

基因组分析方法在微生物分类学中的应用

惠文彦 张和平*

(内蒙古农业大学 乳品生物技术与工程教育部重点实验室 内蒙古 呼和浩特 010018)

摘要: 细菌分类学始于 19 世纪后半叶, 当时主要是以表型标记和生理生化特性为基础的简单分类, 之后 DNA-DNA 分子杂交、16S rRNA 基因序列分析方法的出现给微生物分类带来了极大的便利。尽管如此, 这些分类学方法仍然存在一些局限性, 而基因组时代的到来, 为微生物分类带来了新思路。本文主要介绍了 5 种基于全基因组数据的微生物分类方法, 包括平均核苷酸同源性分析、核心基因组分析、最大唯一匹配指数分析、K 串组分矢量法和基因流动性分析, 并论述了这些方法在微生物分类学中的应用。

关键词: 基因组数据, 微生物分类, 系统发育

Application of genomic analysis in Microbial Taxonomy

HUI Wen-Yan ZHANG He-Ping*

(Key Laboratory of Dairy Biotechnology and Engineering, Ministry of Education, Inner Mongolia Agricultural University, Hohhot, Inner Mongolia 010018, China)

Abstract: The science of bacterial taxonomy started in late 19th century. The main methods of bacterial classification were simply based on phenotypic markers and physiological-biochemical characteristics then. The appearance of DNA-DNA hybridization and 16S rRNA gene sequence analysis brought a huge improvement to the field. However, recently, more and more studies have demonstrated that these approaches have some drawbacks. Fortunately, the emergence of the genomics technology has brought about novel concepts and insights into this area. This paper introduces the principle and application of five bacterial classification methods that are based on whole genome data, including average nucleotide identity analysis, core genome analysis, maximal unique matches index analysis, K-string composition vector method and genomic fluidity analysis.

Keywords: Genomic data, Microorganism taxonomy, Phylogeny

Foundation item: National Natural Science Foundation of China (No. 31430066)

*Corresponding author: Tel: 86-471-4319940; E-mail: hepingdd@vip.sina.com

Received: December 09, 2015; **Accepted:** February 24, 2016; **Published online** (www.cnki.net): March 07, 2016
基金项目: 国家自然科学基金项目(No. 31430066)

*通讯作者: Tel: 86-471-4319940; E-mail: hepingdd@vip.sina.com

收稿日期: 2015-12-09; 接受日期: 2016-02-24; 优先数字出版日期(www.cnki.net): 2016-03-07

细菌分类学始于 19 世纪后半叶, 起初主要根据形态特征、生理生化反应等方面的相似程度将细菌进行简单分类, 例如菌落形态、生长条件或者有无致病可能性等。然而这种经典的表型分类方法仅适用于菌种的常规鉴定, 在揭示细菌间的系统发育关系方面具有一定的局限性^[1]。20 世纪 60 年代到 80 年代间出现了一系列如生化分类法、数值分类法和 DNA-DNA 杂交(DNA-DNA hybridization, DDH)等分类学方法^[2]。尤其是 DDH 的出现, 使生物的基因型相关性成为了微生物分类的可靠依据^[3]。DDH 也一度被称为分类学中的“金标准”。然而这一方法也存在缺陷, 一方面对亲缘关系较远的物种灵敏度不够, 另一方面工作量大耗时长, 不适用于野生菌种^[1]。

20 世纪 80 年代后 DNA 扩增技术的发展为原核生物分类学研究注入了新的动力^[2,4], 许多新的分子标记技术被用来进行细菌的分类和鉴定。其中 16S rRNA 基因因其保守、稳定且不受水平基因转移的影响, 在众多分子标记中脱颖而出成为细菌分类鉴定的重要依据^[2]。随即应用而生的 16S rRNA 基因序列分析方法在微生物分类中得到了广泛应用^[5-7]。但是这种方法在区分两个相似性很高的近缘菌种或亚种时会引起偏差, 因此有学者提出全基因组分析的方法可以克服这个缺点^[8-9]。

