

基于宏组学方法认识微生物群落及其功能

马海霞¹ 张丽丽¹ 孙晓萌¹ 张怀强¹ 何明雄² 陈冠军¹ 王禄山^{1,2*}

(1. 山东大学 微生物技术国家重点实验室 山东 济南 250100)

(2. 农业部农村可再生能源开发利用重点实验室 四川 成都 610041)

摘要: 进入后基因组学时代, 测序技术飞速发展, 测序成本明显下降, 形成了涵盖宏基因组学、宏转录组学和宏蛋白质组学的宏组学技术, 推动了对微生物群落的多样性、结构及潜在基因功能方面的深入研究。最近随着整合的宏组学技术的提出及应用, 全面系统分析微生物群落动态变化及其代谢功能已成为可能, 这将成为微生物生态学研究的新趋势。本文综述了宏组学在研究海洋湖泊、深海热泉、人体肠道、牛瘤胃生境、森林土壤与堆肥生境等环境中微生物群落的结构和功能方面的最新进展与成功应用案例。

关键词: 宏组学, 微生物群落, 宏基因组学, 宏转录组学, 宏蛋白质组学

Understanding microbial communities and their functions by meta-omics approaches

MA Hai-Xia¹ ZHANG Li-Li¹ SUN Xiao-Meng¹ ZHANG Huai-Qiang¹
HE Ming-Xiong² CHEN Guan-Jun¹ WANG Lu-Shan^{1,2*}

(1. State Key Laboratory of Microbial Technology, Shandong University, Jinan, Shandong 250100, China)

(2. Key Laboratory of Development and Application of Rural Renewable Energy, Ministry of Agriculture, Chengdu, Sichuan 610041, China)

Abstract: In the post-genomic era, the exponential progress in sequencing technology evidently reduced the cost of sequencing. Meta-omics technology includes metagenomics, metatranscriptomics and metaproteomics, has emerged in this progress, which has promoted the in-depth study of the diversity, construction and potential functions of microbial communities. Lately, the dynamic changes and metabolic functions of microbial communities can be comprehensively analyzed with the integrated application of meta-omics, and this is becoming a new research trend of microbial ecology. This paper summarized latest advances and successful application using meta-omics technology to study of structure and function of microbe communities in various environments, such as ocean, lakes, deep-sea hydrothermal vents, human gut, bovine rumen, soil and compost.

Keywords: Meta-omics, Microbial communities, Metagenomics, Metatranscriptomics, Metaproteomics

基金项目: 国家科技重大专项项目(No. 2013ZX10004217); 山东省自然科学基金项目(No. ZR2013CM038); 农业部农村可再生能源开发利用重点实验室开放课题基金项目

*通讯作者: Tel: 86-531-88366202; 信箱: lswang@sdu.edu.cn

收稿日期: 2014-12-01; 接受日期: 2015-03-03; 优先数字出版日期(www.cnki.net): 2015-03-04

微生物群落广泛分布在地球上的每个生态系统中,对全球碳、氮和硫等元素的物质循环起着重要作用^[1]。人们对于微生物群落的认识多基于传统微生物学方法,并结合生物化学、分子生物学和遗传学等方法,分析群落中可培养微生物分布与功能。但是,现在已培养的微生物可能不到自然界微生物总量的1%^[2],大量未培养的微生物如同地球上的“暗物质”^[3],在相应生境中存在种类到底有多少,如何发挥功能,传统微生物学方法已经不能全面解答。

宏基因组学(Metagenomics)的出现和兴起扩展了人们对微生物群落的认识。宏基因组学由 Handelsman 等^[4]1998年提出,指对生境中所有微生物的总基因组 DNA 进行研究,避开传统可培养方法的局限性,提高新微生物的发现效率。进入 21 世纪的第 2 个十年,测序技术飞速发展,测序费用大大降低,测序通量迅速扩大,可以对生境中所有微生物的 DNA 进行直接测序,这就使得人们可以对生境中动态微生物种群组成、分布及其动态变化进行全面研究^[5]。结合宏转录组学(Metatranscriptomics^[6])及新一代质谱技术催生下的宏蛋白质组学(Metaproteomics^[7]),人们就可以系统地研究微生物群落及其功能,全面分析其组成、发展规律、进化和功能,本文从宏组学的整体角度出发,综述了人们认识微生物群落及其功能的新方法及新进展。

1 宏基因组学——认识微生物群落中物种多样性及其基因潜力

区别于单个物种的基因组,宏基因组学研究是针对生境中微生物群落中所有 DNA 的分析。一般来说,宏基因组学研究主要分成两个层面,一是分析环境中特定基因,即通过构建宏基因组文库,基于序列筛选或者基于功能筛选分析某种功能基因,并进一步对筛选到的基因进行深度测序;二是针对环境中所有 DNA,进行深度测序,分析该生境中微生物的组成与相关功能。这两方面的相关分析都依

赖于测序的深度和广度。

测序技术的快速发展极大地提高了测序通量与精确性,大大缩短了测序时间,如二代测序 Illumina 平台 HiSeq2000 每天能运行产生 450 Gb 长约 100 个碱基对的数据。测序成本方面 21 世纪初至今已经降低几十万倍,新的方法完成个人基因组测序(30 亿 bp)成本约 1 000 美元。测序技术的快速发展使得宏基因组技术成为通用技术,大大促进微生物生态学的发展,宏基因组学技术一般的研究策略如图 1 所示^[5]。

从环境样品中提取总 DNA,对达到测序要求的总 DNA 样品上机测序,获得序列信息。根据不同的目的采取不同的途径分析序列数据,基因一般按照 cd-hit (例如 95%的相似性)聚类,将同源序列聚类成簇;通过对 16S rDNA 和 18S rDNA 序列标记分析比对可以获得微生物的种类和丰度信息,解析微生物群落的结构和进化关系;借助生物信息学分析软件和蛋白质数据库,例如 PFAM、TIGRFAM、COG (Cluster of orthologous groups of proteins, 蛋白相邻类的聚簇)和 KEGG (Kyoto encyclopedia of genes and genomes, 京都基因与基因组百科全书),可以预测分析蛋白质序列并注释蛋白质功能,分析潜在的基因功能。

宏基因组学技术已经应用于多个环境微生物群落的研究中,如人体肠道、动物瘤胃、土壤、海洋和湖泊中微生物区系变化。自 2003 年以来,研究的环境由简单到复杂,具体重要事件如图 2 所示^[9-18]。

