

基质辅助激光解吸/电离飞行时间质谱 在微生物基因组研究中的应用

李汉东¹ 李飞飞² 张大可² 邓立彬³

(¹ 济宁医学院基础医学与法医学院, 山东 济宁 272067; ² 中国科学院北京基因组研究所, 北京 100029;

³ 南昌大学基础医学院, 江西 南昌 330006)

摘要 质谱已成为分子生物和生物化学研究中一个重要的分析工具, 它的高通量、分析速度、准确性和灵敏度都是传统分析技术所不可比拟的。以基质辅助激光解吸/电离飞行时间质谱为代表的现代生物质谱技术是基因组研究的新兴技术, 可用于核酸高通量的数量和质量分析。在微生物基因组研究中主要应用于微生物基因型鉴定、病毒准种分析, 以及微生物基因组的进化研究、甲基化定量分析和突变检测。本文主要介绍基质辅助激光解吸/电离飞行时间质谱技术的原理及在微生物基因组研究的应用进展。

关键词 基质辅助激光解吸/电离飞行时间质谱; 微生物基因型鉴定; 病毒准种分析

中图分类号: Q657.6; Q812 **文献标志码:** A **文章编号:** 1000-9760(2011)02-063-05

随着人类基因组计划以及多种真核生物基因组测序的完成, 遗传因素在疾病发生中的作用得到深入探讨。人类基因组图谱的完成使得鉴定与各类疾病易感性相关的基因组变异如单核苷酸多态 (SNP) 以及拷贝数多态等成为可能^[1]。此类研究通常基于上千份样本进行分析, 极大促进了高通量测序以及基因分型技术的发展。其中质谱技术以其精确度高、分析速度快及易实现高通量自动化检测等特点, 成为核酸分析的重要技术手段, 并逐渐应用于微生物基因组研究。

1 质谱技术原理和应用

质谱 (MS) 作为一种分析手段已出现几十年, 但直到 1988 年基质辅助激光解吸电离 (MALDI)、电喷雾电离 (ESI) 等软电离技术的出现才为分析强极性、热不稳定性和难挥发性的生物样品提供了可能, 生物质谱从而发展起来, 由于技术分辨率的问题, MALDI 更适合于碱基数较少的短链核酸的分析^[2]。基质辅助激光解吸/电离飞行时间质谱基本原理为将制备的样本分析物 (核酸) 与芯片上一特定的基质分子共结晶, 将芯片放入质谱仪的真空管, 用瞬时纳秒强光激发, 基质分子吸收辐射能量并迅速产热, 使基质晶体升华, 核酸分子转变为亚稳定离子, 这些电荷离子在加速电场中获得相同动能, 飞过自由漂移区到达检测器, 到达检测器的时

间与其分子量成反比。质谱仪根据离子到达检测器的时间推算样本分析物的分子量大小并绘制成峰图, 它的高灵敏性可以将只相差一个碱基的核酸序列区分开来, 我们可以利用峰图提供的信息进行各种核酸分析。目前, 质谱技术在 DNA 分析领域的应用主要包括 SNP 位点的发现及验证、单倍体生物的基因型分析、单体型分析、基因表达量分析以及甲基化分析等。

这种新兴技术在微生物基因组研究中有多个方面的应用, 例如单倍体微生物的基因型分型, 病毒准种中不同基因型病毒的数量分析等, 这些应用有利于寻找更灵敏特异的病原微生物诊断分型手段以及进行快速的疫苗安全性检测, 对临床治疗有重要意义。

2 基于质谱技术的基因组重测序

基于质谱技术的比较序列分析法能以极高的准确度快速确定目标序列含有的 SNP 或其他突变。新突变位点的发现基于位点特异性切割产生的一系列短的核酸片段分子量的不同 (图 1)。对含有目标序列的区段 PCR 扩增, 200-800bp 的 PCR 产物体系用虾碱性磷酸酶 (shrimp alkaline phosphatase, SAP) 处理以除去未反应的 ddNTP 和引物, 然后利用 T7 启动子 (或 T7 和 SP6) 介导体外转录, 并对产生的正负链 RNA 分别进行 T、C