生物学的研究进入基因组时代是在美国基因组研究所完成了对流感嗜血菌(*Haemophilus influenzae*)的全基因组测序工作之后^[10]。随着高通量 DNA 测序技术在微生物领域应用的不断成熟, 可以比表型特征和基因片段序列分析更精确的全基因组数据被应用到了微生物分类学研究中^[4,11]。随之也涌现出一些基于基因组数据的分析方法, 使微生物分类学在精确度、可靠性等方面取得了重要的突破。

本文重点介绍了基于全基因组数据分析的 5 种具有代表性的微生物分类学方法, 旨在为以后的研究提供一定的参考。

1 平均核苷酸同源性分析

平均核苷酸同源性(Average nucleotide identity, ANI)是指两个基因组之间同源基因的相似性^[4,12]。ANI 值可以通过两种运算方法得出: 一种是以 MUMmer 运算法则为基础(ANIm); 另一种是以 BLASTn 方法为基础(ANIb), 相比之下后者应用更为广泛^[11]。普遍认为亲缘关系较近的种群间 ANI 值至少为 70%–75%, 而定义一个种的 ANI 值需要达到 95%–96%以上, 并且引起人们关注的是这些 ANI 的数值与“金标准”——DDH 有紧密的对应关系^[3,13]。ANI 具有方便、耗费工作量少、错误率低、分辨率高的优点, 近年来得到了微生物分类学家们的青睐^[13]。

Zhang 及其研究团队^[14]通过计算比较 1 226 个细菌菌株全基因组序列间的 ANI 值后, 发现处于同一种的菌株间 ANI 为 93.6%, 同一属的菌株间 ANI 值为 83.6%, 而同一科菌株间 ANI 为 78.9%。作者认为在测序技术快速发展的这一机遇下 ANI 可以同时用于可培养和不可培养微生物的种属鉴定中。Kim 等^[12]通过研究 6 787 个基因组序列样本以期确定 16S rRNA 基因序列相似性的值和 ANI 值之间的相关性, 发现用来区分两个不同种的 98.65%的 16S rRNA 基因相似性值与 95%–96%的 ANI 值相对应。在这一方法良好准确性的前提下, ANI 在解决微生物分类的实际问题中发挥着重大的作用。Yi 等^[15]通过计算 ANI 值和 DDH 值来研究枯草芽孢杆菌(*Bacillus subtilis*)三个亚种之间的关系。比较发现, 亚种 *B. subtilis* subsp. *inaquosorum* 和 *B. subtilis* subsp. *spizizenii* 可以归为枯草芽孢杆菌(*Bacillus subtilis*)下的一个种, 且与亚种 *Bacillus subtilis* subsp. *subtitlis* 有较远的亲缘关系。Tamura 等^[16]计算了诺卡氏菌属(*Nocardia*)的 ANI 值(77.2%–95.2%), 并在这些值的基础上构建了诺卡氏菌属(*Nocardia*)的系统发育树, 这一系统发育树显示有 25 个分支, 并且所有分支的自展值都为 100%, 这些结果说明由 ANI 构建的系统发育树可以很好地反映诺卡氏菌属内菌种的进化关系。在大量的研

究结果和基因组数据的支持下,众多研究团队都已开发了 ANI 值的计算工具,比如本地运算软件 Jspecies (<http://www.imdea.uib.es/jspecies>), Gegenees (<http://www.gegenees.org/documentation.html>) 和在线计算工具 ANI caculator (<http://enve-omics.gatech.edu/>), EzGenome (<http://www.ezbiocloud.net/ezgenome/ani>)等。

内蒙古农业大学乳品生物技术与工程教育部重点实验室研究团队利用这一方法重构了不同乳酸菌的系统发育关系。在对 213 株乳杆菌及其相关属的研究中, Sun 等^[17]在计算乳杆菌(*Lactobacillus*) ANI 值的同时,研究了这些数值的分布规律。对这些数值进行统计可以看出:乳杆菌 ANI 值的分布在同一科、属间有差异,但同一纲、目间却有一定的重叠性。这一结果对于建立全新的乳杆菌分类体系有一定的参考作用。另一项研究中, Sun 等^[18]对双歧杆菌属(*Bifidobacterium*)基因多样性和进化关系进行了深入的探索。他们在利用 ANI 这一评价方法验证双歧杆菌属不同种间的分类地位时发现,之前划分为不同遗传谱系的 *B. indicum* 和 *B. coryneforme* 间的 ANI 值高达 98.27%, 远远大于定义种的界限,依据这一结果可以将二者合并为一个单独的种。

