

# 多位点序列分型及其应用

张少敏(综述), 徐建国(指导)

**摘要:** 多位点序列分型(multilocus sequence typing, MLST)是一种基于核酸序列测定的细菌分型方法。这种方法通过 PCR 扩增多个管家基因内部片段并测定其序列,分析菌株的变异。MLST 操作简单,结果能快速得到并且便于不同实验室的比较,已经用于多种细菌的流行病学监测和进化研究。随着测序速度的加快和成本的降低,以及分析软件的发展,MLST 逐渐成为细菌的常规分型方法。

**关键词:** 管家基因; 多位点序列分型; 应用

中图分类号: R34

文献标识码: A

文章编号: 1003-9961(2008)10-0648-03

**Multilocus sequence typing and its application** ZHANG Shao-min, XU Jian-guo. Institute of Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing 102206, China

**Corresponding author:** ZHANG Shao-min, Email: zhang-shaomin@163.com

**Abstract:** Multilocus sequence typing (MLST) was proposed as a nucleotide sequence-based approach that could be applied to many bacterial pathogens. It analyzed allelic variation at multiple housekeeping loci by nucleotide sequencing of internal fragments. MLST combined developments in high-throughput sequencing and bioinformatics with established population genetics techniques to provide a portable, reproducible, and scalable typing system that reflected the population and evolutionary biology of bacterial pathogens.

**Key words:** housekeeping gene; multilocus sequence typing; apply

多位点序列分型 (multilocus sequence typing, MLST)是一种通过直接测定多个管家基因的核苷酸序列来发现细菌变异的分型方法。它由 Maiden 等<sup>[1]</sup>于 1998 年首次运用于脑膜炎奈瑟菌的分型,后来由于其结果精确并且易于传递而广泛应用于其他致病菌、真菌以及一些非致病菌。MLST 由多位点酶电泳 (multilocus enzyme electrophoresis, MLEE)方法发展而来,后者根据多种代谢相关酶的电泳迁移率的不同来描述它们的变异,MLST 则通过直接测定表达这些酶的管家基因的核苷酸序列来发现它们基因水平上的不同。MLST 方法一般测定 6~10 个管家基因内部 400~600 bp 片段的核苷酸序列,每个位点的序列根据其发现的时间顺序赋予一个等位基因编号,每一株菌的等位基因编号按照指定的顺序排列就是它的等位基因谱,也即是这株菌的 ST(sequence type)。这样得到的每个 ST 均代表一组单独的核苷

酸序列信息<sup>[2]</sup>。因为这种基于核苷酸序列测定的方法可以准确地记录细菌基因水平上的变异,并且核酸序列可以通过互联网方便地传递,所以 MLST 被用于研究细菌的流行病学、群体生物学、致病性和进化<sup>[1]</sup>。

## 1 MLST 方案设计

**1.1 菌株的选择** 为了尽量反映细菌的多态性,即涵盖尽量多的信息,一般根据已有的分型方法和流行病学资料,选择 100 株左右具有代表性的菌株作为分析对象。理论上,这些选定的菌株要代表所研究细菌的群体,而不是仅仅包含那些对人致病的克隆<sup>[3]</sup>。

**1.2 管家基因的确定** 全基因组序列的测定大大方便了 MLST 位点的选择,一般选择 10 到 20 株菌评价 10 到 15 个备选位点<sup>[2]</sup>。MLST 最早应用于脑膜炎奈瑟菌就是从 11 个备选基因中选出 6 个基因<sup>[1]</sup>,后来出于流行病学原因增加一个位点(fumC)。以后的大多数 MLST 方案都采用 7 个位点进行后续实验和分析。如果需要提高分辨率,可以加入那些非保守的基因,例如表面抗原基因<sup>[4]</sup>或者毒力基因<sup>[5]</sup>,这样比增加管家基因的个数将更有效。

如果测定整个管家基因的序列将能得到最完整

作者单位: 传染病预防控制国家重点实验室 中国疾病预防控制中心  
传染病预防控制所, 北京 102206

作者简介: 张少敏, 女, 硕士研究生, 主要从事细菌学研究工作

通讯作者: 张少敏, Tel: 010-61739479, Email: zhang-shaomin@163.com

收稿日期: 2008-03-27

的信息,但是出于实际考虑,MLST 只选择管家基因内部 400~600 bp 长度片段。因为在目前的条件下,全自动测序仪单个反应测得的可信的核苷酸片段是 600 bp, 这样每个位点的等位基因序列就能通过正反向各一个测序反应确定,既加快了速度也节省了开支。这一点在处理疫情暴发时尤为重要。

**1.3 引物的设计** 虽然全基因组序列的获得和分析软件的发展给引物设计带来很大便利,但是,在实际工作中,应该将设计出的所有引物调整为同一退火温度,以适用于大范围流行病学监测和群体研究。另外,为了发现引物结合位点可能存在的变异,每个位点都应该设计一对以上引物<sup>[2]</sup>。

