

放线菌的多相分类

王意敏 刘志恒

(中国科学院微生物研究所 北京 100080)

关键词 放线菌, 系统进化, 多相分类

分类号 Q939.13 文献标识码 C 文章编号 0253-2654(1999)-02-0137-40

随着科学技术的发展,人们对放线菌的认识已由早期的描述分类发展到现代的化学分类 (Chemotaxonomy) 和分子分类 (Molecular taxonomy), 以至于现在的多相分类 (Polyphasic taxonomy)。分子分类的规则来源于细胞中的核酸 (即 RNA 和 DNA) 分析, 化学分类特征则来源于大部分细胞化学特性和一些表达特征, 如蛋白质及其功能。而多相分类是 27 年前提出的观点, 它是将表型、基因型以及系统进化分析数据综合起来作为分类和进化分析的依据。多相分类最早由 Colwell^[1] 提出, 可用于所有水平的分类单位的描述。

当人们进行研究时, 对于某种技术在何种水平含有分类信息, 以及这种技术所花费的时间和劳动量都需要进行了解。一般来说, rRNA 或蛋白质序列分析不适用于对大量分类单位进行分析。化学分类技术如脂肪酸分析是快速检测, 可以在短期内对大量菌株进行比较和分类, 而像 DNA 杂交技术只限于少量代表性菌株。

在现在的文献资料中, 所有的注意力都放在了种、属、科分类水平上。当然, 种在细菌分类中的地位相当重要, 但分类的等级结构要求我们考虑至少更高的属和科的分单位。Erko Stackebrandt 等人 (1997 年) 根据 16S rRNA 和 rDNA 基因编码分析, 在纲 (Classis) 和属之间为放线菌提出了一个新的等级分类系统, 并且不考虑分类单位的表现描述, 将在属水平以上相近的分类单位划分到科、亚目、目、亚纲和纲^[2]。

1 放线菌分类的划分

1.1 形态分类 放线菌广泛分布于自然界。大多数为好气型, 少数为人和动植物致病菌的发酵型。放线菌能在实验室培养基上正常生长, 生活周期较一般细菌要长。通常为革兰氏阳性, 目前只有枝动菌 (*Mycoplana*) 一个属为阴性。其形态学可归为三种类型: 1) 生孢放线

菌 (*Sporoactinomycetes*); 2) 诺卡氏菌型放线菌 (*Nocardioform actinomycetes*); 3) 游动放线菌 (*Actinoplanetes*)。

Kluyver & Van Niel^[3] 提出最早的包括放线菌在内的系统进化树, 按照他们的观点, 形态复杂的微生物如放线菌是由形态简单的球菌进化的。这种形态相似性反映进化关系的观点无疑是缺少证据的, 由于适应环境而造成的不同生物体的形态相似性也是常见的。

1.2 化学分类 根据放线菌的细胞化学组成和特性, 又可以将放线菌归纳为 9 个胞壁类型, 4 个全细胞水解糖型, 5 个磷酸类脂类型, 及多种甲基萘醌类型。这些细胞化学特性是放线菌化学分类的主要指征^[4]。

通过细胞壁组分的化学分析, 人们开始对传统的形态分类规则提出挑战。化学分类特征的阐述 (如胞壁糖、氨基酸、脂肪), 证明了原来形态分类划分的大部分放线菌属和几乎所有的科, 都由化学组成不同的分类单位组成。利用化学特征将放线菌重新划分, 澄清了原来分类单位, 同时不断提出一些新的分类单位, 如类诺卡氏菌属 (*Nocardioides*)^[5] 的建立等。但是这些化学特征本身呈多元化, 用它们并不能构建系统进化图。

1.3 分子分类 随着科学技术的进步, 分子生物学技术在现代分类学中已占主导地位。应用于分类的分子主要为 DNA 和 RNA。rRNA 和 DNA 在分类中扮演不同的角色。如在亲缘关系较近的细菌之间比较相似性, DNA 杂交要比测序更容易。以测序为基础的分类模式与表型数据的综合能对已有的分类地位进一步检验。