ANI 在微生物分类的应用中产生了大量可靠的结果,但是在实际操作中也不可避免地出现了一些问题,2009 年 Richter 等^[19]和 2014 年 Kim 等^[12]的研究中发现,在计算 ANI 时会出现不属于同一种的菌株间 ANI 值大于 96% 的异常情况。这种异常情况的发生是由以前分类偏差或者 ANI 自身缺陷引起,尚未有定论,还有待于进一步的研究。总之,由于其具有快速便捷、应用范围广、同时与 DDH 具有良好的关联性等优点,一些学者认为可以把 ANI 作为代替 DDH 成为下一代金标准的候选方法^[12]。

2 最大唯一匹配指数法

MUMi (Maximal unique matches index, MUMi) 是以两个基因组间的最大唯一配对数(Maximal unique and exact matches, MUMs)为基础并结合

MUMmer 这一生物信息软件来计算基因组距离的方法,可用于种内比较。MUMi 值将在 0 到 1 之间变动,值越小代表这两个基因组之间的亲缘关系越近。许多研究发现这种方法在衡量基因组亲缘关系时与 ANI 和 DDH 有很好的关联性, 0.33 ± 0.03 的 MUMi 对应于 $95\% \pm 0.5\%$ 的 ANI 和 70% 的 DDH^[20]。

Touzain 等^[21]使用 MUMi 的数据建立了系统发育树,发现大肠杆菌(*Escherichia coli*)的 MUMi 树有较高的准确性。Tiller 等^[22]在对 7 株分离自野生澳大利亚啮齿动物的布鲁氏菌株(*Brucella*)的研究中,首先将这 7 株菌定义为布鲁氏菌属下的一个新种;在后续研究中,基于 MUMi 的分析结果也显示这 7 株菌之间有很紧密的亲缘关系,这个结果与多位点序列分型的分析结果相一致。Deloger 等^[20]由 MUMi 得出了大肠杆菌(*Escherichia coli*)和金黄色葡萄球菌(*Staphylococcus aureus*)的系统发育关系,并且也得到了与采用多位点序列分型相似的结果。此外, Lugli 等^[23]通过计算双歧杆菌属(*Bifidobacterium*)所有基因组之间的 MUMi 值来验证试验中 ANI 值的准确性,计算得出双歧杆菌属(*Bifidobacterium*)的 MUMi 在 0.11 (菌株 *B. indicum* 188 LMG11587 和菌株 *B. coryneforme* LMG18911) 和 0.99 (菌株 *Bifidobacterium tsurumiense* 189 JCM13495 和菌株 *Bifidobacterium asteroides* LMG10735)之间,这些值均与 ANI 呈良好的对应关系。以上研究表明,在使用 MUMi 分析时,得出结果的准确性不亚于多位点序列分型和 ANI 分析等这些已被科研工作者认可的方法。Ormeño-Orrillo 等^[24]采用了 MUMi 的方法对根瘤菌(*Rhizobium*)进行了研究,并确定了 *Rhizobium tropici* CIAT899、*Rhizobium* sp. PRF 81 与其他根瘤菌的亲缘关系,研究结果表明, CIAT899 和 PRF81 为亲缘关系很近的菌株,而 *Rhizobium rhizogenes* K84 应属于其他分类群。

MUMi 法在探测种内相近菌株的差异性方面有较高的敏感性^[20],因此可用于种内相似菌株的鉴别中,此外还具有方便快捷的优点,并且与 DDH 和

ANI 有紧密的对应关系, 因此具有较大的优势。

3 核心基因组分析

核心基因组(Core genome)的定义是在菌株基因组中的较为稳定、不易于发生水平转移的、包括看家基因在内的基因集, 这些基因大都具有种属特异性^[1,25]。由于核心基因组进化缓慢并且受种内重组的影响小, 所以可以将之用于鉴定细菌菌株之间的亲缘关系中^[26-27]。构建核心基因组的方法是利用 BLAST 工具进行相似性比对, 并且在比对同时设定研究所需要的阈值^[17]。