1.1 肠道菌群及动态演化关系的分析

人体携带着超过自身细胞十倍数量的微生物以及比我们自身的基因数目多 100 倍的微生物基因,仅肠道中就含有 1 000 亿个微生物,包括细菌、真菌和少量的病毒^[19]。随着各国人类微生物组(Human microbiome)计划的开展与完成,人们对微生物与人类疾病的关系的认识和理解更进一步,尤其是肠道微生物在如肠炎、肥胖症、II 型糖尿病、结肠癌等疾病中的重要作用。

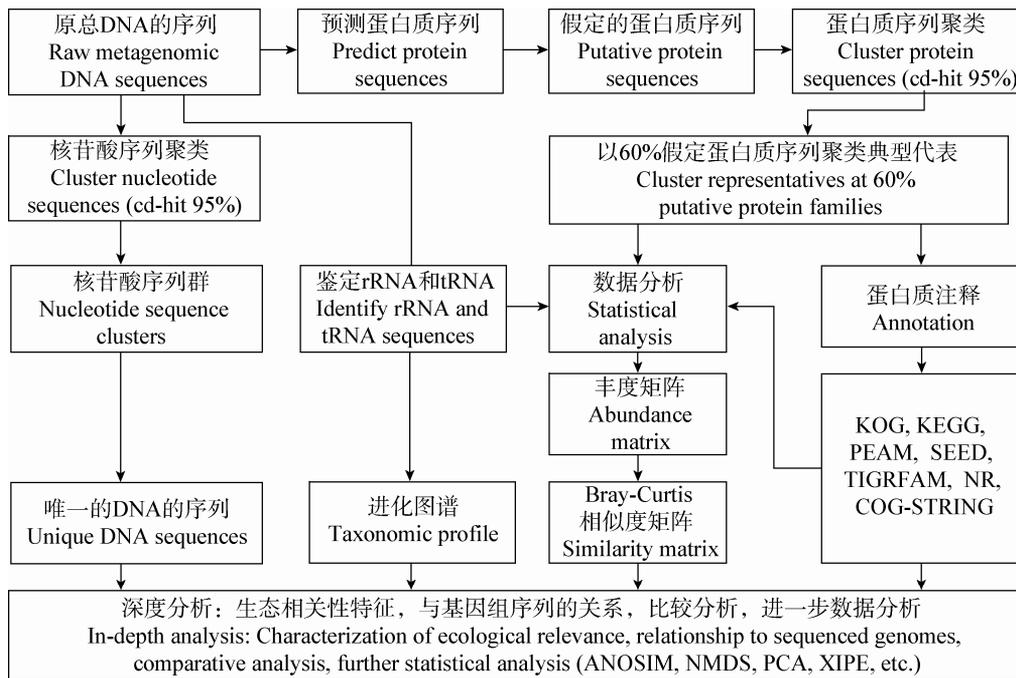


图 1 基于宏基因组序列分析微生物生态的一般研究策略^[5]

Figure 1 The common strategy of analysis of microbial ecosystem based on metagenomic sequence^[5]

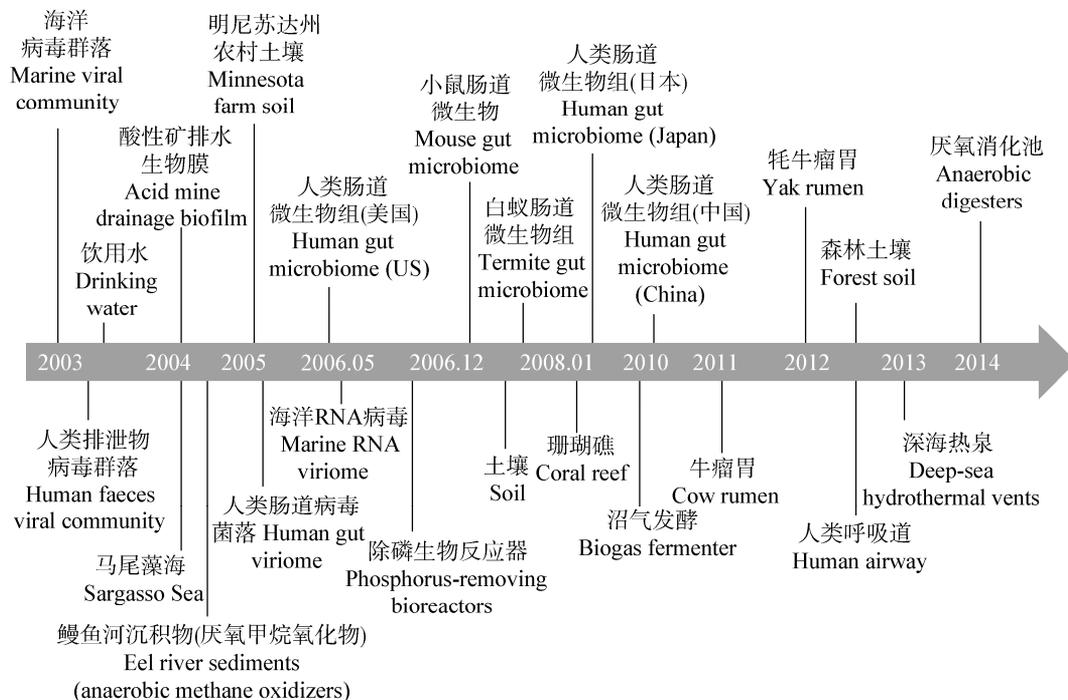


图 2 基于序列的宏基因组项目在多样化的环境样品研究中的应用的重要事件^[8]

Figure 2 The timeline of sequencing-based metagenomics project showing a variety environment sampled since 2003^[8]

肠道菌群的评估和鉴定已成为人类疾病重要的研究领域。宏基因组学已经在肠道微生物多样性和潜在功能基因研究中凸显出了巨大优势。人们已经在人类胃肠道中检测到了 9 个细菌门, 其中占据优势的是拟杆菌门、厚壁菌门和放线菌, 变形菌门较少^[19]。研究人员研究发现, 胎儿时期人体在无菌条件下成长, 出生后微生物开始定殖人类的胃肠系统, 且菌群随着年龄的增长发生变化。在幼儿时期, 母乳喂养的婴儿肠胃系统以双歧杆菌为主; 断奶后菌群数量明显增加; 成年时以拟杆菌门、厚壁菌门和放线菌门微生物为主; 老年时期, 厚壁菌门与双歧杆菌明显减少, 拟杆菌门和变形菌门增加明显^[15]。以上研究, 如结合生理检测到的菌群变化, 可以为人类肠道和身体健康等保健与治疗提供科学依据。

