

doi:10.3969/j.issn.1672-5565.2013.04.07

# 宏基因组的生物信息分析

赵勇<sup>1</sup>, 黄劲松<sup>1</sup>, 宋新蕊<sup>1</sup>, 陈禹保<sup>1\*</sup>, 童贻刚<sup>2\*</sup>

(1.北京市计算中心,北京 100094; 2.军事医学科学院微生物流行病学研究所,北京 100071)

**摘要:**基于高通量测序的宏基因组学研究是近年来的研究热点之一。宏基因组的生物信息分析正在逐渐完善成熟,各种分析软件和流程的开发与应用,极大地促进了宏基因组研究的发展,特别是在遗传与进化、基因发现、宏基因组和人类疾病的相关研究等方面取得了显著成果。本文旨在结合宏基因组学的研究内容和研究方向,对宏基因组的生物信息分析方法进行综述,探讨宏基因组的生物信息分析面临的机遇和挑战。

**关键词:**宏基因组;生物信息学;高通量测序;

**中图分类号:**Q938.1, R857.3 **文献标识码:**A **文章编号:**1672-5565(2013)-04-282-05

## Bioinformatics analysis on metagenome

ZHAO Yong<sup>1</sup>, HUANG Jin-song<sup>1</sup>, SONG Xin-rui<sup>1</sup>, CHEN Yu-bao<sup>1\*</sup>, TONG Yi-gang<sup>2\*</sup>

(1, Beijing Computing Center, Beijing 100094, China;

2. Beijing Institute of Microbiology and Epidemiology, Academy of Military Medical Science, Beijing 100071, China)

**Abstract:** Utilizing the service provided by next generation sequencing, metagenomics has become one of the hotspots in current research. Following the progress on software development and analyzing platform, significant achievements had been made in genetics, gene finding, microbial communities and human disease investigation, and so on. This mini review aims to address the progress in bioinformatics analysis regarding metagenomics research, discuss the opportunities and challenges of modern metagenomics from a bioinformatics perspective.

**Keywords:** Metagenome; Bioinformatics; High Throughput Sequencing

宏基因组学(metagenomics)是通过非微生物培养的方法对环境微生物菌落进行调查研究的-门新兴学科,其主要研究对象为菌落中的细菌、古细菌、真菌和病毒等微生物,其主要目的是通过对微生物菌落中微生物的多样性、种群结构及其动态改变、各成员之间相互关系及与环境之间的相互关系等方面的分析,揭示更深层次的遗传与进化规律<sup>[1]</sup>。环境中微生物群落的研究经历了早期的以革兰氏染色和微生物培养为主要技术的阶段,由于微生物培养的瓶颈的存在,微生物群落的研究始终处于较初级的水平。DNA测序和PCR扩增技术的发展,极大地促进了人们对微生物群落的了解,以16/18S rRNA和OTU等为基础的数据库的建立与应用,使广大科研工作者能够对微生物菌落中物种的多样性

和种群结构等有更加深入的认识<sup>[2]</sup>。2005年以来, DNA二代测序技术的应用和生物信息学的发展,为微生物群落的研究注入了一股强大的动力,宏基因组学渐渐走入了人们的视线并成为当代生物医学研究的热点<sup>[3-8]</sup>。本文旨在通过对过去几年来二代测序在宏基因组学研究方面应用的总结,探讨宏基因组学在未来的研究方向和发展潜力。

## 1 宏基因组学的研究进展

宏基因组学的研究分为两个阶段,基于传统的克隆和鸟枪法测序分析的宏基因组研究和基于二代测序技术结合现代生物信息学分析的宏基因组研究阶段。

收稿日期:2013-07-01;修回日期:2013-07-25。

基金项目:北京市科学技术研究院海外人才专项资助(OTP-2012-001)。

作者简介:赵勇,男,博士,北京市特聘专家。研究方向:生物信息和计算生物学;E-mail:zhaoyong@bcc.ac.cn。

\*通讯作者:童贻刚,男,博士,研究员,教授。研究方向:生物信息学。E-mail:tong.yigang@gmail.com;