对于核心基因组在微生物分类中的应用, 通常是将核心基因组进行比对, 并据此绘制系统发育树。Kaas 等^[28]在研究大肠杆菌(*Escherichia coli*)基因组数据后, 证明了由核心基因组建立的系统发育关系具有良好的分辨率, 并且指出使用核心基因组建树的方法应普遍应用于大肠杆菌的分型中。Leekitcharoenphon 等^[27]从 73 个已公布的肠道沙门氏菌(*Salmonella enterica*)基因组中识别了 2 882 个核心基因, 发现根据不同核心基因建立的同一树比 16S rRNA 基因和多位点序列分型分析方法的分辨率更高。另外 Zhang 及其研究团队^[29]完成了对 28 个乳酸菌基因组的 232 个直系同源基因的系统发育分析, 这一分析采用了两种不同的空位分配标准和两种不同的建树方法, 结果产生了可靠且平行性良好的拓扑结构。他们研究发现这些待检的乳酸菌种大体上可以分为两个种群: 第一个种群包括肠球菌科(Enterococcaceae)和链球菌科(Streptococcaceae); 第二个种群包括乳杆菌科(Lactobacillaceae)和明串珠菌科(Leuconostocaceae), 使用核心基因组分析得到的这些结论与之前的研究结果是一致的^[30-31]。van den Bogert 等^[32]研究了链球菌(*Streptococcus*)核心基因组的子集基因并用来建树, 得到了由 5 个明显的分支构成的系统发育树, 其中 3 个分支的分类与之前使用飞行时间质谱分析和 16S rRNA 基因分析的结果相一致。

本研究团队在研究双歧杆菌属(*Bifidobacterium*)

的 45 个模式菌株的基因多样性和进化关系时, 采用了基于核心基因组的分类学方法。在建立了包含有 402 个基因的核心基因集之后, 以这一基因集为基础绘制了双歧杆菌属(*Bifidobacterium*)的系统发育树。结果表明, 基于核心基因构建的系统发育树具有很高的分辨率^[18]。后续研究中, 团队成员 Sun 等^[17]又以 213 个乳杆菌和 10 个相关属的基因组为研究对象, 鉴定得到由 73 个基因构成的核心基因组, 并以这些基因为基础通过最大似然法构建了乳杆菌属系统发育树。这一结果不仅真实再现了乳杆菌不同种间的系统发育关系, 也为我们理解乳杆菌与相关属间的发育关系提供了确实的证据。例如: 嗜果糖乳酸细菌属(*Fructobacillus*)和魏斯氏菌属(*Weissella*)应是乳杆菌类群的成员。

由以上研究结果可以看出, 核心基因组分析可广泛应用于原核生物种属及属以上微生物的研究中, 并且也可用于揭示各群体内的进化关系。然而, 对于这一方法的争议主要集中在如何定义核心基因组, 定义核心基因组的界限是什么, 到底这些基因需要保守到什么程度才可以列入核心基因组的范畴中。因此, 核心基因组这一方法在微生物分类的应用中还有待进一步完善。

以上 3 种方法在应用时都需要将微生物基因组进行两两比对, 而接下来所提及的两种方法 K 串组分矢量法和基因流动性分析则具有免对比的特点。

4 K 串组分矢量法

K 串(K-string)组分矢量法是一种通过计算蛋白质序列或者 DNA 序列中寡肽的出现频率来推断基因组相关性的方法^[33]。K 串是指在一个长度为 N 的 DNA 或者 RNA 序列中长度为 K 的连续结构, 对于基因序列有 4^K 个可能的 K 串^[34]。K 可以选择 1 到 N 之间的任意数。长 K 值侧重于强调种的特异性, 而短 K 值重点在于反映不同种之间的共同特征^[35]。通过观察由此方法建立的系统发育树可以看出, K=5 或者 K=6 是最适用于细菌系统发育构建的 K 值, 并且也有学者指出没有必要选择大于 7 的 K 值^[36]。

