的发展以及宏基因组学技术的完善,微生物功能酶基因将越来越广泛的应用于工业、农业、医学等领域,使人类生产和生活朝着更经济、更环保、更清洁的“绿色”方向发展。多功能酶是近年来酶学研究的新课题,但目前还没有多功能酶经过宏基因组学技术筛选获得。

但宏基因组学技术并非完美无缺,还存在着不少“瓶颈”问题有待解决。例如,现行的环境样品总基因组DNA提取方法对提取的DNA的纯度、片段长度及其代表性的要求还有待提高;另外,宏基因组文库的高通量筛选,尤其是功能驱动的筛选检出率不高,要做到高通量高效率的筛选仍然是十分棘手的问题。本实验室经过对环境宏基因组的蛋白酶、酯酶的相关大量研究,在宏基因组DNA的提取方面已经能够获得一定的成果,但是宏基因组文库的高通量筛选仍存在一定的难度,文库的构建以及载体和感受态细胞的选择直接影响功能驱动筛选对文库的筛选筛出率,而利用设计兼并引物进行的序列驱动筛选获得了大量不同种属蛋白酶的片段,但较难获得功能酶基因全长。根据不同的研究目的,选择不同的文库构建方法和筛选方法会比较成功获得功能酶新基因。

参考文献

- [1] 杜晨宇,李春,张木,等.工业生物催化过程发展及其展望.化工学报,2003,5(4):456-463.
- [2] Sleator RD, Shortall C, Hill C. Metagenomics. Lett Appl Microbiol, 2008, 47 :361-366.
- [3] Turnbaugh PJ, Gordon JI. An invitation to the marriage of metagenomics and metabolomics. Cell, 2008, 134 :708-713.
- [4] Ferrer M, Beloqui A. Microbial metagenomes :moving forward industrial biotechnology. Chemical Technology & Biotechnology, 2007, 5 (82) :421-423.
- [5] Jiang CJ, Hao ZY, Zeng R, et al. Characterization of a novel serine protease inhibitor gene from a marine metagenome. Mar Drugs, 2011, 9 (9) :1487-1501.
- [6] Handelsman J, Rondon MR, Brady SF, 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.
- [7] Robe P, Nalin R, Capellao C, et al. Extraction of DNA from soil. Journal of Soil Biology, 2003, 4 (39) :183-190.
- [8] Yang ZH, Xiao Y, Zeng GM, et al. Comparison of methods for total community DNA extraction and purification from compost. Appl Microbiol Biotechnol, 2007, 74 (4) :918-925.
- [9] Lodish H, Berk A, Zipursky SL, et al. Molecular cell biology [M] . 4th edition. Section 7.1 DNA Cloning with Plasmid Vectors, New York, 2000.
- [10] Simon C, Daniel R. Metagenomic analyses :past and future trends. Appl Environ Microbiol, 2011, 77 (4) :1153-1161.
- [11] Gabor E, Liebeton K, Niehaus F, et al. Updating the metagenomics toolbox. Biotechnol J, 2007, 2 (2) :201-206.
- [12] Dinsdale EA, Edwards RA, Hall D, et al. Functional metagenomic profiling of nine biomes. Nature, 2008, 452 (7187) :629-632.
- [13] Healy FG, Ray RM, Aldrich HC, et al. Direct isolation of functional genes encoding cellulases from the microbial consortia in a thermophilic, anaerobic digester maintained on lignocellulose. Appl Microbiol Biotechnol, 1995, 43 (4) :667-674.
- [14] Ferrer M, Martinez A, Barca F, et al. Mining genomes and ' metagenomes ' for novel catalysts. Biotechnology, 2005, 6 (16) :1-6.
- [15] Uchiyama T, Miyazaki M. Functional metagenomics for enzyme discovery :challenges of efficient screening. Curr Opin Biotechnol, 2009, 20 (6) :616-622.
- [16] Hoff KJ, Tech M, Lingner T, et al. Gene prediction in metagenomic fragments :a large scale machine learning approach. BMC Bioinformatics, 2008, 9 :217.
- [17] Venter JC, Remington K, Heidelberg JF, et al. Environmental genome shotgun sequencing of the Sargasso Sea. Science, 2004, 304 (5667) :66-74.
- [18] Courtois S, Cappellano CM, Ball M, et al. Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. Appl Environ Microbiol, 2003, 69 (1) :49-55.
- [19] Knietsch 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 Microbiology, 2003, 6 (69) :3048-3060.