经过研究, 人们发现 rRNA 是最合适的分子选择, 它具有保守性, 在细胞中普遍存在, 并且由高度保守区

和可变区组成。利用 16S rRNA 和 RNA 聚合酶,延伸因子 G,ATP 酶等分子所建立的系统进化树相似。另外如细胞色素 C 在一定的范围内也能给出与 rRNA 一致的进化图谱^[6]。综上所述,人们可以从以上任一分子的基因序列分析精确地决定任何生物的系统进化。但是,文献中也有很多例子证明不同的大分子产生不同的系统进化。有人认为主要是因为系统进化树的计算推导理论远远落后于技术的发展。目前关于系统进化分析理论最紧迫的问题就是能使分子发挥最大的信息量,从而使之在一定的分类水平上越来越精确可靠。

2 系统进化研究方法的发展

2.1 DNA 碱基组成比例的测定 (G+C mol%) DNA 的 G+C mol% 测定是典型的基因型指征之一,并且通常被认为属于细菌分类群的标准描述不可或缺的部分。一般认为,G+C mol% 含量在种内不超过 3%,在属内不超过 10%^[7]。

2.2 DNA-DNA 杂交 DNA 杂交百分比和杂交体热稳定性的降低通常用于对种的同源性进行描述。DNA 结合百分比^[8]或 DNA 杂交值或相关结合比例^[9,10]都是两个全基因组之间比较序列相似性的间接参数。现已证明 1% 的错配能产生 1%~2.2% 的热稳定性的降低^[11]。但这些数据是由短寡核苷酸链得到,它们是否可推导出全基因组相似性仍然引起高度争论。

在所有的方中,羟磷灰石法^[9]、光复性速率法^[8]和 S1 核酸酶法^[10]是最普遍的。光复性速率法的优点是 DNA 不需要标记,若 DNA 同源性低于 30% 则在分类学上无较大意义。羟磷灰石法和 S1 核酸酶法两种方法得到的相关结合比例虽然不同,但 ΔT_m 值相似。但是,这些较老的方法需要大量的 DNA 且耗时间较长。少量 DNA 快速杂交方法已经出现^[12],如新的同位素标记或非同位素标记的膜杂交方法,已取代旧的方法。

DNA 杂交的条件是由盐离子浓度、甲酰胺浓度及所用 DNA 的变性温度和 G+C mol% 决定。杂交的最佳温度范围一般较宽(约 5℃),因此复性或杂交需要的温度范围为低于熔点 22~26℃。熔点温度(T_m)可以通过不同的公式从 DNA 碱基比例和盐离子浓度得到。在杂交混合液中每加入 1% 的甲酰胺,熔点温度和杂交温度降低 0.6℃^[13]。

2.3 rRNA 同源性分析 现在通常认为 rRNA 是研究系统进化关系最好的选择。近年来,核糖体的组分一直

是分类研究的主体。新的分子生物学技术的发展使微生物学家把注意力集中在了 rRNA 分子的比较研究上。

2.3.1 16S rRNA 寡核苷酸编目:16S rRNA 编目解释了放线菌中许多分支之间的进化关系,根据 16S rRNA 编目,将革兰氏阳性菌分为两大类,即 G+C mol% 大于 55% 或小于 50%。它的缺点是花费高,技术难度大。

2.3.2 rRNA 测序:1985 年,人们发现了一个新的简单的理论,即直接进行 rRNA 的测序。rRNA 具有较强的二级结构,测序一般有 1%~2% 的误差,虽然不影响较大的分支构造,但对分辨率有影响。