陈禹保,男,教授,研究方向:生物信息学;E-mail:chenyb@bcc.ac.cn。

## 1.1 传统的宏基因组学

传统的宏基因组学研究主要流程为:选择性提取环境中的样品,如土壤等,抽提样品中的 DNA/RNA,将样品中的 DNA 片段化,使用克隆技术将 DNA 随机连接到质粒或粘粒等载体,然后转化细菌并筛选克隆,通过对插入片段的测序和功能性分析,从而达到对微生物菌落的基本了解。在 PCR 技术发明之后,对微生物菌落的研究,尤其是对其多样性和种群结构方面的研究,开始倾向于 16 S/18 S rRNA 的测序和分析<sup>[9]</sup>。考虑到 16 S/18 S rRNA 测序和分析的分辨率不足以满足分析的需求,操作分类单位(Operational taxonomic unit, OTU)开始引入微生物群落的多样性分析<sup>[2]</sup>。由于微生物群落是一个多基因组的混合物,其中可能高达 99% 微生物不能单独分离培养,对个体微生物基因组全面的功能性研究远远落后于微生物群落分类学的发展。另外,传统的宏基因组分析费用高,周期长,极大地限制了宏基因组学研究的进展。

## 1.2 基于二代测序技术的宏基因组学

以 Roche 454 和 Illumina Hiseq 为代表的二代测序技术,成功实现了高通量和低费用的有效结合,将宏基因组学研究推向了前台。不同于第一代以 Sanger 末端终止法为原理的测序技术,二代测序技术基于边合成边测序的理念。Roche 454 主要表现为通过检测在合成过程中释放的焦磷酸信号读取基因的序列, Illumina 则是通过读取掺入染料的荧光信号获得模板 DNA 的序列<sup>[10]</sup>。Roche 454 最大的优势在于高质量测序的读长最长可达 1 000 bp 以上。Illumina Hiseq 系列测序仪最大的特点是在较低费用的基础上一次测序产生最高可达 600 G 的数据量。近两年来, Illumina 推出的 MiSeq 则以简单、快速、准确受到部分终端用户的好评<sup>[11]</sup>。

基于二代测序技术的宏基因组生物信息分析尚面临很多挑战。由于微生物群落包含物种的多样性和复杂性,无论生物信息分析软件和流程的开发,还是当前电脑计算能力的有限性,都对得到精确的分析结果造成了相当的困难。例如,通常情况下的基因组测序是针对单一物种的基因组,而微生物菌落的宏基因组测序事实上没有一个特定的范围<sup>[6]</sup>。微生物群落各成员的丰度不一和基因序列的相似度差异较大等都对目前使用的软件和分析方法构成了相当大的影响<sup>[3-4]</sup>。

## 2 宏基因组的生物信息学分析

总体来说,宏基因组学研究的的问题是:1) 微生物

群落的成分及其动态改变; 2) 微生物群落中的各种微生物的基因结构及其功能; 3) 微生物群落的各个成员之间及其与环境或宿主之间的关系; 4) 微生物群落研究成果在现实生活中的应用。

### 2.1 宏基因组的生物信息学分析内容

微生物群落的多样性分析:微生物群落中物种的多样性依然是目前研究的重点。对群落结构的研究,将有助于了解种群结构的稳定性,进而了解种群内物种间的相互依赖、相互制约的内在联系,为将来构建功能性种群服务。鉴于微生物群落是一个多物种的集合体,其中高达 95% 以上的微生物物种无法分离也不能独立培养,拼装出每个独立个体的基因组现在也无法实现,16 S 测序分析依然是现阶段微生物群落多样性和多态性分析的基石<sup>[4]</sup>。不可忽略的是,虽然基于 PCR 和高通量测序技术的微生物群落多样性分析结果与传统的鸟枪测序结果存在的偏差在可以接受的范围内<sup>[6]</sup>,但其物种的鉴别能力尚有很大的改善空间。因此,发现有效区分物种的基因标记物仍是当前需要追求的目标之一。

微生物群落的成分分析:在物种多样性分析的基础上,一个微生物群落中物种的丰度和密度也是宏基因组学研究工作共同关心的问题之一。某些物种丰度或密度的增减可能会改变种群的功能和特性<sup>[3]</sup>。基于高通量测序的数据分析,可以达到在个位数上了解微生物群落中物种的数目,使分析的结果更精确。对各个微生物种群随着环境变化的动态观察,由于有效对照的缺失,结果往往不是十分精确;目前的成分分析结果大多是基于统计学的推测,唯有微生物群落的主成分分析和群落之间的成分分析是相对可靠的。

