

单基因遗传病分子诊断的新策略和新方法

王蕾 贺静 朱宝生*

(云南省第一人民医院遗传诊断中心、云南省出生缺陷与遗传病重点实验室,云南昆明 650032)

【摘要】 单基因遗传病具有复杂的表型异质性及遗传异质性,基因突变可涉及点突变、片段缺失/重复、整个基因的缺失/重复、以及动态突变等众多类别。分子诊断技术使得单基因遗传病可在分子水平乃至单个碱基发生变异的情况下做出准确诊断,然而其可靠性问题成为备受瞩目的技术难点。针对不同遗传病使用不同的解决方案,是目前单基因遗传病分子诊断的主流思想。基因诊断相关人员不仅要了解遗传疾病致病基因及其突变特点,还要熟知各种检测技术的特点,合理选择。

【关键词】 单基因遗传病;分子诊断;杂交;扩增;测序

【中图分类号】 R714.55 **【文献标识码】** A

单基因遗传病(monogenic disorders)是由位于同源染色体上的一对等位基因之一或者二者发生突变,导致蛋白结构信息错误或基因表达控制异常而出现的身体结构发育异常和/或生理功能异常,该类疾病一般符合孟德尔遗传方式,所以,又称为孟德尔遗传病(mendelian inherited diseases)^[1]。大多数单基因遗传病是罕见病,其中每一种在全球人口中的平均发生率在1/10 000~1/100 000^[2],多则在数千个新生儿中出现一例,少则数十万新生儿中出现一例;但对局部地区或者来自某些地区的特定人群而言,某些单基因遗传病的平均发生率可能高达1%~2%,如在东南亚和云南南部的某些少数民族中,地中海贫血和葡萄糖-6-磷酸脱氢酶缺乏症的发生率就高达2%以上^[3]。由于单基因遗传病种类众多,故即使每种疾病的患者不多,但总体受累人群却不容小觑。据可靠估算,单基因遗传病在全球人口中的发生率约为1%^[4],给千万个家庭及患者带来极大的精神负担和经济负担。

1976年,美籍华裔科学家简悦威成功应用液相DNA分子杂交技术实现了镰形细胞贫血症的基因诊断^[5],是人类遗传病诊断开始进入分子诊断时代

的里程碑。继形态学、生化学及免疫学诊断技术后,分子诊断飞速发展,该技术使得遗传病可在分子水平乃至单个碱基发生变异的情况下做出准确诊断。不仅可为患者提供可靠诊断,还可为表型正常的携带者提供遗传咨询。大致可分为分子杂交检测技术、基因扩增检测技术和基因测序技术三大类。

1 分子杂交检测技术

分子杂交检测技术是通过已知基因序列的探针针对靶序列进行特异性的捕获检测。主要技术有斑点杂交(dot blot, DB)、反向斑点杂交(reverse dot blot, RDB)、Southern印迹、Northern印迹、Western印迹、荧光原位杂交(fluorescence in situ hybridization, FISH)、基因芯片(gene chips)等。

1.1 反向斑点杂交(RDB) Saiki等^[6]于1989年最早发明RDB技术,并证实该技术在突变检测中的确定性。1999年,Chan等^[7]将该技术应用于中国人群 α^{CS} 、 α^{QS} 、 α^{CD59} 和 α^{CD30} 4种非缺失型 α -地中海贫血的突变检测。RDB法为目前临床常用的常规检测方法,原理是将一系列可与疾病相关等位基因杂交的特异性寡核苷酸探针排列在尼龙膜上,经扩增、杂交、洗脱及显色,最后判读受检样品是否携带与膜上探针杂交的基因突变。其优点在于可在一张杂交膜上同时检测样品中多种点突变,结果可靠、可重复性好、灵敏度高,且检测时所耗样品量少、经

doi: 10.13470/j.cnki.cjpd.2016.03.002

基金项目:国家自然科学基金(81260415)、云南省卫生领军人才项目(L-201201)

*通信作者:朱宝生, E-mail: bszhu@aliyun.com

济方便,尤为适用于基因分型、已知基因序列的突变检测。缺点是膜条制备时间长,杂交操作繁杂、很难进行大批量检测。

1.2 荧光原位杂交(FISH) 1977年,Rudkin^[8]发明了FISH技术,在原位杂交技术的基础上使用荧光素取代同位素标记探针,更安全、快速、准确、灵敏,且探针标记更稳定,能同时显示多色,不但能对中期分裂相进行分析,还能显示于间期核中。其基本原理遵循碱基互补原则,将荧光标记的已知探针与受检样品DNA进行杂交,通过检测荧光信号实现受检样品DNA的定性、定位分析。FISH技术自20世纪70年代问世以来已在临床广泛应用,由于其直观性、特异性已成为染色体微缺失/微重复诊断的金标准,在分子领域主要适用于基因定位、肿瘤基因检测等方面^[9]。