位点特异性切割,相当于对正链进行四种碱基的特异性切割,四种切割产物放入 384 孔板的一个孔中用质谱分析,质谱仪综合四种产物的分析图谱产生目标序列的质谱图。SNP 位点会影响位点特异性切割,使产物的分子量发生变化,对比待测序列和参考序列切割后的产物质谱图,其峰值的改变代表新 SNP 位点的存在,这是基于质谱技术重测序的基本原理。单碱基突变理论上可产生 10 处质谱图峰值的变化,插入或缺失同样会造成峰图不同的变化^[3]。

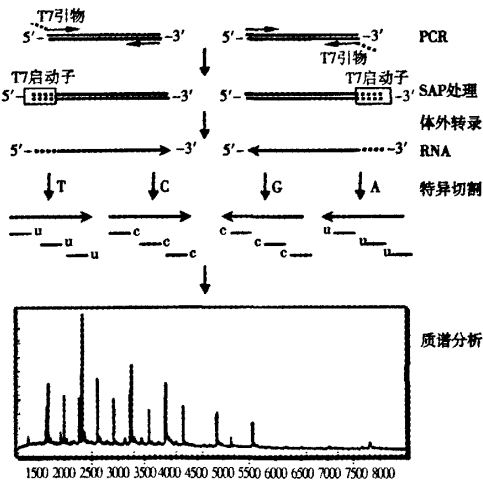


图 1 基于质谱技术的比较序列分析法

PCR 产物经 T7 启动子介导体外转录,转录的正负链 RNA 分别进行 T、C 特异性切割,切割产物进行质谱分析并通过与参考序列质谱图的比对发现新的突变。

另外一种基于质谱技术的重测序方法称为固相捕获双脱氧核苷酸法。这种方法主要特点是利用生物素标记的 ddNTP 在 DNA 片段的任意位置终止 PCR 扩增,产生一系列长度连续分布的片段并进行质谱分析,从而得到 DNA 片段的碱基组成^[4]。由于此种方法在碱基读长上有诸多限制,目前应用较少。

病原微生物基因组的变异是微生物研究的重要内容。突变一方面可能导致抗原决定簇发生变化,使得宿主体内之前产生的抗体失效,从而造成传染性疾病广泛流行;另一方面可以导致临床治疗中病原体产生耐药性。针对微生物基因组进行全基因组或者目标区域的测序是突变分析中的金标准(Golden Standard)。与 Sanger 测序相比,质谱

检测虽然不能直接给出序列信息,但通过与参考序列的比对更容易发现序列变异。此技术直接检测目标序列的分子量,无须任何化学标记;适合大量样本的高通量分析,可同时分析 384 个样本;分析速度快,384 个样品的检测可在 8h 内完成。经验证比较序列分析法可检测到目标序列 98% 的 SNP 位点,并能给出 SNP 的准确位置^[3]。

质谱分析的精确性使其成为突变检测的重要工具,可用于微生物进化研究中突变位点的监测。Christiane Honisch 等人用比较序列分析法测定大肠杆菌 k-12 MG1655 在甘油作为唯一碳源时产生的适应性突变^[5]。在 40d 的实验室模拟进化中,大肠杆菌 k-12 MG1655 经过 700 代的进化达到最快生长状态,利用 MS 重测了这一菌株包括甘油代谢相关基因在内的 4.4% 的序列,确定了 13 个多态位点,主要位于甘油激酶基因中,功能试验进一步证明这些位点在适应性进化中有重要作用。所有突变位点都通过了测序验证,说明这种方法具有很高的准确度。这项研究利用质谱技术确定了大肠杆菌适应甘油基本培养基的遗传基础并观察到突变的动态变化,为微生物进化机制的研究提供了一个很好的方法。