近年来, 我国研究人员在肠道微生物研究中也取得了一定的成果。如 2010 年深圳华大基因研究院团队对 124 名欧洲人的粪便样品 DNA 信息进行分析, 利用 Illumina 测序获得 576.7 Gb 的数据, 获得 330 万个参考基因, 这比人体基因数目多 150 倍。结果分析表明, 其中 99% 的基因来自细菌, 总体包括 1 000–1 150 种微生物, 其中至少 160 种微生物是人体共有的优势物种^[16]。此外, 在肠道微生物与 II 型糖尿病研究方面, 人们基于深度鸟枪测序法对来自 345 名 II 型糖尿病患者和非糖尿病患者对照的粪便样肠道微生物 DNA 进行测序, 结果分析表明 II 型糖尿病患者中肠道微生物生态已中度失调, 丁酸产生菌丰度减少, 条件致病菌增加, 硫酸盐还原与氧化应激性功能的微生物增加^[20]。这些研究可能对实现糖尿病患者的个性化医疗提供了指导。

1.2 生物质降解微生物群落

生物质是重要的生物资源, 其高效降解一直是人们关注的热点。瘤胃生境与堆肥生境是生物质高效降解的两个生境, 其中都含有大量降解生物质的微生物区系^[18,21]。Hess 等^[17]利用测序技术分析牛瘤胃中附着在植物纤维上的微生物宏基因 DNA, 从获得的 268 Gb 数据中鉴定出 27 755 个碳水化合物降解酶相关基因, 并成功表达了 90 个候选基因,

还组装了 15 个未培养微生物的全基因组。相关研究扩展人们对降解纤维素生物质的基因和基因组的认识, 对深入挖掘降解酶基因资源提供了依据, 并且利用相关技术还可组装出多个未培养微生物全基因组, 这也表明了宏基因组测序在认识群落中未培养微生物的重要作用。中国科学院微生物研究所的 Dai 等^[21]基于宏基因组与功能筛选等方法, 研究了牦牛瘤胃中水解纤维素的微生物以及纤维素降解基因, 分析发现有 150 种降解纤维糖苷水解酶 (Glycoside hydrolases, GH) 基因, 69% 成簇并与编码相关功能的基因相连; 对 35 个长度大于 10 kb 水解纤维素的叠连群的分析, 发现 25 个来自于拟杆菌门和 4 个来自于厚壁菌门; 发现 GH5 亚家族的新型内切葡聚糖酶基因在瘤胃中的重要作用。

堆肥生境研究方面, 宏基因组技术在分析微生物群落结构变化上也具有明显优势。Martins 等^[18]通过对圣保罗热带公园堆肥样品总 DNA 测序, 共得到 300 万条序列, 分析结果表明该堆肥生境中变形菌门的乳酸杆菌属占据优势。根据 COG 功能基因预测, 共有 112 个纤维素酶相关(多属于糖苷水解酶 GH5 家族)蛋白质, 32 个具有纤维素结合结构域 (Carbohydrate binding module, CBM)。这是第一次对堆肥生境的总 DNA 进行高通量测序研究, 开启了堆肥生境中高效降解菌群及其基因资源的认识。Gannes 等^[22]利用 454 焦磷酸测序分析了 3 种木质纤维素堆肥系统的主要阶段, 结果显示随着堆肥过程的进行微生物群落发生明显更替, 堆肥嗜温期和腐熟期微生物群落各不相同, 嗜温期微生物在腐熟期并不是先前推测的那样处于休眠状态, 堆肥微生物群落差异中所处发酵阶段的影响比堆肥底物的类型更大。相关研究将有助于堆肥腐熟生物指标的标准化。

1.3 海洋极端环境微生物群落

红海由于其特殊的地理位置, 形成特殊的极端环境, 是地球上温度最高盐度最大的海水环境之一。四十年以来发现了许多盐水池, 并以缺氧、高盐度、热液和金属含量高的环境成为地球上的一些

极端的环境,具有重要的研究价值。Bougouffa 等^[23]在 2013 年对红海中两种具有高盐和超热特定生境中不同深度海水样品微生物区系进行分析,发现海洋对流层中的微生物丰度低但多样性高。细菌群落随着海洋深度的变化分层明显,多变量分析结果显示温度和盐度是影响群落形成的主要因素。此外海水与沉积物中真核生物群落分布特征与阿拉伯海岸的相似,并检测到许多未分类的微生物种类^[24]。通过比较红海中染病与健康海绵生物 *Carteriospongia foliascens* 微生物群落差异,发现健康海绵上微生物群落多样性低,主要以蓝藻细菌、拟杆菌和变形菌为主要菌群;而染病海绵上原有主要蓝藻细菌属明显减少,出现不明分支的蓝藻细菌,并存在向周围海水扩散迹象^[25]。这些研究结果对于人们认识极端环境微生物群落的生存状况和将来研究微生物在胁迫条件下的应对机制具有重要意义。

宏基因组学方法与技术在土壤、海洋和人体等生境研究中广泛应用,将全面揭示微生物种群的结构与其基因潜在功能,这将微生物群落的结构与功能及其与环境的关系等研究带来一场变革。不过宏基因组学技术依赖于总 DNA 提取与测序深度。天然生境复杂多样,环境中的核酸酶及腐殖质都会影响总 DNA 提取,因此需要根据不同生境样品优化提取条件与提取方法。现有的测序技术也各有优缺点,如 Illumin 适应高通量测序,但是序列片段短;太平洋测序生物科学公司 PacBioRS 平台可以完成实时单分子 DNA 测序,读长可达 1 000 个碱基,但是一次性准确性达标的机率低(约 81%–83%),DNA 聚合酶在反应中易降解等。另外测序数据的后续分析也依赖于硬件系统与算法程序的精确性。随着提取条件的优化提升、测序技术的进一步发展,宏基因组学技术将会发挥更大的技术优势。