1.3 微阵列基因芯片 微阵列基因芯片,简称基因芯片,是由千万个DNA或寡核苷酸探针密集排列在小面积载体表面组成的微点阵列。在一定条件下,载体上的DNA或寡核苷酸探针可与来自样品中序列互补的核酸片段杂交,通过检测杂交信号实现样品DNA序列分析^[10]。该技术具有高通量、高分辨率、微型化、自动化程度高等优势,可用于基因定位、已知突变检测、DNA测序、遗传图谱构建等^[11]。但存在着芯片制作技术复杂、检测应用的芯片扫描仪成本高、检测数据分析难度较高的问题,阻碍了其应用于基层医院临床检测中。国产的遗传性耳聋基因检测芯片目前在临床上应用较多,它可以同时检测GJB2、GJB3、SLC26A4和线粒体12S rRNA这4个基因中的9个基因突变热点,是世界上第一个被批准用于临床诊断的基因芯片产品。

2 基因扩增检测技术

临床应用最广泛的单基因遗传疾病分子诊断技术是聚合酶链反应(polymerase chain reaction, PCR)以及其衍生方法。如巢式PCR(nested PCR)、多重PCR(multiplex PCR)、三引物PCR(tri-primer PCR, TP-PCR)、逆转录PCR(reverse transcription PCR, RT-PCR)、实时荧光定量PCR(real-time fluorescence quantitative PCR, RTFQ-

PCR/qPCR)、数字PCR(digital PCR, dPCR)^[12]、变性高效液相色谱分析(denaturing high-performance liquid chromatograph, DHPLC)^[13]、高分辨溶解曲线分析(high-resolution melting curve analysis, HRM)^[14]、多重连接依赖探针扩增(multiplex ligation-dependent probe amplification, MLPA)^[15]等。现今几乎所有的基因突变检测技术都是基于PCR发展的。

2.1 数字PCR dPCR是20世纪初发展起来的新型核酸分子定量检测技术,较传统定量PCR技术具有更高的准确性、灵敏性及可重复性,可对大量背景分子中的微量模板进行定量分析^[12,16]。按分液技术不同分为三类:微孔板^[17]、微流体芯片^[18]和微滴式。由于dPCR的本质是通过单分子扩增将低丰度的基因信号从复杂背景中提取出来,故在无创产前诊断中可对受到大量母体背景DNA影响的低含量胎儿游离DNA进行检测。2010年,Chu等^[19]首次将数字化PCR应用于无创单基因疾病诊断,进行单分子等位基因检测并通过等位基因比值法推断胎儿是否受累。Lam等^[20]于2012年依据父源性的遗传单倍型设计了针对珠蛋白及其外部SNP位点的探针来筛选可提供信息位点,并根据母源性单倍型将SNP分为 α 和 β 两组,借助dPCR技术推断胎儿基因型,成功分析2个 β -地中海贫血家系。Tsui等^[21]应用dPCR检测X染色体上的基因位点,在12个样本中成功检测出7例发生血友病基因突变的男胎。

2.2 变性高效液相色谱分析 DHPLC技术是通过分析异质性双链结构从而对DNA序列变异进行筛选。在不变性的温度条件下可用于分离不同分子量的双链DNA、分析具有长度多态性的片段;在部分变性的条件下,突变型和野生型分别形成同源双链,同时错配产生异源双链,由于杂合与纯合二倍体在柱中保留时间的差异而在色谱图中呈现双峰或多峰的洗脱曲线,以识别突变型^[22]。该技术可分析SNP、检测单个碱基置换、插入或缺失,具有自动化程度高、敏感性和特异性较高等优点。

2.3 高分辨溶解曲线分析 HRM技术是利用突变型片段与野生型片段加热变性时溶解曲线形状和

位置的差异,从而区分突变样本与正常样本,不受突变碱基位点及类型的限制,既可扫描扩增片段的未知突变又可进行已知突变的基因分型^[23]。该技术操作便捷、通量高、成本低,且特异性优于其他突变初筛技术,如单链构象多态性分析(SSCP)技术,适用于单碱基突变及小片段插入、缺失的检测,但无法准确检测具体突变^[24]。