3 微生物种属和基因型的鉴定

细菌和病毒不同种属类型、基因亚型在致病性以及抗药性上有明显差异,对人和动物病原微生物进行基因型鉴定是传染病诊断和防治的重要步骤,有助于掌握其起源、分布及传播规律,选择有效的疫苗和防治措施。传统的病毒细菌分型方法如血清学试验、抗原抗体反应虽然是可靠的,但方法的建立耗时较长,如需要分离鉴定病原并制备成抗原、制备高滴度的诊断血清或单克隆抗体。近年来分子生物学方法也用于鉴定不同基因型,如 DNA 芯片、核酸探针杂交、基因型特异引物 PCR 等,但 these 方法尚存在准确度不高的问题^[6]。

3.1 比较序列分析法

上述比较序列分析法也为病原微生物的分型提供了一种新的途径。从数据库中下载不同基因型的参考序列后,可以利用软件生成相应参考序列的模拟峰图,然后将待测样本的质谱图与模拟峰图进行比对,从而判断出样本微生物的基因型及其种

属类型。这种方法不仅可以自动快速分型,还可报告新的 SNP 位点。

Michael Lefmann 等人利用 16S rDNA 区段 Rnase T1 切割(G 特异切割)产物质谱分析的方法分析分枝杆菌属,研究发现临床常见的 12 种分枝杆菌都有特异性的酶切图谱可以相互区分,通过 24 例样本对此方法进行验证,结果显示质谱分型方法准确度可达 96%,从而为分枝杆菌的种属鉴定提供了一种快速、简便的方法^[7]。Christiane Honisch 等人用质谱技术的比较序列分析法对 100 例脑膜炎双球菌样本进行分型,与测序结果有 98% 的一致性^[6]。另外,质谱技术也已用于奈瑟氏菌属不同种的区分、猿疱疹病毒的鉴定等^[8]。

3.2 单核苷酸引物延伸法

另一种基于质谱技术的微生物分型方法是单核苷酸引物延伸法(图 2)。基本原理是针对基因型特异性位点设计引物,单碱基延伸后的产物进行质谱检测确定该位点的碱基。由于延伸产物比引物多一个碱基,而且掺入不同碱基分子量会有所差异,可根据质谱峰图判断基因型特异性位点的碱基。多重 PCR/引物延伸反应还能够实现针对多个 SNPs 的检测,可以提高通量、降低成本,目前最高可达 40 重,即同时分析检测 40 个位点,每组 PCR 反应只需 5~10ng DNA。反应体系为非杂交依赖性,不存在潜在的杂交错配干扰,不需要各种标记物,准确度高,而且每个检测点只需 3~5s。Ju Luan 等用 HBV 病毒常见的 60 个突变位点对多重引物延伸进行了检验,结果这种方法可检测到 53 个(88.3%)突变,检测到的微小变异是传统测序的两倍,说明引物延伸法有很高的准确性和灵敏度^[9]。

单核苷酸引物延伸法在微生物基因型分析中已有很多应用。人类乳头瘤状病毒是引起子宫颈癌的主要原因,它主要有 14 种基因型,其基因型与疾病的发生、预防和疫苗的使用密切相关。Anna So derlund-Strand 等人尝试用单碱基延伸结合质谱技术对人类乳头瘤状病毒(HPV)进行分型,他们根据 14 种基因型特有的 SNP 分别设计延伸引物,并分析了 532 例子宫颈样本,检测出了传统方法检测不到的 5 个 HPV 阳性样本,并且确定了所有阳性样本中病毒的基因型,检测方法表现出很高

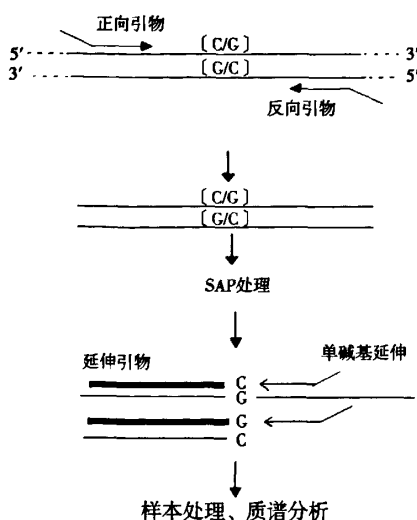


图 2 单核苷酸引物延伸法

扩增含有 SNP 位点的 DNA 片段并经 SAP 处理,在 SNP 的上游或下游设计一条延伸引物,使引物的 3 端紧邻 SNP 位点,由于与 SNP 互补的 dd NTP,引物的 3 端掺入一个碱基,产生等位基因特异性引物产物。经质谱分析根据分子量确定 SNP 位点的碱基组成。