在2014年Chan等^[37]将熵原理应用到组分矢量方程中并对原有方法进行了改善,使组分矢量法更加准确,应用范围更加广泛。为了研究人员更好地利用这一方法,网络服务平台CVTree (<http://tlife.fudan.edu.cn/cvtree/>)在2004年初步构建,并在2009年得到完善^[38-39]。

Qi等^[33]利用组分矢量法分析了原核生物的系统发育关系,得到的结果与利用16S rRNA基因分析相似。另外,在Chan等^[40]的研究中,在K=5时绘制的不动杆菌属(*Acinetobacter*)种系发生关系图在种的水平上与由ANI建立的种系发生关系大体一致。以上结果均证明了K串组分矢量法能够很有效地应用于原核生物的系统发育分析中。在早期的表型实验和近期的若干分子研究中,志贺氏杆菌(*Shigella*)和大肠杆菌(*Escherichia coli*)被认为是相关性很高的分类群,因此将志贺氏杆菌(*Shigella*)归于大肠杆菌(*Escherichia coli*)中。然而Zuo等^[35]使用组分矢量法进行了一系列的试验后发现4个志贺氏杆菌种并不属于大肠杆菌,准确说来应该是大肠杆菌属中的另一个种,这一研究说明K串组分矢量法可以在一定程度上修正一些分类偏差。

虽然K串组分矢量法在种系发生关系分析中取得了重要的成就,并且适用范围广,可用于细菌、古菌及真菌的鉴定中^[33]。然而研究发现并不是所有的K串都可以有效地用于种系发生树的建立^[41],并且水平基因转移也会影响到这一方法的准确性^[35]。

5 基因流动性分析

基因流动性 ϕ (Genome fluidity)是指在一组个数为N的基因组中,独特基因数和总体基因的比值,它是属于一种基因容量的比对方法,可简单认为是基因组之间非重叠部分的度量,其最初主要用于衡量微生物种群内基因多样性^[42-43]。基因流动性的数值越小代表对应基因组的特异性程度越小,当基因流动性的值 ϕ 为0.1的时候,代表一对基因组有10%的独特基因,相应应有90%的共同基因^[43-44]。基因流动性这一方法在计算时,它的优点在于不会

因提供的基因序列过短或是鉴定基因同源性方法的不同而影响其结果的精确性^[42]。

Zwick等^[44]计算了蜡状芽孢杆菌(*Bacillus cereus*)的基因流动性,其结果显示蜡状芽孢杆菌(*Bacillus cereus*)的基因流动性的值 $\phi=0.22$,说明这个属的分类地位处于世界性物种如脑膜炎奈瑟氏菌(*Neisseria meningitidis*)或者大肠杆菌(*Escherichia coli*) ($\phi=0.3$)和狭分布物种如金黄色葡萄球菌(*Staphylococcus aureus*) ($\phi=0.15$)之间。除了实验事实之外,Lan等^[45]介绍了一个基因组信息的简易比较平台POGO-DB (<http://pogo.ece.drexel.edu/>),基因流动性是此平台所使用的比对方法之一。

基因流动性的准确性同样也会受水平基因转移的影响^[40],在一定程度上制约了这一方法的应用。

6 展望

上文中论述的这几种基于基因组数据的方法,虽然它们的应用领域各有侧重,但在微生物分类中均有大量成功的实例。ANI分析因与DDH有良好的关联性而有望成为下一代的金标准。MUMi分析和ANI同为计算基因组指数的方法,与DDH也同样有一定的联系这一特性增加了其应用优势。核心基因组分析由于其所选基因有较强的保守性,因而在建树中有着较高的分辨率和准确性。基于基因含量的两种方法——K串组分矢量法和基因流动性分析,虽然目前在微生物分类中应用还比较少,但是由于其算法简便、免对比的特点因而有一定的应用前景。在这5种方法中ANI分析和核心基因组分析的应用最为成熟。

微生物分类学的最终目标是建立一个可以反映“自然秩序”的系统。因此,微生物的分类技术在不断的探索和争论中不断提高。目前已出现的分类学方法已经在很大程度上修正了之前存在的偏差,但是这些方法仍存在较大的改善空间。另一方面,由于目前对基因组序列信息所包含的生物学意义了解得不够透彻,所以在应用这些基因组数据时应更加慎重。总体来说,以基因组数据为基础的这

