

根瘤菌系统发育分类方法研究进展*

郑君芳 朱万孚**

(北京大学基础医学院病原生物学系 北京 100083)

摘要: 根瘤菌系统发育地位的判断在根瘤菌新属种的确认中起重要作用。其中, 16S rRNA 序列分析是关键技术。然而, 新近的研究对采用 16S rRNA 全序列推测的系统发育关系的准确性提出质疑。本实验室的工作表明基因组物理结构分析将能更客观反映其系统发育关系。基因组序列分析在帮助澄清某些根瘤菌属种的系统发育地位中也起重要作用。

关键词: 根瘤菌, 系统发育, 分类方法

中图分类号: Q93 **文献标识码:** A **文章编号:** 0253-2654 (2004) 02-0126-04

根瘤菌与宿主植物共生结瘤固氮, 从而免于使用氮肥, 因而对农业生产和环境保护起重要作用。自根瘤菌发现 100 多年来, 其资源调查和分类的研究一直是个热点。近年来新分类方法的应用促进了根瘤菌新分类群的不断发现, 也使根瘤菌的分类渐趋科学合理。各种分类方法互相验证, 共同构成多相分类体系, 即首先采用依据表型特征和遗传特征的分类方法进行根瘤菌的初步聚群, 然后选取群内代表菌株与同群内菌株及相关属种的代表菌株进行杂交, 并测定其 G + C mol% 和 16S rRNA 序列, 判断其系统发育地位。系统发育地位的判断在根瘤菌新属种的确认中起重要作用。

1 依据基因序列分析的系统发育分类方法

依据基因序列分析进行分类是目前根瘤菌的主要系统发育分类方法。所依据的基因有核糖体基因、其它染色体基因和质粒基因 3 类, 其中以核糖体基因中的 16S rRNA 序列为主, 其余基因对其进行验证或补充。

1.1 16S rRNA 基因 16S rRNA 基因进化速率缓慢, 基因产物功能保守, 因而应用于细菌的自然分类, 即依据其进化关系(系统发育关系, phylogenetic relationships)进行分类。Woese 在 20 世纪 70 年代末开始应用 16S rRNA 序列分析技术对细菌进行系统发育分类的研究^[1]。1991 年 Young 等开始依据 16S rRNA 基因序列比较分析对根瘤菌进行系统发育分类研究^[2]。随后, 在确认许多根瘤菌的新属种时, 16S rRNA 基因起到主要作用, 如将 *Sinorhizobium* 属从 *Rhizobium* 属中分出^[3], 建立新属 *Allorhizobium*^[4], 合并 *Agrobacterium* 和 *Allorhizobium* 到 *Rhizobium*^[5], 分开 *B. japonicum* 和 *B. elkanii*^[6], 鉴别 *Methylobacterium nodulans*^[7] 和变形杆菌门 (Proteobacteria) β 纲根瘤菌 *Burkholderia* sp.^[8] 等。

但也不断有研究对 16S rRNA 基因的分类作用提出质疑。首先, 有研究表明不同种的菌株 16S rRNA 基因存在重组现象^[9,10], 说明依据 16S rRNA 基因推论系统发育关系有可能出现错误。或者因为同一基因组内 16S rRNA 序列本身可能就是嵌合序列^[11], 使其不同片段与不同种的菌株近似, 从而降低了利用 16S rRNA 基因进行分类的可靠性。其

* 教育部教育振兴行动计划专项基金 (985 项目)

** 联系人

收稿日期: 2003-06-09, 修回日期: 2003-08-05

次, 16S rRNA 分子较小, 包含的信息量较少, 对于亲缘关系较近的细菌可能分辨率不足, 所以尽管 16S rRNA 基因可明确地将根瘤菌分类到属并定义属间关系, 但常不能分辨关系较近的种间关系^[12]。最后, 虽然依据 16S rRNA 序列可将根瘤菌属从系统发育上彼此分开, 但可能由于细菌之间发生的基因横向转移, 使一些依据传统分类方法无属的属种与根瘤菌属聚在一起, 如 *Rhodospseudomonas* (光合细菌) 和 *Afpia* (病原菌)^[10]。