2 宏转录组学——认识微生物群落的基因表达与调控

宏基因组学虽然能够提供微生物区系(尤其是未培养的微生物)潜在的活动信息,但仍不能揭示在特

定的时空环境下,微生物群落基因的动态表达与调控等问题。要解决这一问题,就需要在转录与表达水平上进一步研究。宏转录组是指在某个特定时刻生境中所有微生物基因转录体的集合,这是原位衡量宏基因组表达的一种方法。传统技术可利用微阵列等来分析微生物群落的基因表达,可是设计和构建微阵列不仅耗时,而且费用昂贵,还不能检测新基因。随着高通量测序技术的发展,生境中 RNA 也可以直接测序分析,不过其组装过程却是重要的挑战之一。首先,与可利用重叠区域组装基因组 DNA 不同,RNA 深度测序得到的序列大小范围更广,例如小 RNA 等就不需要组装,大片的 RNA 则需要组装;其次,不同于 DNA 的双链都可以测序,RNA 只有一条有义链是有效的;第三,由于外显子的存在,来自同一个基因的 RNA 可能有不同的拼接方式,因此现有 RNA 组装过程需要依赖已有的基因组数据库。随着 RNA-seq 测序技术的发展及其相关新算法的发明,可以对序列从头组装,不再依赖于参考基因组,但是仍存在错误率高的缺点^[26]。尽管如此,在现有技术条件下,宏转录组学在多种生境中的应用仍然给人们带来了许多新发现。

宏转录组分析的一般策略是:采集环境样品,提取总 RNA,去除残余的 DNA,扩增 mRNA,由于 mRNA 极易降解,需要将 RNA 反转录为 cDNA,然后对 cDNA 进行测序,测序平台同样是二代测序平台如 454 测序和 Illumina 测序平台或者新一代单分子实时测序平台等。真核微生物基因转录产生的 mRNA 携带有 Poly-A 尾巴,依此可将原核生物与真核生物的表达区别开来。近年来宏转录组学在海洋和土壤等生境研究中得到广泛应用。

2.1 认识微生物间共生与协作

自然界中存在大量未培养微生物,其中部分原因是微生物间存在共生关系,很难分离到共生关系中的一方,而事实上相关微生物在生境中起着重要作用,利用 RNA 转录组分析可以获得未培养微生物间的生理功能与相互协作。Bomar 等^[27]利用 RNA-seq 技术解析了 *Hirudo verbana* (一种医用水

蛭)肠道中宏转录组的特征, 结果表明类 *Rikenella* 的细菌是丰度最高且未培养的共生体, 以硫化、唾液酸化和粘化的聚糖作为能量供应的来源。其次是 *Aeromonas veronii*, 根据转录组学数据预测它可利用粘蛋白或者碳水化合物为能源。基于相关信息, 研究者设计了含有粘蛋白聚糖培养基, 培养出类 *Rikenella* 的细菌。该研究说明宏转录组学在解析微生物群落功能方面应用, 可以利于培养基的设计, 完成更多未培养微生物的分离与培养。

2.2 认识微生物群落中生物质主要降解者与代谢过程

真核生物对森林土壤的肥力保持和生物质降解起着重要作用, Damon 等^[28]对山毛榉和红杉森林土壤中的宏转录组中真核 cDNA 进行测序, 发现 52%云杉森林的 cDNA 与 60%山毛榉森林的 cDNA 在 GenBank/EMBL/DDJB 数据库中没有找到同源序列, 对 12 个全长序列的碳水化合物活性酶组分系统发育分析发现, 大部分与真菌没有关系。如部分 GH45 内切纤维素酶与软体动物亲缘关系更近, 而 GH7 纤维二糖水解酶与甲壳纲动物亲缘关系最近, 这说明在土壤有机质的降解中非真菌的真核生物可能也起着重要作用。Radax 等^[29]利用宏转录组学对海绵中微生物群落的“双 RNA”(rRNA 和 mRNA) 进行分析, 共获得 262 298 个 RNA-tags, 109 325 个 rRNARibo-tags 揭示转录基因来自 10 个门, 其中海绵杆菌门和酸酞菌门是主要的菌群, 丰度最高的 mRNA 是来自氨氧化古菌有关硝化作用的主要代谢酶。

综上, 宏转录组学技术可以帮助人们认识微生物在特定环境下的代谢过程, 不仅限于认识微生物相关基因, 还可以认识微生物群落的代谢活动。该方法仍存在许多技术上的挑战, 如 mRNA 半衰期短、富集非常困难、cDNA 合成和扩增存在偏差等。除了样品处理与存储等问题外, 宏转录组学方法和测序平台优化提升可以解决部分问题。随着样品处理与测序技术的进步, 宏转录组学技术将给微生物群落研究带来新的变革。

3 宏蛋白质组学——微生物群落功能的直接表征

尽管宏基因组学研究能够给出微生物群落中潜在基因的功能, 宏转录组学能够深入了解基因表达和活性的程度, 但其细胞定位与活动调节是发生在蛋白质水平上的, 因此从转录组与蛋白质组得到的信息将大不相同。研究微生物群落的功能与代谢途径必须对蛋白质组进行研究, 通过揭示环境中蛋白质的组成与丰度、蛋白质的不同修饰、蛋白质和蛋白质之间的相互关系, 可以认识微生物群落的发展、种内相互关系、营养竞争关系等^[21]。

宏蛋白质组技术分析因蛋白质复杂多样的特点, 因此对检测技术要求具有高通量、高灵敏度、动态范围广、质量估计精确等特点, 而质谱分析是最合适条件的检测平台。早期蛋白质组研究通用方法是 2D (二维)凝胶电泳和 MS (Mass spectrometry, 质谱)结合, 一次实验中凝胶要经过多重的染色分析与鉴定。近年来, 多重色谱分离与 MS 联用技术可以对上万条多肽片段信息进行定性与定量分析^[30], 特别是轨道离子阱(Orbitrap)质谱的应用, 可利用痕量样品就可以鉴定蛋白质, 可以用于环境样品中低丰度的蛋白质分析检测。蛋白质由 20 个氨基酸组成, 如肽段中含有 6 个氨基酸, 蛋白质序列空间就会有 $20^6=6400$ 万种, 因此利用质谱测出肽序列至 6-10 个氨基酸的片段, 就足以鉴定一种蛋白质^[31]。不仅如此, 还可以通过鉴定到的肽段, 设计 DNA 引物序列, 直接从生境中扩增相关基因。