些分类方法具有快速、准确、节省人力等优点, 相信会得到更广泛的应用。在今后的分类学方法研究中, 为了提高微生物分类的合理性和准确性, 除了对已有方法改进完善并且探索新方法之外, 也需要考虑将这些基因组数据的方法同形态学、生理生化特性等表型标记方法联合使用。

参 考 文 献

- [1] Schleifer KH. Classification of *Bacteria* and *Archaea*: past, present and future[J]. *Systematic and Applied Microbiology*, 2009, 32(8): 533-542
- [2] Ramasamy D, Mishra AK, Lagier JC, et al. A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species[J]. *International Journal of Systematic and Evolutionary Microbiology*, 2014, 64(Pt 2): 384-391
- [3] Goris J, Konstantinidis KT, Klappenbach JA, et al. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities[J]. *International Journal of Systematic and Evolutionary Microbiology*, 2007, 57(Pt 1): 81-91
- [4] Konstantinidis KT, Tiedje JM. Prokaryotic taxonomy and phylogeny in the genomic era: advancements and challenges ahead[J]. *Current Opinion in Microbiology*, 2007, 10(5): 504-509
- [5] Yu J, Wang WH, Menghe BLG, et al. Diversity of lactic acid bacteria associated with traditional fermented dairy products in Mongolia[J]. *Journal of Dairy Science*, 2011, 94(7): 3229-3241
- [6] Liu WJ, Bao QH, Jirimutu, et al. Isolation and identification of lactic acid bacteria from Tarag in Eastern Inner Mongolia of China by 16S rRNA sequences and DGGE analysis[J]. *Microbiological Research*, 2012, 167(2): 110-115
- [7] Zhang WY, Lü Q, Xu HY, et al. Phylogenetic analysis of closely related *Enterococcus faecalis*, *Enterococcus faecium* species based on partial housekeeping genes[J]. *Microbiology China*, 2014, 41(2): 297-303 (in Chinese)
张文羿, 吕端, 徐海燕, 等. 粪肠、屎肠球菌及相近种部分持家基因的系统发育分析[J]. *微生物学通报*, 2014, 41(2): 297-303
- [8] Rosselló-Móra R, Amann R. The species concept for prokaryotes[J]. *FEMS Microbiology Reviews*, 2001, 25(1): 39-67
- [9] Tindall BJ, Rosselló-Móra R, Busse HJ, et al. Notes on the characterization of prokaryote strains for taxonomic purposes[J]. *International Journal of Systematic and Evolutionary Microbiology*, 2010, 60(Pt 1): 249-266
- [10] Fleischmann RD, Adams MD, White O, et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd[J]. *Science*, 1995, 269(5223): 496-512
- [11] Chun J, Rainey FA. Integrating genomics into the taxonomy and systematics of the *Bacteria* and *Archaea*[J]. *International Journal of Systematic and Evolutionary Microbiology*, 2014, 64(Pt 2): 316-324
- [12] Kim M, Oh HS, Park SC, et al. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes[J]. *International Journal of Systematic and Evolutionary Microbiology*, 2014, 64(Pt 2): 346-351
- [13] Rodríguez-R LM, Konstantinidis KT. Bypassing cultivation to identify bacterial species[J]. *Microbe*, 2014, 9(3): 111-118
- [14] Zhang W, Du PC, Zheng H, et al. Whole-genome sequence comparison as a method for improving bacterial species definition[J]. *The Journal of General and Applied Microbiology*, 2014, 60(2): 75-78
- [15] Yi HN, Chun J, Cha CJ. Genomic insights into the taxonomic status of the three subspecies of *Bacillus subtilis*[J]. *Systematic and Applied Microbiology*, 2014, 37(2): 95-99
- [16] Tamura T, Matsuzawa T, Oji S, et al. A genome sequence-based approach to taxonomy of the genus *Nocardia*[J]. *Antonie van Leeuwenhoek*, 2012, 102(3): 481-491
- [17] Sun ZH, Harris HMB, McCann A, et al. Expanding the biotechnology potential of lactobacilli through comparative genomics of 213 strains and associated genera[J]. *Nature Communications*, 2015, 6: 8322