2.4 多重连接依赖探针扩增 MLPA技术是荷兰学者 Schouten 等^[15]于2002年在多重扩增探针杂交(multiplex amplifiable probe hybridization, MAPH)技术^[25]的基础上改进设计的,其原理是对可与样本DNA正确杂交并被连接酶连接的探针进行扩增和半定量分析。该技术融合了DNA探针杂交和PCR技术,且引入了连接依赖,其优越性是特异性高、精确度高、重复性强、操作简便、通量高。基于MLPA技术可检测出若干人类基因拷贝数的变异,故现已广泛应用于存在基因缺失/重复突变的人类遗传疾病的分子诊断,如假肥大型肌营养不良症(Duchenne/Becker muscular dystrophy, DMD/BMD)、脊肌萎缩症(spinal muscular atrophy, SMA)、DiGeorge综合征(DiGeorge syndrome)等^[26]。但遗憾的是该产品至今未申请中国的市场准入,仅能少量用于科研工作,而不能应用于基因诊断和产前诊断。

3 基因测序技术

基因测序技术作为现代生命科学的核心技术之一,不仅为遗传信息的揭示和基因表达调控的基础生物学研究提供重要数据,而且在疾病基因检测与分子诊断等临床应用研究中发挥着重要作用。

3.1 Sanger 测序法 1977年,英国生物化学家 Sanger 等^[27]发明了双脱氧链末端终止法(dideoxy chain-termination method),可以检测物种或细胞的核酸序列,通过与基因库进行比对分析被检测物种或细胞的特性。Sanger 测序法作为最经典的测序方法广泛应用于基因组DNA、cDNA等多重复序列的检测。但该技术通量低、成本高、耗时长。1998年 Ronaghi^[28]发明了焦磷酸测序法,其基本原理是利用DNA聚合酶、ATP 硫酸化酶、荧光素酶、三磷

酸腺苷双磷酸酶的协同作用,检测引物延伸过程中所释放的荧光,通过峰值高低判断与其匹配的碱基数量。该方法在通量、成本、耗时等方面优于 Sanger 法,广泛运用于 SNP 位点、等位基因突变检测等。但由于遗传病的异质性及较大片段基因的检测需求,新一代测序技术应运而生。

3.2 第二代测序技术(next-generation sequencing, NGS) 单基因遗传病具有明确的临床、生化指征和/或已知突变热点,则对靶区域进行 Sanger 测序可精确高效做出分子诊断;然而筛选未知候选基因是一大难题。同时,利用连锁分析和关联分析定位遗传病致病基因的传统方法耗时费力,且当家系中受累亲属人数有限、或疾病遗传不外显或外显不全,亦或是患者为自发突变,则连锁分析法难以进行^[29]。新兴的第二代测序技术是用内切酶将基因组DNA处理成一定长度范围的DNA片段,然后进行文库制备并扩增文库DNA,测序过程可分为两类:使用DNA聚合酶和荧光标记或H⁺标记的4种dNTP进行单碱基扩增延伸反应的合成测序;采用DNA连接酶和荧光标记的寡核苷酸探针进行杂交反应的连接测序^[30]。NGS可快速进行大量靶基因筛选、人类全外显子组测序、全基因组测序,通量更高、检测更快^[31]。

3.3 基于NGS技术的高通量测序检测

3.3.1 全外显子组捕获技术(whole exome sequencing, WES) 现今已报道的绝大多数全基因组研究都仅集中在整个基因组序列的2%编码区,且大约85%的已知孟德尔遗传病致病基因都位于蛋白质编码区域^[32]。近年来全外显子组捕获技术的发展为遗传病快速诊断创造了条件。2009年,Ng等^[33]对4例无亲缘关系 Freeman-Sheldon 综合征患者及8例正常个体进行外显子组捕获和第二代测序,在世界上首次证实外显子组捕获—第二代测序在确定罕见致病基因方面具有可行性和应用价值。随后 Ng 等^[34]利用相同的方法策略检测出 Miller 综合征的致病基因 DHODH,这也是国际上首次成功应用外显子组捕获—第二代测序技术检出遗传病的未知致病基因。该技术所需样本数量少、通量高、耗费低,大大推进了人类单基因遗传病的分子诊断。

外显子组捕获集中于外显子区域的测序,不能获得完整的基因组信息,如启动子区、增强子区、mi-croRNAs 编码区等;其次,外显子组捕获过程中存在打断再拼接,读长较短将导致大的插缺难以拼接,故 DNA 结构变异无法检测;再者,难以检出位于高度同源区的变异体,易出现假阳性、假阴性结果^[35]。