宏蛋白质组学技术的研究策略与一般的流程总结如图 3 所示^[32]。在采集环境样品后, 提取环境中的总蛋白质, 富集后一般采用双向电泳分离, 利用蛋白酶降解(一般采用胰酶酶解)形成混合肽段样品, 经液相色谱分离, 利用 MALDI 或者 EI 技术将相关肽段离子化, 利用一级或二级质谱(MS/MS)分析不同片段的质荷比(m/z), 提交 Mascot 和 SEQUEST 软件或从头测序软件包完成相关片段的鉴定。由于蛋白质是基因的表达产物, 因此分析宏

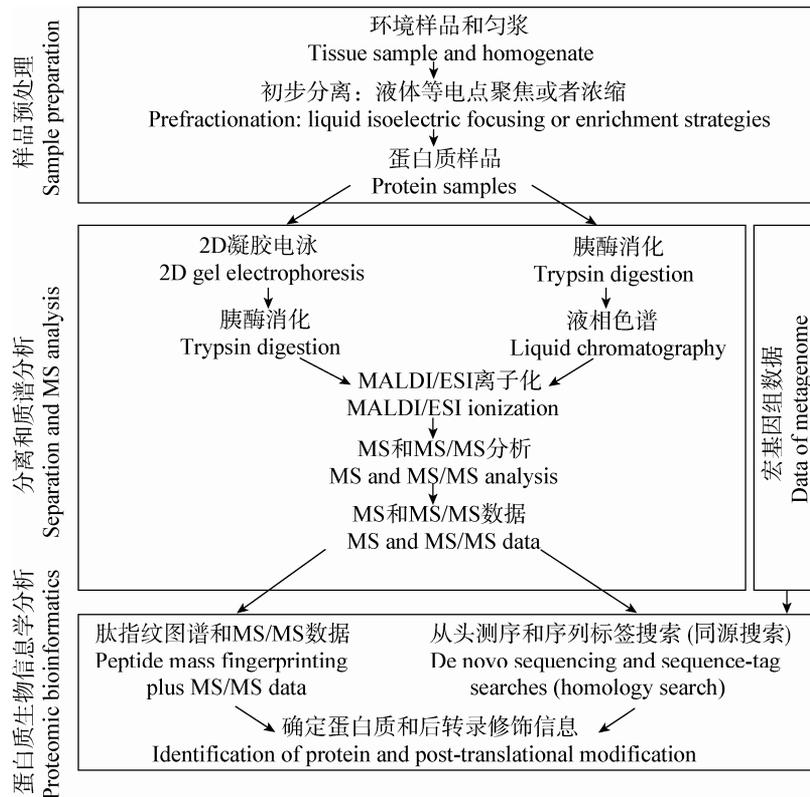


图3 宏蛋白质组研究策略一般流程

Figure 3 Overview of different proteomic work flows

注: 2D: 二维凝胶电泳; MALDI: 基质辅助激光解吸电离; ESI: 电喷雾电离; MS: 质谱; MS/MS: 串联质谱^[32]。

Note: 2D: Two-dimensional gel electrophoresis; MALDI: Matrix-assisted laser desorption ionization; ESI: Electrospray ionization; MS: Mass spectrometry; MS/MS: Tandem mass spectrometry^[32].

蛋白质组数据时可以参照相应宏基因组数据或者其系统发育数据集^[32]。研究相关宏蛋白质组学组分的功能, 还可以利用 COG 或 KEGG 等生物信息学数据库工具。

近年来, 宏蛋白质组学的研究主要集中在 3 种生态系统: 水域(海洋和湖泊)、陆地(土壤和沉积物)、真核宿主微生物(白蚁、小鼠、植物和人体)。对陆地生态系统的研究, 主要是土壤中微生物在碳氮循环中的作用、污染修复与冻土代谢相关潜力分析。近年来, 土壤原位分离蛋白质技术的提升加快了宏蛋白质组研究的广度与深度。

3.1 人体微生物组方面的研究进展

目前为止, 人们已开展了 3 个部位的宏蛋白质组分析: 口腔(唾液)、膀胱(尿液)和肠道(排泄物, 不同肠道灌洗)。由于肠道微生物的宏基因组研究较

多, 因此宏蛋白质组学的研究主要集中在肠道微生物方面^[33]。第一例关于人体肠道微生物的宏蛋白质组研究对象来自于出生 100 d 的两个婴儿。当时由于缺乏肠道微生物基因组测序数据库, 因此只是对蛋白质进行了从头测序分析, 研究发现双歧杆菌转醛酶与婴儿排泄物常出现的一个蛋白质相似, 并在宏蛋白质组中发挥重要功能^[34]。Kolmeder 等^[35]采集了 3 个混食性女性的排泄物, 基于质谱分析结果表明, 来自同一个体蛋白质组相似性比来自两个体的蛋白质组相似性要高, 支持人类宏蛋白质组具有个性。通过瘦型和胖型个体的肠道微生物宏蛋白质组比较分析, 结果显示拟杆菌分泌的蛋白质在胖型个体中更为丰富, 相关假说仍需要更多群体研究结果的支持。此外, 鉴定的 613 种蛋白质中有 505 种在两种表型个体中共同存在^[36]。宏蛋白质组间的比

较对认识个体差异提供了丰富的信息。

Jagtap 等^[37]在 2012 年首次深入研究健康人体唾液中微生物宏蛋白质组,发现含有来自未培养微生物 TM7 的蛋白质,共检测 65 个菌属,其中来自链球菌(*Streptococcus*)、罗氏菌属(*Rothia*)、放线菌(*Actinomyces*)、普氏菌(*Prevotella*)、奈瑟氏菌(*Neisseria*)、乳杆菌(*Lactobacillus*)和月形单胞菌(*Selenomonas*)等的蛋白质含量丰富。此外,通过 KEGG 数据库比对,得到 20 条 KEGG 代谢通路,代表性的有碳水化合物途径、氨基酸代谢途径、能量代谢途径、膜转运蛋白、信号因子等,并且糖酵解途径代谢活跃。相关成果为将来研究人体唾液中微生物蛋白质组的变化与口腔疾病的关系提供了良好的基础。

3.2 海洋能量代谢与物质代谢相关研究

在海洋水域宏蛋白质组研究中, Morris 等^[38]比较分析了在南大西洋海水中不同营养成分浓度下微生物群落膜上的宏蛋白质组,分析发现主要的膜蛋白质是依赖 TonB 型转运蛋白(TBDTs),该蛋白质是利用质子动力转运营养素穿过革兰氏阴性细菌的外膜。另外数据分析还发现革兰氏阴性菌细胞膜的宏蛋白质组中检测到了吸收光的视紫红质,这为视紫质蛋白(可表征世界海洋透光层中能量流)主要源头又提供了新的证据。蛋白质系统发生分析表明 TBDTs 与视紫红质为同一世系中,光营养的细菌浮游生物有可能利用光能提供转运动力。此外,相关研究还发现在南大西洋上层流中具有病毒蛋白与古细菌氨单加氧酶,推测古细菌在富含营养的水域中进行硝化作用。