- [18] Sun ZH, Zhang WY, Guo CY, et al. Comparative genomic analysis of 45 type strains of the Genus *Bifidobacterium*: a snapshot of its genetic diversity and evolution[J]. *PLoS One*, 2015, 10(2): e0117912
- [19] Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2009, 106(45): 19126-19131
- [20] Deloger M, Karoui ME, Petit MA. A genomic distance based on MUM indicates discontinuity between most bacterial species and genera[J]. *Journal of Bacteriology*, 2009, 191(1): 91-99
- [21] Touzain F, Denamur E, Médigue C, et al. Small variable segments constitute a major type of diversity of bacterial genomes at the species level[J]. *Genome Biology*, 2010, 11(4): R45
- [22] Tiller RV, Gee JE, Frace MA, et al. Characterization of novel *Brucella* strains originating from wild native rodent species in North Queensland, Australia[J]. *Applied and Environmental Microbiology*, 2010, 76(17): 5837-5845
- [23] Lugli GA, Milani C, Turroni F, et al. Investigation of the evolutionary development of the genus *Bifidobacterium* by comparative genomics[J]. *Applied and Environmental Microbiology*, 2014, 80(20): 6383-6394
- [24] Ormeño-Orrillo E, Menna P, Almeida LGP, et al. Genomic basis of broad host range and environmental adaptability of *Rhizobium tropici* CIAT 899 and *Rhizobium* sp. PRF 81 which are used in inoculants for common bean (*Phaseolus vulgaris* L.)[J]. *BMC Genomics*, 2012, 13: 735
- [25] Lefebvre T, Bitar PDP, Suzuki H, et al. Evolutionary dynamics of complete *Campylobacter* pan-genomes and the bacterial species concept[J]. *Genome Biology and Evolution*, 2010, 2: 646-655
- [26] Korona R. Gene dispensability[J]. *Current Opinion in Biotechnology*, 2011, 22(4): 547-551
- [27] Leekitcharoenphon P, Lukjancenko O, Friis C, et al. Genomic variation in *Salmonella enterica* core genes for epidemiological typing[J]. *BMC Genomics*, 2012, 13: 88
- [28] Kaas RS, Friis C, Ussery DW, et al. Estimating variation within the genes and inferring the phylogeny of 186 sequenced diverse *Escherichia coli* genomes[J]. *BMC Genomics*, 2012, 13: 577
- [29] Zhang ZG, Ye ZQ, Yu L, et al. Phylogenomic reconstruction of lactic acid bacteria: an update[J]. *BMC Evolutionary Biology*, 2011, 11: 1
- [30] Claesson MJ, van Sinderen D, O'Toole PW. Lactobacillus phylogenomics—towards a reclassification of the genus[J]. *International Journal of Systematic and Evolutionary Microbiology*, 2008, 58(Pt 12): 2945-2954
- [31] Callanan M, Kaleta P, O'callaghan J, et al. Genome sequence of *Lactobacillus helveticus*, an organism distinguished by selective gene loss and insertion sequence element expansion[J]. *Journal of Bacteriology*, 2008, 190(2): 727-735
- [32] van den Bogert B, Boekhorst J, Herrmann R, et al. Comparative genomics analysis of *Streptococcus* isolates from the human small intestine reveals their adaptation to a highly dynamic ecosystem[J]. *PLoS One*, 2013, 8(12): e83418
- [33] Qi J, Wang B, Hao BL. Whole proteome prokaryote phylogeny without sequence alignment: a K-string composition approach[J]. *Journal of Molecular Evolution*, 2004, 58(1): 1-11
- [34] Eloumi M, Zomaya AY. Algorithms in Computational Molecular Biology: Techniques, Approaches and Applications[M]. Hoboken, New Jersey: John Wiley & Sons, 2011: 601-619
- [35] Zuo GH, Zhao X, Hao BL. *Shigella* strains are not clones of *Escherichia coli* but sister species in the genus *Escherichia*[J].