## 2 MLST 数据库

MLST 方法的初衷就是通过互联网免费而且快速传递细菌的遗传信息,用于公共卫生和基础研究。用户可以免费从数据库中下载 PCR 标准操作程序和菌株信息,也可以在数据库中比对序列。用户向 MLST 数据库提交序列之后就能得到它的等位基因编号,再根据所有等位基因编号最终获得细菌的序列型(ST)。如果用户提交的是新的等位基因序列,数据库管理员会在检查序列正确无误后赋予这个序列新的等位基因编号,并把它添加到数据库中。

随着新序列的不断加入,数据库将包含更多的细菌多态性信息。早期的 MLST 数据库因为菌株少,所有的信息包含在一个数据库里,包括等位基因谱、等位基因序列、ST 编号以及菌株基本信息,这样,一方面造成信息的重复。例如,等位基因谱就包括等位基因序列和 ST 编号,另一方面 MLST 的特点要求所有信息都能随时快速从互联网获得,这样势必会增加网站维护的负担。随着 MLST 方案的增加,目前 MLST 站点把等位基因信息和菌株信息分列于两个数据库中,这样更便于用户查询。

## 3 MLST 结果分析

MLST 有两种分析方法,一种是以等位基因编号和 ST 编号为基本分析单位,还有一种就是直接分析核苷酸序列。选择哪种方法取决于所要分析的细菌群体的构成<sup>[6]</sup>,在分析 MLST 数据之前首先要确定所研究细菌群体的克隆程度,然后再根据需要选择相应的分析方法。最简单的方法就是检验种系发生关系,如果每个位点的种系发生关系一致或者近似一致,就可以直接分析核苷酸序列;如果不一致,则要采用分析等位基因编号的方法,如 eBURST 等<sup>[7]</sup>。

MLST 首次应用于脑膜炎奈瑟菌是用聚类方法分析 ST 编号<sup>[1,8]</sup>,但是这种方法完全忽视了细菌群体的实际克隆模式,只是重建了不同菌株之间关系。后来发展的 eBURST 方法克服了这一缺点<sup>[7]</sup>。eBURST 适用于那些主要由基因水平转移导致变异的细菌。因为不论每个等位基因有多少个碱基的改变,它们的权重是相同的。特别是对于那些重组率高的细菌,亲缘关系接近的菌株之间的不同多是由一次重组事件引起的多个碱基的改变,而不是由单碱基突变的积累引起的变异。但是对于那些很少发生重组或者不重组的细菌,它们的变异是多个位点的单碱基突变引起的,这时就需要直接分析核苷酸序列,得到它的种系发生关系树。START (sequence type analysis and recombinational tests) 软件包包含很多 MLST 数据分析软件<sup>[9]</sup>,可以用于数据总结、重组率计算和克隆群的确等。

## 4 MLST 的应用

MLST 方法简单,通过 PCR 扩增和核酸序列测定即可快速得到结果,其结果易于传递和比较,逐渐成为一种常规的分型方法。MLST 通过测定细菌的 4~8 个管家基因的一段核苷酸序列分析它们的进化和群体生物学特征<sup>[10]</sup>。

MLST 方法可以用于传染病的实验室检验。流行性脑脊髓膜炎(流脑)能够引起脑膜炎、败血症等严重的临床症状,还可以引起人群中的暴发,所以,其早期诊断对临床和公共卫生有重要意义。运用 MLST 方法,通过 PCR 直接检测脑脊液等临床标本<sup>[11]</sup>,无需分离出菌株,速度快、结果准确,为流脑的临床诊断和流行病学监测提供依据。

MLST 应用于流行病学上可以揭示细菌毒力与基因型之间的关系。脑膜炎奈瑟菌、肺炎链球菌和金黄色葡萄球菌等致病菌引起侵袭性疾病的克隆只占很小的一部分,大部分只是使人类处于无症状携带状态。大多数侵袭性脑膜炎球菌性疾病都是由脑膜炎奈瑟菌的某些高毒力的克隆所致,这些高毒力的克隆一般是某些特定的 ST,属于特定的 ST 复合群。而那些不引起症状的脑膜炎奈瑟菌则具有多种不同的基因型<sup>[12]</sup>。但是,有研究显示金黄色葡萄球菌的毒力与 MLST 基因型没有相关性<sup>[13]</sup>。应用 MLST 方法同样可以将沙门氏菌、霍乱弧菌、猪链球菌和汉赛巴尔通体的毒力克隆群区别出来<sup>[14]</sup>。

MLST 还可以用来追溯耐药菌株的来源和传播。例如,大部分耐甲氧西林的金黄色葡萄球菌

(MRSA) 是由对甲氧西林敏感的金黄色葡萄球菌 (MSSA) 进化而来<sup>[15]</sup>, 那些对万古霉素敏感性降低的菌株则是来源于这些 MRSA。用 MLST 方法加上一到两个耐药基因的序列分析可以很好地分辨出那些产超广谱内酰胺酶 (ESBL) 的大肠埃希菌<sup>[16]</sup>。