3.3 土壤微生物群落活性酶组分分析

在土壤微生物群落研究中, Schneider 等^[39]采集澳大利亚森林里不同营养成分的凋落物,提取其蛋白质组经 1D-SDS-PAGE 分析后,利用液相色谱和串联质谱检测与生物信息学比对分析,发现凋落物的营养浓度及 C:N:P 的比例影响着微生物群落的结构和相关活性。其中,真菌是胞外酶的主要分泌者,

深入分析表明微生物更替受到营养物浓度的明显影响;较高营养浓度下微生物活性也更高,表现在胞外酶组分的种类多样与活性提高。这为人们认识生境中微生物群落的更替提供了新技术与新方法。

宏蛋白质组学对研究生境中微生物区系的活动非常具有潜力,但是环境样品(如土壤、沉积物等)成分复杂,蛋白质提取是该技术最大挑战。Keiblinger 等^[40]利用 4 种不同处理方式提取土壤中的蛋白质,分别是 SDS-不加入苯酚、NaOH-苯酚法、SDS-苯酚法提取、在 SDS-苯酚法提取前清洗样品,然后用质谱检测提取的质量,发现用 SDS-苯酚法提取获得的蛋白质数量最多。此外,蛋白质不能扩增,稳定性差,环境中蛋白质相对含量少等因素,也影响提取效果与提取效率。随着技术的改善和提高,宏蛋白质组将为人们提供更准确的信息。

4 整合宏组学技术及其相关研究进展

进入后基因组时代,宏基因组、宏转录组和宏蛋白质组学等宏组学带给微生物学研究新的变革,拓展了人们对于微生物群落的认识,基于这 3 种方法的整合宏组学从系统生物观出发研究微生物群落,已经成为新的研究趋势。

用整合宏基因组学、宏转录组学和宏蛋白质组学的宏组学技术研究微生物群落结构与功能的流程图如图 4 所示。宏基因组学可以确定微生物群落的组成,也可以提供潜在功能信息。而宏转录组学可以用来评估基因表达和推测关键的代谢通路。宏蛋白质组学确定表达中的细胞定位和调控、后转录修饰等。

4.1 宏组学在研究人体微生物中的应用典例

在人体研究方面, Lim 等^[41]以囊胞性纤维症病人呼吸道中的微生物为研究对象,排除人体 DNA 的影响,采用宏基因组学方法和宏转录组学两种方法对微生物的 DNA 和转录 mRNA 测序分析研究,结果表明每一个病人拥有一个独特的微生物组,在一些病人呼吸道内绿脓杆菌(*Pseudomonas aeruginos*)被其他的条件性致病菌代替。尽管每个人

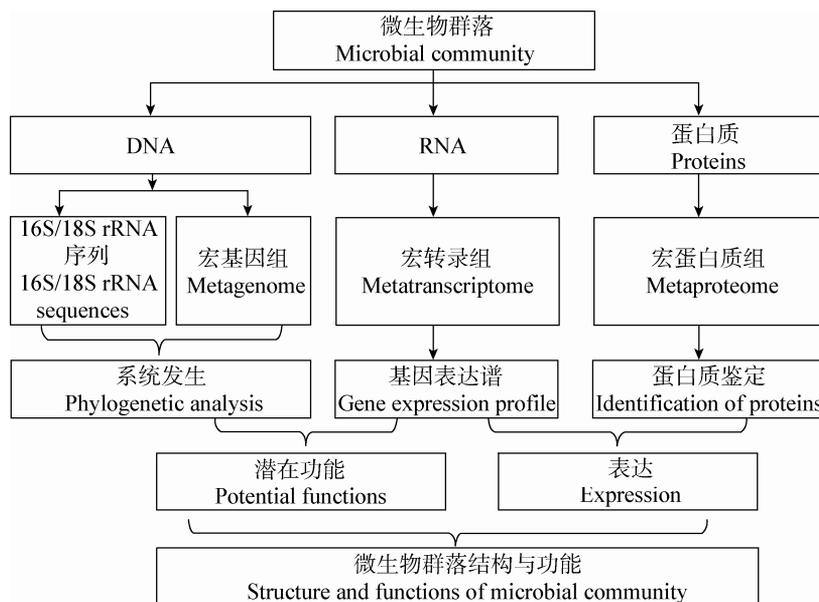


图 4 宏组学技术研究微生物群落的结构与功能

Figure 4 Meta-omics to study the structure and function of microbial community

的微生物区系组成不同,但是每个群落编码产生的代谢潜力是相同的。病毒群以噬菌体为主,是感染囊胞性纤维症病人的主要病原。宏转录组揭示了健康状况不同,微生物群落代谢编码也是不同的。这是第一个同时进行微生物组、病毒与群落宏转录组相关分析的研究工作。

Fouts 等^[42]在 2012 年第一次对无症状性尿液和尿路感染人体尿液对比研究其微生物蛋白质组,也分析了两组中的微生物多样性。研究表明,健康的人体尿液并不是先前认为无菌的,女性体内尿液主要是乳杆菌,男性体内尿液主要是棒状杆菌属。宏蛋白质组分析结果表明,感染病原菌的尿液中出现人体防御机制需要的蛋白抗体等。

4.2 整合宏组学在研究深海热泉微生物中的应用

Urlich 等^[43]在研究深海热泉这一地球上独特的环境(它周围形成利用地球化学能源的化能无自养微生物错综复杂的生态系统)时,在挪威格陵兰海扬马延岛采集了两个深海热泉口周围的沉积物样品,采用整合宏组学方法(宏转录组方法、宏基因组学和宏蛋白质组学)进行研究。两个样品分别获得