微生物群落基因组组装分析:相较于微生物个体的基因组而言,目前的高通量测序依然是小片段测序,微生物群落中各个微生物个体的基因组拼装是当前无法完成的任务<sup>[12]</sup>。如果能够找到合适的方法,无疑将为开展微生物群落的研究提供极大的帮助。当前宏基因组的拼装原则上采用 assembling、Blast search 和 ORF 相结合的方法<sup>[7]</sup>。

微生物群落中个体的功能性分析:一个微生物个体在群落中的存在既有其偶然性,也有其必然性。近期受到广泛关注的人的肠道微生物课题(Human microbial project)的研究显示,肠道微生物群落中包含有与抗菌素的耐药性、药物的代谢和应激反应等相关的基因,这些基因对药物的代谢和人体的健康都有着很大的影响<sup>[13]</sup>。微生物和宿主之间相关关系的研究也揭示出一些新的信号传递系统和代谢系统<sup>[14]</sup>。基因的功能预测和实验验证依然是需要

解决的问题。单个基因的功能预测原则上依然是采用同源基因比对的原理,如 Blast 是基于序列比对,而 pfam, HMM, CDD, COG 等则是将基因序列与 profile 比对。

微生物群落的代谢通路分析:将微生物群落作为一个整体进行代谢通路分析是一个新发展方向。现有的代谢通路和信号传递系统都是基于单细胞的蛋白质相互作用级联反应构建的。微生物群落成员间、微生物群落和环境(或宿主)间存在有互动,例如代谢产物的互享互助,如何有效的建立微生物群落的代谢通路是目前尚待解决的问题<sup>[14]</sup>。

## 2.2 宏基因组常用生物信息分析软件

各种生物信息学分析软件有很多已经被应用于宏基因组的数据分析<sup>[26-27]</sup>,常规分析包括基因比对、序列装配、基因预测、统计分析等(见表1),这些分析都是根据数据的特性选择合适的软件并辅以适当的自主开发脚本来完成。除上述可以独立运行的软件外,有很多网站可以提供在线软件分析服务

表1 宏基因组分析相关的常用软件及网站

Table 1 Metagenomics analysis related software and website

软件功能	软件名称
基因比对	BLAST, MegaBLAST, BLAT, LAST, SSAHA2, SOAP2, BWA, BWA-SW, Bowtie, FR-HIT, MAUVE
序列装配	Meta-Velvet, META-IDBA, IDBA-UD, Genovo, MetAMOS, SOAPdenovo
基因预测	FragGeneScan, MetaGeneAnnotator (MGA), Glimmer, Glimmer-MG, HMMer3, BLAST, RAPSearch2, RAST, XBASE, RDP classifier, NBC, CARMA3, MEGAN, SOrt-ITEMS, GeneMark, MetaPhyler
统计分析	QIIME, CD-Hit, Mothur, EstimateS, SPADE, ShotgunFunctionalizeR, R-package, metagenomeSeq, GSEA, MetaPath, MetaPhyler, STAMP, HUMAnN
在线分析	MG-RAST <sup>[28]</sup> , IMG/M <sup>[29]</sup> , EBI metagenomics, Galaxy, RPD classifier
质量控制	FastQC, Trimmomatic, NGS QC Toolkit
数据库	NCBI NR/NT, UniRef100, KEGG, RDP Classifier, SEED databases, COG, eggNOG, TIGRFAMs
多样性分析	QIIME <sup>[24]</sup> , CloVR-16S, FANTOM <sup>[15]</sup>
功能性分析	MEGAN <sup>[15]</sup> , SmashCommunity <sup>[16]</sup> , STAMP <sup>[17]</sup> , shotgunFunctionalize R <sup>[18]</sup> , CloVR-metagenomics, CloVR-microbe <sup>[19]</sup>

## 3 宏基因组研究的应用前景

宏基因组研究有着广泛的应用前景,在理论研究和实际应用中都有很重要的研究价值。其中主要表现在以下几个方面:

进化分析:研究显示,物种之间的差异既有该物种本身基因组的差异,也有与其伴随的微生物的差异,某些相关微生物与宿主之间是共生关系。水平基因转移的发生是广泛存在的现象还是在特定条件下的产物?物种间的共生关系导致物种间水平基因转移的机制及其意义的阐明将对生物物种的进化理

(见表1)。对于在线分析服务,如果不考虑数据的安全性,数据传输速度及数据处理的容量仍是进行快速有效分析的限制因素。

### 2.3 宏基因组生物信息分析存在的问题

当前主流测序仪生成的数据原则上都需要进行严格的质量控制,数据质控的差异是造成实验验证差异的主要原因。数据的格式转换则是目前解决得最好的内容。对数据质控形成一个统一的标准并非不可能,我们相信,随着测序质量的提高,业界将有望达成这个目标。

尽管宏基因组数据存在多样化的问题,高通量数据生物信息分析流程化仍然是目前的总体趋势。流程化分析将有效增加数据分析的可靠性和结果间的可对比性。不可否认的是当前的数据分析流程多有一定的局限性和较高的错误率,特别是在序列拼装方面,但基于流程化的分析很容易通过相关参数的调整或软件的升级提高分析的精度。

论产生重要的影响<sup>[23]</sup>。

基因发现:基于高通量测序的宏基因组学研究为新基因的发现打开了一条新的通路。就种类和数量而言,微生物都是最大的基因资源库。传统的方法是在基因装配的基础上预测基因的存在,然后通过同源基因的比对和实验验证来确认基因的发现。对于宏基因组而言,由于测序的主要成分为多物种的混合物,基因组的装配尚无法完成。无论基于现有的哪个软件,所预测的基因的实验验证依然是个主要问题。宏基因组代谢与信号传导通路的研究是研究基础最贫乏的方向,KEGG、IPA 等数据库的构建都是基于单一物种,多物种之间的信号传导通路



尚处于不明朗状态,只能借鉴单一物种的数据。噬菌体是一类以特定细菌为宿主的病毒,在当前抗生素耐药菌泛滥而新的抗生素的研发遭遇瓶颈的情况下,新的噬菌体的发现有望为抗生素耐药菌感染提供新的治疗途径。噬病毒体(virophage)<sup>[20-22]</sup>是近年来的一个新发现,由于病毒本身的生物特性,开发以噬病毒体为基础的新药很有可能为流感、AIDS等病毒源性疾病的治疗带来新的曙光。

环境与生态研究:微生物群落生态系统的调查,特别是环境微生物的存在与环境生态改变的关系,是当前最受重视的课题之一<sup>[25]</sup>。目前的研究显示,冷热酸碱等极端环境都对环境微生物群落有着多种多样的影响。在环境污染的情况下,某些特殊功能的微生物群落也许对于难以降解的污染具有特殊的功效。环境微生物的存在与农业植物栽培和林业培育关系密切,固氮菌的存在对大豆生长的影响是农业栽培的实例。地球微生物组计划(Earth microbiome project)旨在广泛的开展对环境中的微生物的研究,以达到造福全球和全人类的目的<sup>[30]</sup>。

疾病和个体化医疗:目前粗略估计人体内微生物的数量是人的细胞数目数十倍甚至于百倍。大量的研究证据显示,人体微生物的种群和多样性与人体疾病的发生有着显著的相关性,例如肥胖、心血管疾病和肿瘤等。人类微生物组计划(Human microbiome project)的调查显示,肠道微生物中存在大量与药物代谢和分解相关的细菌,提示在个体化医疗方面不仅仅要考虑宿主基因组中药物的代谢相关基因,同时也需要考虑到消化道中微生物群落的存在和组成<sup>[27]</sup>。

生物信息软件研发和分析平台构建:随着测序技术的高速发展,优质的第三代测序仪将面世并且发展势头强劲,其特性表现为:更高的测序通量,更精确的测序质量,更长的测序长度。这些特性将为生物信息分析提出更多的需求。后两个特性将使得宏基因组的拼装成为可能,而前一个特性毫无疑问将需要电脑硬件和软件的进一步整合,否则,计算能力将成为分析的瓶颈<sup>[27]</sup>。