最近 van Berkum 等通过对 α 变形杆菌门一组代表菌株 *rm* 操纵子内 3 个位点 (16S rRNA, 23S rRNA 和 IGS) 进行序列分析, 并分别用 3 个序列构建系统发育树, 发现 16S rRNA 基因所建的系统发育树与 23S rRNA 和 IGS 所建的系统发育树明显不同, 因而认为依据 16S rRNA 基因序列差异得出的系统发育地位值得怀疑, 据此变更细菌名称及其分类地位不一定合理。另外, *Mesorhizobium* 和 *Bradyrhizobium* 属细菌的 16S rRNA 基因序列比对、*Mesorhizobium* 和 *Sinorhizobium* 属细菌的 16S rRNA 基因序列比对结果表明这两对细菌的 16S rRNA 序列存在重组现象^[13]。因此, 在估计基因组间的系统发育关系时, 应采取更保守的方法, 即应依据更多基因位点, 并进行比较分析来得出分类结论^[13]。如目前尝试的 23S rRNA 基因序列分析, 将提供更多的系统发育方面的信息。

1.2 其它染色体基因 依据染色体保守基因 *atpD*、*recA* 和 GSI 推论的根瘤菌系统发育关系基本与 16S rRNA 基因推论的系统发育关系一致^[14,15], 但染色体保守基因 GSII 序列推论的系统发育关系表明, *R. leguminosarum* 与慢生根瘤菌 *B. japonicum* 和快生根瘤菌 *S. fredii* 间的进化距离没有多大差别, *M. huakuii* 在 *Mesorhizobium* 群外^[14]。这种关系与 16S rRNA 基因推论的系统发育关系很不一致, 可能是由于全基因替换 (replacement) 所致^[14]。这说明染色体基因也不能十分准确地表明根瘤菌的系统发育关系。

1.3 质粒基因 根瘤菌的共生质粒基因包括结瘤 (*nod*) 和固氮 (*nif*) 基因。依据根瘤菌共同结瘤基因 *nodABC* 和其调节基因 *nodD* 得到的系统发育关系主要是反映了其宿主范围的差异, 而不能反映当前根瘤菌的系统发育分类状况^[16], 因此在系统发育分类中的作用不大。

依据特定基因序列分析的系统发育分类方法还存在一个共同的问题, 即如果采用的算法或分子序列不同, 则计算出的系统发育关系也很可能有差异^[17]。因此, 依据基因序列进行系统发育分类研究时, 应结合多项指标。

2 依据基因组结构分析的系统发育分类方法

由于 16S rRNA 基因系统发育分类方法存在上述问题, 研究者一直在寻求更为稳定的基因组特征来对根瘤菌进行更客观的系统发育自然分类。在目前还不能获得大量细菌全基因组序列的情况下, 最好的办法是对一些代表性细菌全基因组序列进行序列分析, 对其它大量细菌绘制物理图谱, 并加以比较。

最近, *Salmonella* 及 *Pasteurella* 等菌属的基因组分析初步表明, 细菌基因组结构在进化上较稳定, 可以在某种程度上反映细菌间的系统发育关系^[18], 从而有望据此完善细菌系统发育分类系统^[18,19]。这种方法的理论依据为: 细菌中存在着系统发育群 (phylogenetic clusters), 这反映了它们具有共同的祖先; 细菌的某些基因组特征在进化过程中相对稳定, 这样的特征可用作系统发育标记; 而有些进化事件, 如碱基替换, 基因组重排, 获取外源 DNA 和祖先基因组 DNA 部分片段的缺失, 可逐步改变基因组, 使细菌从祖先种系分离, 并最终适应新环境成为新种。对基因组进行分析, 获得基因组物