3.3.2 分子倒位探针平行测序(MIPs) 分子倒位探针(molecular inversion probes, MIPs)技术所采用的探针是与基因组靶点互补的单链 DNA 分子,可与基因组靶点杂交并将其捕获。MIPs 探针由两侧向内部共包含 7 个基本结构^[36-38]:位于两端的 2 段与目的基因互补的序列;向内为 2 段对所有 MIPs 通用的 PCR 引物序列,可视为探针释放位点,通常还包含一个限制酶切位点;以及一段探针特异性标签序列和一个限制酶切位点,作为标签释放位点。当探针的互补序列与基因组靶点结合后环化,在内切酶 I 的作用下释放探针形成倒位序列,然后以通用引物完成靶序列富集,内切酶 II 处理释放 Tag 标签序列,进而实现目的基因测序。该技术具有高通量、高灵敏度、高准确性,目前多重 MIPs 分析技术已可在同一体系检测超过 55 000 个基因座^[36],适用于大规模检测。分子倒位探针平行测序技术已广泛应用于 SNP 分型、拷贝数变异检测、等位基因失衡分析等。

单基因遗传病具有复杂的表型异质性及遗传异质性,基因突变可涉及点突变、片段缺失/重复、整个基因的缺失/重复、以及动态突变等众多类别,针对不同遗传病使用不同的解决方案,是目前单基因遗传病分子诊断的主流思想。随着分子遗传学诊断技术的飞速发展,基因诊断新技术将广泛应用于出生缺陷与遗传病诊断中。基因诊断相关人员不仅要了解遗传疾病致病基因及其突变特点,还要熟知各种检测技术的特点,合理选择。

参 考 文 献

[1] 傅松滨.医学遗传学[M].第3版.北京:北京大学医学出版社,2013.
[2] 马端,周文浩,黄国英.罕见病并不罕见[J].中国循证儿科杂志,2011,6(2):83-86.

[3] Jie Zhang, Bao-Sheng Zhu, Jing He, et al. The spectrum of α - and β -thalassemia mutations in Yunnan Province of South-western China [J]. Hemoglobin, 2012, 36(5): 464-473.
[4] Genes and human disease [R/OL]. <http://www.who.int/genomics/public/geneticdiseases/en/index2.html>. [2014-01-16]
[5] Kan YW, Golbus MS, Dozy AM. Prenatal diagnosis of alpha-thalassemia. Clinical application of molecular hybridization[J]. N Engl J Med, 1976, 295 (21): 1165 -1167.
[6] Saiki RK, Walsh PS, Levenson CH, et al. Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotide probes[J]. Proc Natl Acad Sci USA, 1989, 86: 6230-6234.
[7] Chan V, Yam I, Chen FE, et al. A reverse dot-blot method for rapid detection of non-deletion alpha thalassaemia[J]. Br J Haematol, 1999, 104: 513-515.
[8] Rudkin GT, Stollar B. High resolution detection of DNA-RNA hybrids in situ by indirect immunofluorescence[J]. Nautre, 1977, 265: 472-473.
[9] Manzur A, Muntoni F. Diagnosis and new treatments in muscular dystrophies[J]. J Neurol Neurosurg Psychiatry, 2009, 80(7): 706-714.
[10] Miller MB, Tang YW. Basic concepts of microarrays and potential applications in clinical microbiology[J]. Clin Microbiol Rev, 2009, 22(4): 611-633.
[11] Yoo SM, Choi JH, Lee SY, et al. Applications of DNA microarray in disease diagnostics [J]. J Microbiol Biotechnol, 2009, 19(7):635-646.
[12] Hindson CM, Chevillet JR, Briggs HA, et al. Absolute quantification by droplet digital PCR versus analog real-time PCR[J]. Nature Methods, 2013, 10(10): 1003-1005.
[13] Crépin M, Pigny P, Escande F, et al. Evaluation of denaturing high performance liquid chromatography for the mutational analysis of the MEN1 gene[J]. J Mol Endocrinol, 2006, 36(2): 369-376.
[14] Wittwer CT, Reed GH, Gundry CN, et al. High-resolution genotyping by amplicon melting analysis using LCGreen[J]. Clin Chem, 2003, 49: 853-860.
[15] Schouten JP, McElgunn CJ, Waaij R, et al. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification [J]. Nucleic Acids Res, 2002, 30(12): e57.
[16] Hudecova I. Digital PCR analysis of circulating nucleic acids [J]. Clinical Biochemistry, 2015, 48(15): 948-956.
[17] Vogelstein B, Kinzler KW. Digital PCR[J]. Proc Natl Acad