- [20] Brockhurst MA, Colegrave N, Rozen DE. Next-generation sequencing as a tool to study microbial evolution. *Mol Ecol*, 2011, 20 (5) : 972-980.
- [21] Steele HL, Jaeger KE, Daniel R, et al. Advances in recovery of novel biocatalysts from metagenomes. *J Mol Microbiol Biotechnol*, 2009, 16 (1-2) : 25-37.
- [22] 杜丽琴, 庞浩, 胡媛媛, 等. 蔗糖富集环境土壤微生物宏基因组分析及蔗糖水解相关酶基因克隆. *应用与环境生物学报*, 2010, 16 (3) : 403-407.
- [23] Heath C, Hu XP, Cary SC, et al. Identification of a novel alkaliphilic esterase active at low temperatures by screening a metagenomic library from Antarctic desert soil. *Appl Environ Microbiol*, 2009, 75 : 4657-4659.
- [24] 谌容, 王秋岩, 杨兵, 等. 温泉环境宏基因组文库中醛脱氢酶基因的克隆及分析. *杭州师范大学学报: 自然科学版*, 2010, 9 (5) : 379-384.
- [25] 唐咸来, 车志群, 李双喜, 等. 沼气池宏基因组文库中未培养微生物 β -葡萄糖苷酶基因的克隆与鉴定. *广西植物*, 2008, 28 (6) : 790-794.
- [26] 芦晓飞, 赵志祥, 谢丙炎, 等. 西藏米拉山高寒草甸土壤微生物 DNA 提取及宏基因组 Fosmid 文库构建. *应用与环境生物学报*, 2009, 15 (6) : 824-829.
- [27] Sogorb MA, Vilanova E, Carrera V. Future application of phosphotriesterases in the prophylaxis and treatment of organophosphorus insecticide and nerve agent poisonings. *Toxicol Lett*, 2004, 151(1) : 219-233.
- [28] Navarro-Fernández J, Nechitaylo TY, Guerrero JA, et al. A novel platelet-activating factor acetylhydrolase discovered in a metagenome from the earthworm-associated microbial community. *Environ Microbiol*, 2011, 13 (11) : 3036-3046.
- [29] Singh BK, Macdonald CA. Drug discovery from uncultivable microorganisms. *Drug Discov Today*, 2010, 15 (17-18) : 792-799.
- [30] Sessitsch A, Hardoim P, Döring J, et al. Functional characteristics of an endophyte community colonizing rice roots as revealed by metagenomic analysis. *Mol Plant Microbe Interact*, 2012, 25 (1) : 28-36.
- [31] Pushpam PL, Rajesh T, Gunasekaran P. Identification and characterization of alkaline serine protease from goat skin surface metagenome. *AMB Express*, 2011, 1 (1) : 3.
- [32] Warnecke F, Luginbühl P, Ivanova N, et al. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature*, 2007, 450 (7169) : 560-565.
- [33] Edwards JE, McEwan NR, Travis AJ, et al. 16SrDNA library-based analysis of ruminal bacterial diversity. *Antonie van Leeuwenhoek*, 2004, 3 : 263-281.
- [34] 王佳堃, 安培培, 刘建新. 宏基因组学用于瘤胃微生物代谢的研究进展. *动物营养学报*, 2010, 22 (3) : 527-535.
- [35] Dethlefsen L, McFall-Ngai M, Relman DA. An ecological and evolutionary perspective on human-microbe mutualism and disease. *Nature*, 2007, 449 (7164) : 811-818.
- [36] NIH HMP Working Group, Peterson J, Garges S, Giovanni M, et al. The NIH Human Microbiome Project. *Genome Res*, 2009, 19(12) : 2317-2323.
- [37] Turnbaugh PJ, Ley RE, Hamady M, et al. The Human Microbiome Project. *Nature*, 2007, 449 : 804-810.
- [38] Virgin HW, Todd JA. Metagenomics and personalized medicine. *Cell*, 2011, 147(1) : 44-56.

(责任编辑 狄艳红)