109 728 个宏基因组序列和 41 260 个环境转录组学序列,分析表明两个沉积物具有相似的种群结构,包括 ϵ -变形菌门、 γ -变形菌门和 δ -变形菌门,还有纤毛虫、线虫和多种多样的古菌群。通过转录和多肽分析鉴定了高活性的代谢通路,与硫氧化、甲烷氧化和碳固定代谢通路相关的基因大量表达,此外,有氧和无氧(硝酸盐和硫酸盐)呼吸链相关基因表达量也很高。趋化性和鞭毛基因的高效表达反映了微生物群在富含生理生化物质的动态栖息地中的生活方式。主要代谢途径分布在不同分类群中,因此可以推断相关原核生物和真核生物区系的部分功能。该研究充分展示出利用宏组学技术认识微生物群落代谢途径的优势。

4.3 整合宏组学分析厌氧消化池中微生物群落的结构、相互作用

复杂的生态系统(如大规模沼气发酵系统)现在可以利用整合宏组学进行全面深入的研究。已有研究表明:沼气发酵系统中厚壁菌门与拟杆菌门物种可能负责多糖的降解和发酵,其中梭菌属与产甲烷菌存在互养关系^[44]。通过对厌氧消化池微生物群落的宏转录组学分析,发现厌氧消化池中转录基

因多为编码底物水解、酸化和乙酸形成的相关酶类, 另外古细菌的转录活性很高^[45]。通过宏基因组与宏蛋白质组的分析, 发现与降解植物碳水化合物有关的少量厚壁菌门物种表达的糖苷水解酶与大量拟杆菌门表达的糖转运蛋白。宏基因组数据表明产甲烷菌只是群落中相对较少的种群, 而宏蛋白质学数据表明甲烷生成所需要的关键酶是高度表达的^[46]。这也从侧面反映了微生物群落中占据数量优势, 但由于受环境等因素影响, 其基因表达分泌量并不一定占据优势, 存在种类数量少但是功能酶分泌量大的情况。

近年来, 微生物技术国家重点实验室相关研究室也利用宏组学技术开展了天然堆肥生境中微生物群落的系统研究。利用 454 焦磷酸测序平台分析了农作物秸秆堆肥生境中不同空间深度微生物群落中真菌和细菌的动态变化, 发现在堆肥不同空间深度上, 主导微生物菌群明显, 在氧气充足外层, 真菌以嗜热丝孢菌属为优势菌属, 细菌以喜热裂孢菌属、嗜热多孢菌属和高温芽孢杆菌为主要菌群^[47], 利用宏蛋白质组分析也发现了相关优势菌群分离的主要糖苷水解酶类。这对于发现新的木质纤维素降解酶基因资源及全面认识生物质降解代谢通路具有重大意义, 也会对生物质资源的开发与工业应用提供理论和实践指导。

综上所述, 宏基因组学、宏转录组学、宏蛋白质组学技术在 DNA、RNA 和蛋白质三个层次上揭示微生物群落的结构、系统发生、代谢功能、调控规律等。随着宏组学技术提取方法的改进与高通量测序技术的进一步发展, 相信这些技术将会在多种生态领域得到快速的发展和运用, 可以使人们从系统角度全面认识微生物群落与其功能, 利用其规律、挖掘新的微生物菌种及其酶资源等, 这将是未来微生物生态学研究的新趋势与新研究方向。

参 考 文 献

[1] Madsen EL. Microorganisms and their roles in fundamental biogeochemical cycles[J]. *Current Opinion in Biotechnology*, 2011, 22: 456-464

- [2] Amann RI, Ludwig W, Schleifer KH. Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation[J]. *Microbiological Reviews*, 1995, 59(1): 143-169
- [3] Jansson JK, Prosser JI. The life beneath our feet[J]. *Nature*, 2013, 494(7435): 40-41
- [4] Handelman J, Rondon MR, Brady SF, et al. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products[J]. *Chemistry & Biology*, 1998, 5(10): 245-249
- [5] Gilbert JK, Dupont CL. Microbial metagenomics: beyond genomes[J]. *The Annual Review of Marine Science*, 2011, 3: 347-371
- [6] Leininger S, Urich T, Sehlter M, et al. Archaea predominate among ammonia-oxidizing prokaryotes in soils[J]. *Nature*, 2006, 442: 806-809
- [7] Rodriguez VF. Environmental genomics, the big picture?[J]. *FEMS Microbiology Letters*, 2004, 231(2): 153-158
- [8] Hugenholtz P, Tyson GW. Microbiology: metagenomics[J]. *Nature*, 2008, 455: 481-483
- [9] Tyson GW, Chapman J, Hugenholtz P, et al. Community structure and metabolism through reconstruction of microbial genomes from the environment[J]. *Nature*, 2004, 428: 37-43
- [10] Edwards RA, Rohwer F. Viral metagenomics[J]. *Nature Reviews Microbiology*, 2005, 3: 504-510
- [11] DeLong EF, Preston CM, Mincer T, et al. Community genomics among stratified microbial assemblages in the ocean's interior[J]. *Science*, 2006, 311: 496-503
- [12] Turnbaugh PJ, Ley RE, Mahowald MA, et al. An obesity-associated gut microbiome with increased capacity for energy harvest[J]. *Nature*, 2006, 444: 1027-1031
- [13] Yooseph S, Sutton G, Rusch DB, et al. The sorcerer II global ocean sampling expedition: expanding the universe of protein families[J]. *PLoS Biology*, 2007, 5: e16
- [14] Warnecke F, Luginbühl P, Ivanova N, et al. Metagenomic and functional analysis of hind gut microbiota of a wood-feeding higher termite[J]. *Nature*, 2007, 450: 560-565
- [15] Yin YS. Intestinal Microbiome: a "Functional Organ" Forgotten[M]. *Lifeomics*, 2013, 9(58): 1-46 (in Chinese) 尹业师. 肠道菌群——一个被遗忘的“功能器官”[M]. *生命奥秘*, 2013, 9(58): 1-46
- [16] Qin JJ, Li RQ, Raes J. A human gut microbial gene catalogue established by metagenomics sequencing[J]. *Nature*, 2010, 464: 59-65
- [17] Hess M, Sczyrba A, Egan R, et al. Metagenomic discovery of biomass-degrading genes and genomes from cow rumen[J]. *Science*, 2011, 331: 463-467
- [18] Martins LF, Antunes LP, Pascon RC, et al. Metagenomic analysis of a tropical composting operation at the Sao Paulo zoo park reveals diversity of biomass degradation functions and organisms[J]. *PLoS One*, 2013, 8(4): e61928
- [19] Human microbiome project consortium. Structure, function and diversity of the healthy human microbiome[J]. *Nature*, 2012, 486(7402): 207-214
- [20] Qin JJ, Li YR, Cai ZM, et al. A metagenome-wide association study of gut microbiota in type 2 diabetes[J]. *Nature*, 2012, 490(7418): 55-60
- [21] Dai X, Zhu Y, Luo Y, et al. Metagenomic insights into the fibrolytic microbiome in Yak rumen[J]. *PLoS One*, 2012, 7(7): e40430
- [22] Gannes V, Eudoxie G, Hickey WJ, et al. Prokaryotic successions and diversity in composts as revealed, by 454-pyrosequencing[J]. *Bioresource Technology*, 2013, 133: 573-580