的灵敏性。丙肝病毒(HCV)不同的基因型对抗病毒治疗的反应性不同,治疗前病毒分型检测可以指导临床治疗。Elena N. Ilina 等人挑选 HCV 基因组 5,非翻译区的 2 个基因型特异 SNPs 并进行单碱基延伸反应,成功区分了亚洲流行的 5 种 HCV 病毒,可用于快速判断 HCV 阳性样本中的病毒类型。Malin I. L. Sjolholm 等人针对人类疱疹病毒的不同基因型设计了多重特异性引物,可以一次性检测样品中包含的多种病毒类型,共分析了 882 例病毒样本,结果准确度达到 95.6%。Maxim V. Afanas'ev 等人用此种技术分析已报道的结核杆菌候选抗药突变,确定了 13 个与利福平抗药性相关的 SNPs。此外,疟原虫药物敏感株的鉴定也是通过二氢叶酸还原酶基因 4 个 SNPs 的单碱基延伸反应实现的,沙门氏菌 9 个不同的血清型也可通过单碱基延伸反应区分^[10-11]。

3.3 在病毒准种分析中的应用

病毒 RNA 复制酶或者反转录酶的低保真性决定了病毒基因组的高突变,共存在于宿主体内的病毒不具有相同的基因组,这些基因序列存在异质性的同种病毒群体即为准种。准种是病毒得以适

应宿主体内环境变化的重要机制,掌握准种的变化是研究病毒进化、病毒-宿主互作的重要手段。准种分析的传统方法为克隆测序。在 MALDI-TOF MS 分析中,样本量的多少和峰图中相应峰的面积成正比,在群体遗传学中,可以将不同个体的 DNA 样本混合(pooling)在一起进行检测,从而估计某些位点不同碱基的出现频率。这种策略同样可以应用于病毒准种的分析^[12]。Chunming Ding 等人通过 MALDI-TOF MS 分析 21 个患者体内 HBV 基因组 YMDD 区突变的出现及其与野生型的相对比例,从而监测长期服用拉米夫定的慢性乙肝患者体内病毒准种 YMDD 区抗药性突变的动态变化,MS 有很高的准确性和灵敏度,可检测到含量仅为 5% 的突变株^[13]。

此外,在病毒减毒活疫苗中,少量的突变或回复突变就会影响疫苗的安全性,腮腺炎病毒减毒活疫苗由两种基因型不同的减毒株按确定比例构成。Georgios Amexis 等人检测腮腺炎病毒减毒活疫苗的安全性,他们用质谱分析代替了常规的 PCR 和限制性酶切分析,只用了基因组中五个不同的 SNP 位点便将两者区分开,含量比例与常规检测一致,这表明 MALDI-TOF MS 可用来进行疫苗安全检测^[14]。

4 微生物基因组的甲基化分析

甲基化分析的基本方法是亚硫酸氢盐处理待测 DNA,可使 DNA 未甲基化的 C 转变为 U,甲基化的 C 则不会改变,对比处理前后序列的改变,可以判断位点的甲基化状态。亚硫酸氢盐处理待测 DNA,利用 T7 启动子的引物对含有甲基化位点的序列进行扩增并体外转录,转录产物用碱基特异性酶(一般为 U 特异酶)进行切割,切割产物用质谱分析。未甲基化的 C 变为 T 后,分子量增大 16Dalton,通过质谱仪分析每个片段在处理前后分子量的差异,从而可以发现甲基化位点以及分析甲基化程度(图 3)。Marcel W. Coolen 等人用已知甲基化程度不同的结肠癌细胞和癌旁正常肠黏膜细胞对质谱技术检测甲基化进行了验证和改进。他们的研究表明质谱技术可以灵敏精确地对甲基化进行分析,检测出 5% 甲基化程度的 CpG 位点^[15]。