- Sci USA, 1999, 96(16): 9236-9241.
- [18] Liu J, Hansen C, Quake SR. Solving the "world-to-chip" interface problem with a microfluidic matrix[J]. Anal Chem, 2003, 75(18): 4718-4723.
- [19] Chu T, Bunce K, Hogge WA, et al. Statistical considerations for digital approaches to non-invasive fetal genotyping[J]. Bioinformatics, 2010, 26(22): 2863-2866.
- [20] Lam KW, Jiang P, Liao GJ, et al. Noninvasive prenatal diagnosis of monogenic diseases by targeted massively parallel sequencing of maternal plasma: application to β -thalassemia [J]. Clin Chem, 2012, 58(10): 1467-1475.
- [21] Tsui NB, Kadir RA, Chan KC. Noninvasive prenatal diagnosis of hemophilia by microfluidics digital PCR analysis of maternal plasma DNA[J]. Blood, 2011, 117(13): 3684-3691.
- [22] Liu W, Smith DI, Reichtzgel KJ, et al. Denaturing high performance liquid chromatography (DHPLC) used in the detection of germline and somatic mutations[J]. Nucleic Acids Res, 1998, 26(6): 1396-1400.
- [23] Reed GH, Kent JO, Wittwer CT. High-resolution DNA melting analysis for simple and efficient molecular diagnostics [J]. Pharmacogenomics, 2007, 8(6): 597-608.
- [24] Reed GH, Wittwer CT. Sensitivity and specificity of single-nucleotide polymorphism scanning by high-resolution melting analysis[J]. Clin Chem, 2004, 50: 1748-1754.
- [25] White S, Kalf M, Liu Q, et al. Comprehensive detection of genomic duplications and deletions in the DMD gene, by use of multiplex amplifiable probe hybridization [J]. Am J Hum Genet, 2002, 71: 365-374.
- [26] Liborio Stuppia, Ivana Antonucci, Giandomenico Palka, et al. Use of the MLPA assay in the molecular diagnosis of gene copy number alterations in human genetic diseases[J]. Int J Mol Sci, 2012, 13: 3245-3276.
- [27] Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors[J]. Proc Natl Acad Sci USA, 1977, 74(12): 5463 - 5467.
- [28] Ronaghi M, Uhlen M, Nyren P. A sequencing method based on real-time pyrophosphate[J]. Science, 1998, 281(5375): 363.
- [29] Franke L, van BH, Fokkens L, et al. Reconstruction of a functional human gene network, with an application for prioritizing positional candidate genes [J]. Am J Hum Genet, 2006, 78: 1011-1025.
- [30] Mardis ER. Next-generation DNA sequencing methods [J]. Annu Rev Genomics Hum Genet, 2008, 9(1): 387-402.
- [31] Metzker ML. Sequencing technologies-the next generation [J]. Nat Rev Genet, 2010, 11(1): 31-46.
- [32] Botstein D, Risch N. Discovering genotypes underlying human phenotypes: past successes for Mendelian disease, future approaches for complex disease [J]. Nat Genet, 2003, 33 (Suppl): 228-237.
- [33] Ng SB, Turner EH, Robertson PD, et al. Targeted capture and massively parallel sequencing of 12 human exomes [J]. Nature, 2009, 461(7261): 272-276.
- [34] Ng SB, Buckingham KJ, Lee C, et al. Exome sequencing identifies the cause of a mendelian disorder [J]. Nat Genet, 2010, 42(1): 30-35.
- [35] Gilissen C, Hoischen A, Han G Brunner, et al. Disease gene identification strategies for exome sequencing [J]. Eur J Hum Genet, 2012, 20: 490-497.
- [36] Turner EH, Ng SB, Nickerson DA, et al. Method for genomic partitioning [J]. Annu Rev Genomics Hum Genet, 2009, 10: 263-284.
- [37] Hardenbol P, Banér J, Jain M, et al. Multiplexed genotyping with sequence-tagged molecular inversion probes [J]. Nat Biotechnol, 2003, 21(6): 673-678.
- [38] Hardenbol P, Yu F, Belmont J, et al. Highly multiplexed molecular inversion probe genotyping: over 10,000 targeted SNPs genotyped in a single tube assay [J]. Genome Res, 2005, 15(2): 269-275.

(收稿日期:2016-07-22)

编辑:宋文颖