但是, 由于某些特定血清群菌株的管家基因序列较一致或者变异积累太慢, 这个方法并不适合于对这些特定血清群的菌株进行分型<sup>[17]</sup>。对结核分枝杆菌<sup>[18]</sup>和鼠疫耶尔森菌<sup>[19]</sup>的管家基因测序发现二者均具有高度一致的序列, 这可能是因为它们在上属于“年轻”的种, 所以只包含一个克隆系。

随着高通量和低成本核酸序列测定的实现, MLST 已经成为微生物分型的基本方法之一。它通过对进化相对较慢的管家基因序列的分析, 研究大量具有代表性的菌株的流行病学和进化特征, 从而为阻止病原菌的传播提供科学依据。

#### 参 考 文 献

- [1] Maiden MCJ, Bygraves JA, Feil E, et al. Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms[J]. Proc Nation Acad Sci USA, 1998, 95(6):3140-3145.
- [2] Maiden MCJ. Multilocus sequence typing of bacteria [J]. Annu Rev Microbiol, 2006, 60:561-588.
- [3] Urwin R, Maiden MCJ. Multi-locus sequence typing: a tool for global epidemiology[J]. Trends Microbiol, 2003, 11(10):479-487.
- [4] Feavers IM, Gray SJ, Urwin R, et al. Multilocus sequence typing and antigen gene sequencing in the investigation of a Meningococcal disease outbreak [J]. J Clin Microbiol, 1999, 37(12):3883-3887.
- [5] Lacher DW, Steinsland H, Blank TE, et al. Molecular evolution of typical enteropathogenic *Escherichia coli*: Clonal analysis by multilocus sequence typing and virulence gene allelic profiling [J]. J Bacteriol, 2007, 189(2):342-350.
- [6] Perez-Losada M, Browne EB, Madsen A, et al. Population genetics of microbial pathogens estimated from multilocus sequence typing (MLST) data[J]. Infect Genet Evol, 2006, 6 (2):97-112.
- [7] Feil EJ, Li BC, Aanensen DM, et al. eBURST: Inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data [J]. J Bacteriol, 2004, 186 (5):1518-1530.
- [8] Enright MC, Spratt BG. Multilocus sequence typing[J]. Trends Microbiol, 1999, 7(12):482-487.
- [9] Jolley KA, Feil EJ, Chan MS, et al. Sequence type analysis and recombinational tests (START)[J]. Bioinformatics, 2001, 17 (12): 1230-1231.
- [10] Thierry Wirth, Daniel Falush, Ruiting Lan, et al. Sex and virulence in *Escherichia coli*: an evolutionary perspective[J]. Mol Microbiol, 2006, 60(5):1136-1151.
- [11] Birtles A, Hardy K, Gray SJ, et al. Multilocus sequence typing of *Neisseria meningitidis* directly from clinical samples and application of the method to the investigation of meningococcal disease case clusters[J]. J Clin Microbiol, 2005, 43(12): 6007-6014.
- [12] Jolley KA, Kalmusova J, Feil EJ, et al. Carried meningococci in the Czech Republic: A diverse recombining population[J]. J Clin Microbiol, 2000, 38(12):4492-4498.
- [13] Feil EJ, Cooper JE, Grundmann H, et al. How clonal is *Staphylococcus aureus*?[J]. J Bacteriol, 2003, 185(11): 3307-3316.
- [14] Iredell J, Blanckenberg D, Arvand M, et al. Characterization of the natural population of *Bartonella henselae* by multilocus sequence typing[J]. J Clin Microbiol, 2003, 41(11):5071-5079.
- [15] Enright MC, Robinson DA, Randle G, et al. The evolutionary history of methicillin-resistant *Staphylococcus aureus* (MRSA)[J]. Proc Natl Acad Sci USA, 2002, 99(11):7687-7692.
- [16] Lucia L, Nemoy ea. Multilocus sequence typing versus pulsed-field gel electrophoresis for characterization of extended-spectrum beta-lactamase-producing *Escherichia coli* isolates [J]. J Clin Microbiol, 2005,
- [17] Harbottle H, White DG, McDermott PF, et al. Comparison of multilocus sequence typing, pulsed-field gel electrophoresis, and antimicrobial susceptibility typing for characterization of *Salmonella enterica* serotype Newport isolates [J]. J Clin Microbiol, 2006, 44(7):2449-2457.
- [18] Sreevatsan S, Pan X, Stockbauer KE, et al. Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination [J]. Proc Natl Acad Sci USA, 1997, 94(18):9869-9874.
- [19] Achtman M, Zurth K, Morelli C, et al. *Yersinia pestis*, the cause of plague, is a recently emerged clone of *Yersinia pseudotuberculosis* [J]. Proc Natl Acad Sci USA, 1999, 96(24): 14043-14048.