5S rRNA 含有 120 个碱基,由于它的一级与二级结构较为简单,首先建立了它的数据库。但分子的大小并不决定一切,结构也是一个很重要的因素。所有的分子都是由结构单位(如 RNAs 中的发卡结构)构成。大部分的功能单位通常以随机的方式变化,并不改变基本结构。但有时((虽然很少)成串的碱基发生变化能使有些结构单位发生改变(重构现象)。在 5S 和 16S rRNA 中已证明了这一点。如果分子中含有这种结构单位较少,那么分子重构就会大大改变菌株的进化位置,并且使随后的进化分析复杂化。5S rRNA 只有 5 个左右的配对区域,而 16S rRNA 却几乎是它的十倍,因此,重构对 16S rRNA 要比 5S rRNA 小得多。短芽孢杆菌(*Bacillus brevis*)曾由于 5S rRNA 的重构导致进化位置的错误。通常,一个分子区域如果经过了重构,那么进化分析时就应将其省略。

也有人将希望寄托在 23S rRNA 上,认为它可能带有比 rRNA 小亚基更多的遗传信息。在有限的范围内的确如此,但是 23S rRNA 的碱基突变速率要比 16S rRNA 快得多。因此,对于较近的亲缘关系,23S rRNA 还能提供较有价值的信息。在一定条件下,由于 16S rRNA 和 23S rRNA 功能上并不相互独立,用它们构建的进化树比较相似^[6]。

在 rRNA 分子中,保守区主要用于识别同源序列,但对识别进化位置毫无用处。轻度突变区可用于鉴别进化分支末端,高度突变区用于定义亲缘更近的分支。但当这些高度突变区应用于进化分析时,可能会增加随机噪音和系统错误。经验显示,只有对这些可变区正确处理,才能更精确地确定进化位置。

在分析时,除了考虑碱基变化的速率外,还要想到碱基突变的性质。rRNA 的小亚基主要是碱基的替代。

也有一些特定的区域发生碱基的插入和缺失,它们能引起进化位置排列的不确定,为了不影响后面的结果分析,要将其排除掉。

综合上述因素,由于 16S rRNA 的分子大小适中,又有一定的信息量,所以很多研究者采用了 16S rRNA 作为研究对象。Lane 首先用反转录酶和测序引物对 16S rRNA 进行测序,一般能获得 90% 的序列。现在则采用 PCR 技术和合适的引物来测全序列,并建立了数据库,它们提供了系统进化树的框架。

与 rRNA 的测序技术相比,进化位置的计算和推导理论的发展显然落后了很多。现在关于进化地位的理论假说都或多或少难与 16S rRNA 的序列内容相吻合。最常用于放线菌的方法是基于 Jukes 和 Cantor 修正的距离矩阵法。它通过对重叠的替代碱基的修正把序列的相似性转换成进化的距离值。这种经过修正的距离值可输入程序中,进行系统树中分类单位中成对的距离值的处理(最小平方分析)。当序列较相近和没有较大的碱基差异或谱系专一碱基突变速率效应时,这种方法较好^[14]。

2.3.3 DNA 为基础的分类方法: DNA 为基础的分类通常是指在种及种以下水平分类。以前,种间和种以下分类是利用生化或血清免疫反应,抗生素敏感性、噬菌体或细菌素分型等表型分类方法来进行的。这些技术普遍适用,操作简单,重复性好,分辨率较高。尽管下面所列的几种方法优点还不如表型分类,在许多实验室分子分类已代替了表型分类。

全基因限制性片段分析(RFLP)和质粒 DNA 分析是首先被应用的方法。在前者中,要提取全基因 DNA 并进行酶切,然后将这些片段进行琼脂糖凝胶电泳分离和比较,酶切片段的长度呈多样性。它的缺点是产生的 DNA 片段的图谱较复杂,比较起来很困难。

质粒分析的缺点是菌株并不总是含有质粒,且大部分菌株经常只属于几个类群。质粒的限制性酶切片分析把两种技术结合了起来,可以产生更简单的电泳图谱;同时,也建立了相同分子量质粒的鉴定。它减少了用于比较的 DNA 片段,增加了可靠性和分辨率。