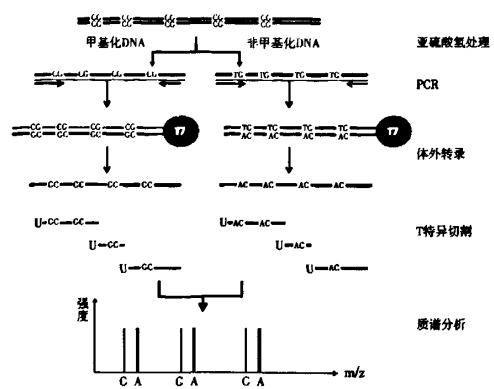


图 3 甲基化分析流程图

亚硫酸氢盐处理 DNA,用含有 T7 启动子的引物对进行扩增、体外转录,产物经 U 特异酶切割并用质谱分析。甲基化 C 变为 T 分子量增大 16Dalton,根据质谱图可确定甲基化程度。

MALDI-TOF MS 可用于研究细菌 DNA 的限制修饰机制。细菌利用限制修饰系统来保护它们不受外来 DNA 的入侵,甲基化酶将内源 DNA 特定序列上腺嘌呤或胞嘧啶甲基化,外源 DNA 侵入细胞时,细菌的限制酶能将其降解,内源 DNA 由于甲基化而被保护。5-羟基胞嘧啶和 5,6-二氢胸腺嘧啶修饰的 DNA 片段的质谱分析阐明了大肠杆菌核酸限制性内切酶 III 的作用机制,主要是通过水解作用断开 DNA 骨架^[16]。

原核生物的 16S-23S-5S 3 种 rRNA 位于同一条转录的 rRNA 分子中,转录后由核酸酶 III 切割为成熟 rRNA,也会产生几种 tRNA 分子。MALDI-TOF 也用于细菌 rRNA 修饰研究,例如嗜热菌 16S rRNA 的所有修饰位点的鉴定、新发现的大肠杆菌甲基化酶 YebU 蛋白对 16S rRNA 修饰位点的研究等^[17]。甲基化可能对加工起引导作用,前体 rRNA 中被甲基化的部位在加工过程中并未被切除,而是一直保持到成熟的 rRNA,另外,如果人为地阻断前体 rRNA 的甲基化,前体 rRNA 的成熟加工也被阻断,因此推测前体 rRNA 的甲基化对 rRNA 的加工具有指导作用^[17]。

病毒的甲基化研究对于疾病病情进展和病毒调节机制有重要意义。西班牙科学家 Agustin F. Fernandez 对三种病毒所导致的癌症的研究表明癌症症状越严重病毒的甲基化程度越深。研究

人员推测甲基化可沉默部分病毒基因使得病毒蛋白表达下降,从而逃避宿主免疫攻击^[18]。HBV CpG 岛的甲基化可抑制病毒 mRNA 的产生及蛋白的表达^[19]。目前质谱技术在病毒甲基化研究中的应用较少,随着人们对病毒甲基化研究的重视和质谱技术的成熟,质谱很可能成为病毒甲基化研究的重要工具。

综上所述,基质辅助激光解吸/电离飞行时间质谱与多种分子技术的结合能对核酸进行多方面分析,由于其快速、精确、高通量等优点,已成为后基因组时代一种主要的研究手段。在微生物基因组研究中的应用目前主要集中于各种病原微生物的分型、病毒准种中各种序列类型的定量分析和适应性突变的发现,以及甲基化程度对病毒致病性的影响等。微生物尤其是病原微生物基因组的研究对于疾病的控制及发现新的治疗方法有重要意义,随着人们对微生物基因组研究的深入,质谱技术无疑在微生物的基因组研究中起到重要作用。

参考文献:

- [1] Ragoussis J, G P Elvidge, K Kaur, et al. Matrix-assisted laser desorption/ionisation, time-of-flight mass spectrometry in genomics research[J]. PLoS Genet, 2006, 2(7):100.
- [2] 彭好益,李强,质谱技术在生物医学中的应用探析[J]. 技术与市场, 2009, 16(2):2.
- [3] Stanssens P, M Zabeau, G Meersseman, et al. High-throughput MALDI-TOF discovery of genomic sequence polymorphisms[J]. Genome Res, 2004, 14(1):126-33.
- [4] Edwards J R, H Ruparel and J Ju. Mass-spectrometry DNA sequencing[J]. Mutat Res, 2005, 573(1-2):3-12.
- [5] Honisch C, A Raghunathan, C R Cantor, et al. High-throughput mutation detection underlying adaptive evolution of Escherichia coli-K12[J]. Genome Res, 2004, 14(12):2495-502.
- [6] Honisch C, Y Chen, C Mortimer, et al. Automated comparative sequence analysis by base-specific cleavage and mass spectrometry for nucleic acid-based microbial typing[J]. Proc Natl Acad Sci U S A, 2007, 104(25):10649-54.
- [7] Lefmann M, C Honisch, S Bocker, et al. Novel mass spectrometry-based tool for genotypic identification of mycobacteria[J]. J Clin Microbiol, 2004, 42(1):339-46.
- [8] Ilina, E N, A D Borovskaya, M M Malakhova, et al. Direct bacterial profiling by matrix-assisted laser desorption-ioniza-

- tion time-of-flight mass spectrometry for identification of pathogenic Neisseria[J]. J Mol Diagn, 2009, 11(1):75-86.
- [9] Luan J, J Yuan, X Li, et al. Multiplex detection of 60 hepatitis B virus variants by maldi-tof mass spectrometry[J]. Clin Chem, 2009, 55(8):1503-9.
- [10] Roumagnac P, F X Weill, C Dolecek, et al. Evolutionary history of Salmonella typhi[J]. Science, 2006, 314(5803):1301-4.
- [11] Baker S, K Holt, E van de Vosse, et al. High-throughput genotyping of Salmonella enterica serovar Typhi allowing geographical assignment of haplotypes and pathotypes within an urban District of Jakarta Indonesia[J]. J Clin Microbiol, 2008, 46(5):1741-6.
- [12] Vignuzzi M, J K Stone, J J Arnold, et al. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population[J]. Nature, 2006, 439(7074):344-8.
- [13] Ding C, V W Wong, K C Chow, et al. Quantitative subtyping of hepatitis B virus reveals complex dynamics of YMDD motif mutants development during long-term lamivudine therapy[J]. Antivir Ther, 2006, 11(8):1041-9.
- [14] Amexis G, P Oeth, K Abel, et al. Quantitative mutant analysis of viral quasispecies by chip-based matrix-assisted laser desorption/ionization time-of-flight mass spectrometry[J]. Proc Natl Acad Sci U S A, 2001, 98(21):12097-102.
- [15] Coolen M W, A L Statham, M Gardiner-Garden, et al. Genomic profiling of CpG methylation and allelic specificity using quantitative high-throughput mass spectrometry: critical evaluation and improvements[J]. Nucleic Acids Res, 2007, 35(18):119.
- [16] D'Ham C, A Romieu, M Jaquinod, et al. Excision of 5, 6-dihydroxy-5, 6-dihydrothymine, 5, 6-dihydrothymine and 5-hydroxycytosine from defined sequence oligonucleotides by Escherichia coli endonuclease III and Fpg proteins: kinetic and mechanistic aspects[J]. Biochemistry, 1999, 38(11):3335-44.
- [17] Emmerechts G, L Maes, P Herdewijn, et al. Characterization of the posttranscriptional modifications in Legionella pneumophila small-subunit ribosomal RNA[J]. Chem Biodivers, 2008, 5(12):2640-53.
- [18] Fernandez A F, C Rosales, P Lopez-Nieva, et al. The dynamic DNA methylomes of double-stranded DNA viruses associated with human cancer[J]. Genome Res, 2009, 19(3):438-51.
- [19] Vivekanandan P, D Thomas and M Torbenson. Methylation regulates hepatitis B viral protein expression[J]. J Infect Dis, 2009, 199(9):1286-91.

(收稿日期 2011-01-11)