## 参考文献(References)

- [1] Chen K, Pachter L. Bioinformatics for Whole-genome Shotgun Sequencing of Microbial Communities [J]. *PLoS Computational Biology*, 2005, 1(2): 106-112.
- [2] Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ, Kulam-Syed-Mohideen AS, McGarrell DM, Marsh T, Garrity GM, Tiedje JM. The Ribosomal Database Project: Improved Alignments and New Tools for rRNA Analysis [J]. *Nucleic Acids Research*, 2009, 37(suppl 1): D141-D145.
- [3] Ley RE, Turnbaugh PJ, Klein S, Gordon JI. Microbial Ecology: Human Gut Microbes Associated with Obesity [J]. *Nature*, 2006, 444(7122): 1022-1023.
- [4] Huber JA, Mark Welch DB, Morrison HG, Huse SM, Neal PR, Butterfield DA, Sogin ML. Microbial Population Structures in The Deep Marine Biosphere [J]. *Science*, 2007, 318(5847): 97-100.
- [5] Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, Schlegel ML, Tucker TA, Schrenzel MD, Knight R, Gordon JI. Evolution of Mammals and Their Gut Microbes [J]. *Science*, 2008, 320(5883): 1647-1651.
- [6] Bartram AK, Lynch MD, Stearns JC, Moreno-Hagelsieb G, Neufeld JD. Generation of Multimillion-sequence 16s rRNA Gene Libraries From Complex Microbial Communities by Assembling Paired-End Illumina Reads [J]. *Applied and Environmental Microbiology*, 2011, 77(11): 3846-3852.
- [7] Mitreva M. Structure, Function and Diversity of the Healthy Human Microbiome [J]. *Nature*, 2012, 486: 207-214.
- [8] Diehl GE, Longman RS, Zhang JX, Breart B, Galan C, Cuesta A, Schwab SR, Littman DR. Microbiota Restricts Trafficking of Bacteria to Mesenteric Lymph Nodes by CX(3)CR1(hi) Cells [J]. *Nature*, 2013, 494(7435): 116-120.
- [9] Rondon MR, August PR, Bettermann AD, Brady SF, Grossman TH, Liles MR, Loiacono KA, Lynch BA, MacNeil IA, Minor C, Tiong CL, Gilman M, Osburne MS, Clardy J, Handelsman J, Goodman RM. Cloning the Soil Metagenome: A Strategy for Accessing the Genetic and Functional Diversity of Uncultured Microorganisms [J]. *Applied and environmental microbiology*, 2000, 66(6): 2541-2547.
- [10] Mardis E R. Next-Generation DNA Sequencing Methods [J]. *Annual Review of Genomics and Human Genetics*, 2008, 9: 387-402.
- [11] Sikkema-Raddatz B, Johansson LF, de Boer EN, Almomani R, Boven LG, van den Berg MP, van Spaendonck-Zwarts KY, van Tintelen JP, Sijmons RH, Jongbloed JD, Sinke RJ. Targeted Next-Generation Sequencing can Replace Sanger Sequencing in Clinical Diagnostics [J]. *Human Mutation*, 2013, 34(7): 1035-1042.
- [12] Dinsdale EA, Edwards RA, Hall D, Angly F, Breitbart M, Brule JM, Furlan M, Desnues C, Haynes M, Li L, McDaniel L, Moran MA, Nelson KE, Nilsson C, Olson R, Paul J, Brito BR, Ruan Y, Swan BK, Stevens R, Valentine DL, Thurber RV, Wegley L, White BA, Rohwer F. Functional Metagenomic Profiling of Nine Biomes [J]. *Nature*, 2008, 452(7187): 629-632.
- [13] Maurice CF, Haiser HJ, Turnbaugh PJ. Xenobiotics Shape the Physiology and Gene Expression of the Active Human Gut Microbiome [J]. *Cell*, 2013, 152(1): 39-50.
- [14] De Filippo C, Ramazzotti M, Fontana P, Cavalieri D. Bioinformatic Approaches for Functional Annotation and Pathway Inference in Metagenomics Data [J]. *Briefings in Bioinformatics*, 2012, 13(6): 696-710.
- [15] Sanli K, Karlsson FH, Nookaew I, Nielsen J. FANTOM: Functional and Taxonomic Analysis of Metagenomes [J]. *BMC Bioinformatics*, 2013, 14(1): 38-43.
- [16] Arumugam M, Harrington ED, Foerster KU, Raes J, Bork P. SmashCommunity: A Metagenomic Annotation and Analysis Tool [J]. *Bioinformatics*, 2010, 26(23): 2977-2978.
- [17] Parks D H, Beiko R G. Identifying Biologically Relevant Differences