低频限制性酶切片分析(LFRFA)选用酶的特点是识别特殊的六到八个碱基,酶切位点较少,因此酶切片较长,利用琼脂糖凝胶电泳不能分离。故 LFRFA 技术的发展只能依靠电泳技术的发展,如脉冲电场凝

胶电泳,被认为是分辨率最好的 DNA 分类方法。

核酸分型方法(ribotyping methods):将复杂的 DNA 酶切片段转移到膜上,和标记的探针(rRNA)进行杂交,比较杂交片段。探针可以在序列和标记技术上有不同,如可用 16S 或 23S rRNA,也可用 rRNA 的保守寡核苷酸片段。另外,延长因子 Tu,核糖体蛋白 S12 和鞭毛蛋白在类似的方法中都可作为探针。

由于 PCR 技术的引入,产生了许多不同的 DNA 分类方法。在这些不同的方法中,一般用短的随机序列作为 PCR 的引物:20 个碱基的寡核苷酸是随机引物 PCR 的引物;10 个碱基的寡核苷酸用于随机扩增多型性 DNA 分析;5 个碱基的寡核苷酸用于 DNA 扩增性指纹图谱。革兰氏阳性菌或阴性菌基因内重复序列的互补片段或 tRNA 的基因片段都可作引物。由于 PCR 反应条件的限制,最新的以 PCR 为基础的方法一般用来对种和种间进行分类。

扩增性 rDNA 限制性酶切片段分析方法(ARDRA):PCR 为基础的分类技术和限制性酶切分析相结合而产生。PCR 产物是 16S 或 23S rDNA 或它们的基因片段。引物则为 rRNA 基因的保守序列。扩增产物利用选好的几种内切酶进行酶切,再进行比较。ARDRA 一般应用于种间水平。

扩增性片段长度多型性分析(AFLP)技术:它的基本原理是先进行酶切,选择特殊的 DNA 片段进行 PCR 扩增。然后分析长度多型性。通常选用两种内切酶,产生两种不同的粘性末端 DNA 片段,短的寡核苷酸链(连接分子)通过末端连接起来形成 PCR 的模板。选择性扩增反应是通过两种不同的引物来执行,与连接分子的序列相同但包括紧邻引物酶切位点的一个或多个选择性碱基。只有与引物完全配对的片段才能扩增。扩增过程经常产生 30~40 个 DNA 片段,其中一些具有种特异性,另外一些是菌株独有的。因此,这种技术可用于种和种以下水平的鉴定和分类。

另外,还有低分子量 RNA 电泳分析法,它的指纹图谱由 5S rRNA 和 tRNA 构成,可用于鉴别真细菌的一些主要类群^[15]。

3 以 16S rRNA 测序为基础的放线菌进化关系

放线菌在细菌进化树中形成一个大的分支,分支的进化深度是由亲缘关系最远的成员之间的 16S

rRNA 的序列差异决定, 主要分支之间的关系仍不能确定, 因为当标准序列的组成和分析方法发生改变时, 它们之间的位置也发生改变。一些形态和化学特征如肽聚糖、甲基萘醌、磷脂、胞壁糖、脂肪酸的类型都对属的特征描述有意义。但在更高级的分类水平, 如在科的水平, 有不同的表型相似性并且保守性较小。除了枝菌酸外, 大部分表型特征都是多元的。将这些特征综合起来很有价值, 对于属的确定较重要。

如果用 rRNA 中可忍受突变的 40% 序列的碱基替代速率对应于地质年代中的主要生态事件作图, 它们的关系呈曲线型。就像所预期的那样, 第一个亿年的碱基替代速率(每 6.2×10^7 年 1% 的突变)要比以后的快(每 $1.2 \sim 1.6 \times 10^8$ 年 1% 的突变)。最近的 0.5 亿年, 碱基替代速率则为每 $5.0 \sim 5.5 \times 10^7$ 年 1% 的突变, 这个数据接近于内共生菌和它们的宿主化石记录比较 rRNA 序列差异所得结果。显然, 碱基替代速率即使有小小的不同也会引起进化年代几百万年的增减。考虑到这些因素, 人们提出下列放线菌进化的主要事件时间尺度: 放线菌起源于大约 2 亿年前(约 20% 的碱基差异); 最早的分支发生在大约 1.9 亿年前(15%~17% 的碱基差异); 最主要的分支发生于不到 1 亿年前(12% 的序列差异); 主要的类群如节杆菌 (*Arthrobacter*) 分化于不到 5 千万年前(它们之间序列差异不到 8%)^[14]。