- [23] Bougouffa S, Yang JK, Lee OO, et al. Distinctive microbial community structure in highly stratified deep sea brine water columns[J]. Applied and Environmental Microbiology, 2013, 79(11): 425-3437
- [24] Wang Y, Zhang WP, Qian PY, et al. Diversity and distribution of eukaryotic microbes in and around a brine pool adjacent to the Thuwal cold seeps in the Red Sea[J]. Frontiers in Microbiology, 2014, 5: 37
- [25] Gao ZM, Wang Y, Qian PY, et al. Pyrosequencing reveals the microbial communities in the Red Sea sponge *Carteriospongia foliascens* and their impressive shifts in abnormal tissues[J]. Microbial Ecology, 2014, 68: 621-632
- [26] Martin JA, Wang Z. Next-generation transcriptome assembly[J]. Nature Reviews, 2011, 12: 671-682
- [27] Bomar L, Maltz M, Colston S, et al. Directed culturing of microorganisms using metatranscriptomics[J]. mBio, 2011, 2(2): e00012-11
- [28] Damon C, Lehenbre F, Oger-Desfeux C, et al. Metatranscriptomics reveals the diversity of genes expressed by eukaryotes in forest soils[J]. PLoS One, 2012, 7(1): e28967
- [29] Radax R, Rattei T, Lanzan A, et al. Metatranscriptomics of the marine sponge *Geodiabarretti*: tackling phylogeny and function of its microbial community[J]. Environmental Microbiology, 2012, 14(5): 1308-1324
- [30] Hettich RL, Sharma R, Chourey K. Microbial metaproteomics: identifying the repertoire of proteins that microorganisms use to compete and cooperate in complex environmental communities[J]. Current Opinion in Microbiology, 2012, 15: 373-380
- [31] Liebler DC. Introduction to Proteomics: Tools for the New Biology[M]. Beijing: Science Press, 2005: 19-20 (in Chinese)
利布莱尔 DC. 蛋白质组学导论—生物学的新工具[M]. 北京: 科学出版社, 2005: 19-20
- [32] Tomanek L. Environmental proteomics: changes in the proteome of marine organisms in response to environmental stress pollutants, infection, symbiosis and development[J]. Annual Review of Marine Science, 2011, 3: 373-399
- [33] Kolmeder CA, Devos WM. Metaproteomics of our microbiome—developing insight in function and activity in man and model systems[J]. Journal of Proteomics, 2014, 97: 3-16
- [34] Klaassens ES, de Vos WM, Vaughan EE. Metaproteomics approach to study the functionality of the microbiota in the human infant gastro intestinal tract[J]. Applied and Environmental Microbiology, 2007, 73: 1388-1392
- [35] Kolmeder CA, de Been M, Nikkilä J, et al. Comparative metaproteomics and diversity analysis of human intestinal microbiota testifies for its temporal stability and expression of core functions[J]. PLoS One, 2012, 7: e29913
- [36] Ferrer M, Ruiz A, Lanza F, et al. Microbiota from the distal guts of lean and obese adolescents exhibit partial functional redundancy besides clear differences in community structure[J]. Environmental Microbiology, 2013, 15: 211-226
- [37] Jagtap P, McGowan T, Bandhakavi S, et al. Deep metaproteomic analysis of human salivary supernatant[J]. Proteomics, 2012, 12: 992-1001
- [38] Morris RM, Nunn BL, Frazar CH, et al. Comparative metaproteomics reveals ocean-scale shifts in microbial nutrient utilization and energy transduction[J]. The ISME Journal, 2010, 4: 673-685
- [39] Schneider T, Keiblinge KM, Schmid E, et al. Who is who in litter decomposition? Metaproteomics reveals major microbial players and their biogeochemical functions[J]. The ISME Journal, 2012, 6: 1749-1762
- [40] Keiblinger KM, Wilhartitz IC, Schneider T, et al. Soil metaproteomics—a comparison of extraction protocols[J]. Soil Biology & Biochemistry, 2012, 54: 14-24
- [41] Lim YW, Schmieder R, Haynes M, et al. Metagenomics and metatranscriptomics: windows on CF-associated viral and microbial communities[J]. Journal of Cystic Fibrosis, 2013, 12: 154-164
- [42] Fouts DE, Pieper R, Szpakowski S, et al. Integrated next-generation sequencing of 16S rDNA and metaproteomics differentiate the healthy urine microbiome from a symptomatic bacteriuria in neuropathic bladder associated with spinal cord injury[J]. Journal of Translational Medicine, 2012, 10: 174
- [43] Urich T, Lanzén A, Stokke R, et al. Microbial community structure and functioning in marine sediments associated with diffuse hydrothermal venting assessed by integrated meta-omics[J]. Environmental Microbiology, 2013, 16(9): 2699-2710
- [44] Jaenicke S, Ander C, Bekel T, et al. Comparative and joint analysis of two metagenomic datasets from a biogas fermenter obtained by 454-pyrosequencing[J]. PLoS One, 2010, 6: 1-15
- [45] Zakrzewski M, Goesmann A, Jaenicke S, et al. Profiling of the metabolically active community from a production-scale biogas plant by means of high-throughput metatranscriptome sequencing[J]. Journal of Biotechnology, 2012, 158(4): 248-258
- [46] Hanreich A, Schimpf U, Zakrzewski M, et al. Metagenome and metaproteome analyses of microbial communities in mesophilic biogas-producing anaerobic batch fermentations indicate concerted plant carbohydrate degradation[J]. Systematic Applied Microbiology, 2013, 36(5): 330-338
- [47] Jia YY. Analysis of microbial dynamic change in compost by a metagenomic approach[D]. Jinan: Master's Thesis of Shandong University, 2012 (in Chinese)
贾洋洋. 利用宏基因组方法分析堆肥生境中微生物区系的变化[D]. 济南: 山东大学硕士学位论文, 2012