理结构,可检测基因组变化和表明这些事件,因此基因组物理结构可以作为一种有用的系统发育标记。具体方法为:用 *I-CeuI* 核酸内切酶切割细菌基因组 DNA,然后进行脉冲场凝胶电泳分离消化后的大片段,依据不完全消化产物探明细菌基因组结构^[20]。此技术可反映研究个体间全基因组的差异信息,清晰表明细菌系统发育群,即亲缘关系较近的细菌有细微又明确的基因组物理结构上的差异,且操作技术简单,不需要繁琐的分子克隆手段即可在普通实验室进行,能高效地对根瘤菌的大量菌株进行全基因组信息比较并分群,从而对根瘤菌进行系统发育分类。基因组结构聚群初步研究结果表明,*I-CeuI* 酶切确定的根瘤菌模式菌株基因组结构分类结果基本与 16S rRNA 序列分析结果一致,但又各有特点^[21]。

由于这项研究刚刚开始,尚需要分析大量根瘤菌基因组结构来证实其系统发育分类价值,但基因组结构分析有望成为未来的广泛应用于根瘤菌的系统发育分类研究的方法。

3 依据基因组序列分析的系统发育分类方法

对细菌全基因组进行测序,然后通过比较研究推论其进化历史似乎是最理想的自然分类方法,因为对一个种进行多株菌的全序列分析可揭示定义种的共同基因核心^[22]。随着更多的根瘤菌全基因组序列分析结果的公布,应用基因组序列对某些根瘤菌属种的系统发育位置进行确认已变得逐渐可行。目前已完成了 4 株根瘤菌基因组 (*Sinorhizobium meliloti*^[23]、*Mesorhizobium loti*^[24]、*Bradyrhizobium japonicum*^[25] 和 *Agrobacterium tumefaciens*^[26,27]) 和 5 个共生质粒 (*Sinorhizobium* sp. NGR234, *S. meliloti* *SymA* 和 *SymB* 和 *M. loti* 的 2 个大质粒^[24]) 的全序列分析。根瘤菌全基因组序列分析可对其基因组及其结构提出新认识^[28],也对其分类地位的阐明起到重要作用。全基因组序列比较表明, *A. tumefaciens* 和 *S. meliloti* 的基因组存在广泛的直系同源基因 (ortholog) 和核苷酸同线 (nucleotide colinearity) 现象,及假想的复制起始点和终止点周围的重排现象。保守蛋白的系统发育也表明它们祖先相同,且种系分离时间不长,不同的仅是它们的基因组结构和毒力基因成分^[27]。因而二者亲缘关系更近。相反, *S. meliloti* 和 *M. loti* 没有这样大量的保守基因顺序,亲缘关系较远,因而分类地位不同。另外,染色体结构的比较和 RIME 元件的有无在阐明 *Agrobacterium* 作为一个独立的属的分类地位时也起到了重要作用。

根瘤菌的染色体携带大多数管家基因 (housekeeping gene), 共生质粒携带部分必需基因 (essential gene)^[23,28]。根瘤菌株 MAFF303099 目前在分类学上归属于 *M. loti*。但 Turner 等根据染色体基因定义种,以及共生基因定义生物型的理论,发现依据染色体基因 *rrn*, *glnA*, *glnII* 和 *recA* 来推论根瘤菌株 MAFF303099 的系统发育关系时,其与 *M. huakuii* 关系更近,而依据 *nodA* 序列推论其系统发育关系时,则与 *M. loti* 关系更近,因而此株属于 *Mesorhizobium huakuii* biovar *loti* 而不是 *Mesorhizobium loti*^[22]。用此方法也将帮助阐明其它根瘤菌的分类地位。

总之,根瘤菌基因组和共生质粒全序列的分析是根瘤菌比较基因组学的一大进步,也为系统发育分类提供了更多的契机。但是目前全基因组序列分析需要耗费大量资金和时间,因此还不能广泛应用于大量细菌,而只能帮助澄清某些根瘤菌属种的系统发育地位。

4 展望

随着测序技术的发展,细菌全基因组序列分析逐渐变得较为容易,根瘤菌未来的系统发育分类将依据对大量细菌进行序列分析然后进行比较的方法,或是通过能够反映序列变异的芯片方法。相信不断的国际合作能加速这一方法的实现。