4 总结与展望

生物的多样性(Biodiversity)表现在它能产生大量不同的分子, 这就是为什么单相分类系统注定不能成功应用的原因。虽然, 以 rRNA 测序为基础的系统进化树是研究进化的一大进步, 但主要分支单位之间的模式仍未确定。rRNA 进化树稳定性有望利用序列的累积来提高, 而一个新的放线菌的描述需要综合它的表现特征和基因特征, 即多相分类。一些新的放线菌谱系被发现, 填补了进化树中的空白, 同时也稳定了已存在的类群之间的关系。然而, 亲缘关系较近的谱系分支位置仍然不确定。rRNA 测序的分辨率不足以分开这些类群。寻找一种不同的分子来确定这些关系是分子分类学家当前面临的任务。适当的表观特征也能使进化树更加稳定。有些研究者认为任何系统进化体系都要呈现表型的一致性。实际上, 它也是多相分类的基础。

目前, 多相分类是纯经验性的, 不具有严格的规则, 可利用任一较重要的信息, 满足大部分但不是全部

的要求。多相分类并不因任何单独的概念偏差受到影响, 信息越多, 结果越能反应实际情况。

大部分细菌在鉴定时, 仍先采用表型鉴定的方法, 但对新的分离菌株, 一种方法是不可靠的, 需要几种信息的综合。现在最直接的方法就是先把它放在进化的框架中, 再用多相分类的方法确定精确的位置。多相分类代替任何单相分类已经成为不可避免的现实。只有对它进行更多的研究, 微生物分类研究才能对生物科技、生物多样性研究、环境研究作出更大的贡献。

参 考 文 献

- [1] Colwell R R. J Bacteriol, 1970, 104:410~433.
- [2] Stackebrandt E, Fred A R, Naomil W R. Int J Syst Bacteriol, 1997, 47:479~491.
- [3] Kluyver A J, Van Niel C B. Zentralbl Bakteriell Parasitenkr Infektionskr Hyg Abt, 1936, 294:396~403.
- [4] Goodfellow M, Minikin D E. Chemical Methods in Bacterial Systematics. London: Academic Press, 1985.
- [5] Prauser H. Int J Syst Bacteriol, 1976, 26:58~65.
- [6] Gary J O, Woese C R. The FASEB J, 1993, 7: 113~123.
- [7] Stackebrandt E, Liesack W. Nucleic acids and Classification. In: Goodfellow M, O'Donnell A G ed. Handbook of New Bacterial Systematics. London: Academic Press Ltd, 1993.
- [8] De Ley J, Cattoir H, Reynaerts A. Eur J Biochem, 1970, 12:133~142.
- [9] Brenner D J, Fanning G R, Rake A V et al. Anal Biochem, 1969, 28:447~459.
- [10] Grimont P A D. Can J Microbiol, 1988, 34:541~546.
- [11] Stackebrandt E, Goebel B M. Int J Syst Bacteriol, 1994, 44:846~849.
- [12] Jahnke K D. J Microbiol Methods, 1994, 20:273~288.
- [13] Johnson L L. DNA Reassociation experiments. In: Stackebrandt E, Goodfellow M ed. Nucleic acid techniques in bacterial systematics. United Kingdom: John Wiley & Sons Ltd. Chichester, 1991.
- [14] Embley T M, Stackebrandt E. Annu Rev Microbiol, 1994, 48:257~89.
- [15] Vandamme P, Pot B, Gills M et al. Microbiol Rev, 1996, 60:407~438.