- between Metagenomic Communities[J]. *Bioinformatics*, 2010, 26(6): 715–721.
- [18] Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blomberg SP, Webb CO: Picante: R tools for Integrating Phylogenies and Ecology[J]. *Bioinformatics*, 2010, 26(11): 1463–1464.
- [19] Angiuoli SV, Matalka M, Gussman A, Galens K, Vangala M, Riley DR, Arze C, White JR, White O, Fricke WF. CloVR: A Virtual Machine for Automated and Portable Sequence Analysis from the Desktop Using Cloud Computing[J]. *BMC Bioinformatics*, 2011, 12: 356.
- [20] Fischer MG, Suttle CA. A Virophage at the Origin of Large DNA Transposons[J]. *Science*, 2011, 332(6026): 231–234.
- [21] Boyer M, Yutin N, Pagnier I, Barrassi L, Fournous G, Espinosa L, Robert C, Azza S, Sun S, Rossmann MG, Suzan-Monti M, La Scola B, Koonin EV, Raoult D. Giant Marseillevirus Highlights the Role of Amoebae as a Melting Pot in Emergence of Chimeric Microorganisms [J]. *Proceeding of the National Academy of Sciences*, 2009, 106(51): 21848–21853.
- [22] La Scola B, Desnues C, Pagnier I, Robert C, Barrassi L, Fournous G, Merchat M, Suzan-Monti M, Forterre P, Koonin E, Raoult D. The Virophage as a Unique Parasite of the Giant Mimivirus [J]. *Nature*, 2008, 455(7209): 100–104.
- [23] Schönknecht G, Chen WH, Ternes CM, Barbier GG, Shrestha RP, Stanke M, Bräutigam A, Baker BJ, Banfield JF, Garavito RM, Carr K, Wilkerson C, Rensing SA, Gagneul D, Dickenson NE, Oesterhelt C, Lercher MJ, Weber AP. Gene Transfer from Bacteria and Archaea Facilitated Evolution of an Extremophilic Eukaryote[J]. *Science*, 2013, 339(6124): 1207–1210.
- [24] Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. QIIME Allows Analysis of High-throughput Community Sequencing Data[J]. *Nature Methods*, 2010, 7(5): 335–336.
- [25] Sachdev DP, Cameotra SS. Biosurfactants in Agriculture [J]. *Applied Microbiology Biotechnology*, 2013, 97(3): 1005–1016.
- [26] Teeling H, Glöckner FO. Current Opportunities and Challenges in Microbial Metagenome Analysis-A Bioinformatic Perspective [J]. *Briefings in Bioinformatics*, 2012, 13(6): 728–742.
- [27] Morgan XC, Huttenhower C. Human Microbiome Analysis [J]. *PLoS Computational Biology*, 2012, 8(12): e1002808.
- [28] Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M, Paczian T, Rodriguez A, Stevens R, Wilke A, Wilkening J, Edwards RA. The Metagenomics RAST Server-A Public Resource for the Automatic Phylogenetic and Functional Analysis of Metagenomes[J]. *BMC Bioinformatics*, 2008, 9: 386–394.
- [29] Markowitz VM, Ivanova NN, Szeto E, Palaniappan K, Chu K, Dalevi D, Chen IM, Grechkin Y, Dubchak I, Anderson I, Lykidis A, Mavromatis K, Hugenholtz P, Kyrpides NC. IMG/M: A Data Management and Analysis System for Metagenomes [J]. *Nucleic Acids Research*, 2008, 36(suppl 1): D534–D538.
- [30] Jansson JK, Prosser JI. Microbiology: The Life Beneath our Feet [J]. *Nature*, 2013, 494(7435): 40–41.