- Genomics, Proteomics & Bioinformatics, 2013, 11(1): 61-65
- [36] Hao BL. CVTree: A Whole-genome-based and alignment-free approach to microbial phylogeny[J]. International Journal of Modern Physics: Conference Series, 2012, 9: 1-10. doi: 10.1142/S2010194512005041
- [37] Chan RH, Chan TH, Yeung HM, et al. Composition vector method based on maximum entropy principle for sequence comparison[J]. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2012, 9(1): 79-87
- [38] Qi J, Luo H, Hao BL. CVTree: a phylogenetic tree reconstruction tool based on whole genomes[J]. Nucleic Acids Research, 2004, 32(Suppl 2): W45-W47
- [39] Xu Z, Hao BL. CVTree update: a newly designed phylogenetic study platform using composition vectors and whole genomes[J]. Nucleic Acids Research, 2009, 37(Suppl 2): W174-W178
- [40] Chan JZM, Halachev MR, Loman NJ, et al. Defining bacterial species in the genomic era: insights from the genus *Acinetobacter*[J]. BMC Microbiology, 2012, 12: 302
- [41] Yuan JB, Zhu QM, Liu B. Phylogenetic and biological significance of evolutionary elements from metazoan mitochondrial genomes[J]. PLoS One, 2014, 9(1): e84330
- [42] Kislyuk AO, Haegeman B, Bergman NH, et al. Genomic fluidity: an integrative view of gene diversity within microbial populations[J]. BMC Genomics, 2011, 12: 32
- [43] Snipen L, Ussery DW. A domain sequence approach to pangenomics: applications to *Escherichia coli*[J]. F1000Research, 2012, 1: 19
- [44] Zwick ME, Joseph SJ, Didelot X, et al. Genomic characterization of the *Bacillus cereus* sensu lato species: backdrop to the evolution of *Bacillus anthracis*[J]. Genome Research, 2012, 22(8): 1512-1524
- [45] Lan YM, Morrison JC, Hershberg R, et al. POGO-DB—a database of pairwise-comparisons of genomes and conserved orthologous genes[J]. Nucleic Acids Research, 2014, 42(D1): D625-D632

2016 年中国微生物学会及各专业委员会学术活动计划表(2-2)

序号	会议名称	主办/协办单位	时间	人数	地点	联系方式
13	第七届中国微生物学大会暨微生物学与免疫学论坛	中国微生物学会临床微生物学专业委员会	9 月	400	待定	王苗苗 18758810661
14	2016 年全国青年病毒学者学术年会	中国微生物学会病毒学专业委员会	9 月	200	待定	吴莹 010-64807688
15	首届临床微生物学与医院感染论坛	中国微生物学会临床微生物学专业委员会	9 月	350	待定	王苗苗 18758810661
16	2016 年微生物与人类健康学术研讨会	中国微生物学会医学微生物学与免疫学专业委员会	9 月	200	上海	胡福泉 13594616136
17	第十一届中国微生物学会兽医微生物学专业委员会委员会议	中国微生物学会兽医微生物学专业委员会	10 月	400	待定	丁家波 13683505108
18	第十三届国际工业微生物遗传学大会	中国微生物学会	10 月 16-20 日	400	湖北武汉	孙雪 027-68756642
19	2016 年中国微生物学会学术年会	中国微生物学会	10 月	600	陕西西安	杨海花 010-64807200
20	食品酿造技术与产业发展学术报告会	中国微生物学会酿造分会	10 月	200	广东汕头	张秀梅 13503213265
21	第 14 届中日韩国际酶工程学术会议	中国微生物学会酶工程专业委员会	11 月	200	广西南宁	欧阳浩森 010-64807420
22	第十九次全国环境微生物学学术研讨会	中国微生物学会环境微生物学专业委员会	11 月	500	重庆	蒋建东 13915976780
23	中国微生物与白酒酿造技术研讨会	中国微生物学会工业微生物学专业委员会	12 月	200	待定	010-53218310
24	第六届全国微生物基因组学学术研讨会	中国微生物学会农业微生物学专业委员会	12 月	200	海南乐东	吴悦 027-87287254