致谢 感谢本室硕士研究生艾效曼在文章修改方面的建议。

参考文献

- [1] Woese C R. *Microbiol Rev*, 1987, **51** (2): 221 ~ 71.
- [2] Young J P, Downer H L, Eardly B D. *J Bacteriol*, 1991, **173** (7): 2271 ~ 2277.
- [3] de Lajudie P, Willems A, Pot B, *et al.* *Int J Syst Bacteriol*, 1994, **44**: 715 ~ 733.
- [4] de Lajudie P, Laurent-Fulele E, Willems A, *et al.* *Int J Syst Bacteriol*, 1998, **48** Pt 4: 1277 ~ 1290.
- [5] Young J M, Kuykendall L D, Martinez-Romero E, *et al.* *Int J Syst Evol Microbiol*, 2001, **51** (Pt 1): 89 ~ 103.
- [6] Willems A, Coopman R, Gillis M. *Int J Syst Evol Microbiol*, 2001, **51** (Pt 1): 111 ~ 117.
- [7] Sy A, Giraud E, Jourand P, *et al.* *J Bacteriol*, 2001, **183** (1): 214 ~ 220.
- [8] Moulin L, Munive A, Dreyfus B, *et al.* *Nature*, 2001, **411** (6840): 948 ~ 950.
- [9] Sullivan J T, Eardly B D, van Berkum P, *et al.* *Appl Environ Microbiol*, 1996, **62** (8): 2818 ~ 2825.
- [10] Young J P W, Haukka K E. *New Phytol*, 1996, **133**: 87 ~ 94.
- [11] Eardly B D, Wang F S, van Berkum P. *Plant Soil*, 1996, **186**: 69 ~ 74.
- [12] Peng G X, Tan Z Y, Wang E T, *et al.* *Int J Syst Evol Microbiol*, 2002, **52** (Pt 2): 457 ~ 462.
- [13] van Berkum P, Terefework Z, Paulin L, *et al.* *J Bacteriol*, 2003, **185** (10): 2988 ~ 2998.
- [14] Turner S L, Young J P. *Mol Biol Evol*, 2000, **17** (2): 309 ~ 319.
- [15] Gaunt M W, Turner S L, Rigottier-Gois L, *et al.* *Int J Syst Evol Microbiol*, 2001, **51** (Pt 6): 2037 ~ 2048.
- [16] Laguerre G, Nour S M, Macheret V, *et al.* *Microbiology*, 2001, **147** (Pt 4): 981 ~ 993.
- [17] Young J M. *Int J Syst Evol Microbiol*, 2001, **51**: 945 ~ 953.
- [18] Liu S L, Schryvers A B, Sanderson K E, *et al.* *J Bacteriol*, 1999, **181** (21): 6747 ~ 6755.
- [19] Liu S L, Liu G R, Li S X, *et al.* *J Peking Univ [Med]*. 2002, **34** (5): 457 ~ 463.
- [20] Liu S L, Hessel A, Sanderson K E. *Proc Natl Acad Sci U S A*, 1993, **90** (14): 6874 ~ 6878.
- [21] 郑君芳, 陈文峰, 刘柱荣, 等. *微生物学通报*, 2003, **30** (4): 126 ~ 130.
- [22] Turner S L, Zhang X X, Li F D, *et al.* *Microbiology*, 2002, **148** (Pt 11): 3330 ~ 3331.
- [23] Galibert F, Finan T M, Long S R, *et al.* *Science*, 2001, **293** (5530): 668 ~ 672.
- [24] Kaneko T, Nakamura Y, Sato S, *et al.* *DNA Res*, 2000, **7** (6): 331 ~ 338.
- [25] Kaneko T, Nakamura Y, Sato S, *et al.* *DNA Res*, 2002, **9** (6): 189 ~ 197.
- [26] Gootner B, Hinkle G, Gattung S, *et al.* *Science*, 2001, **294** (5550): 2323 ~ 2328.
- [27] Wood D W, Setubal J C, Kaul R, *et al.* *Science*, 2001, **294** (5550): 2317 ~ 2323.
- [28] Downie J A, Young J P. *Nature*, 2001, **412** (6847): 597 ~ 598.