

性发育异常疾病遗传学诊断研究进展

Advances in genetic diagnosis for disorders of sexual development

李佳楠¹, 杨季云^{1,2△}

LI Jia-nan, YANG Ji-yun

1. 电子科技大学医学院, 四川 成都 610054; 2. 四川省医学科学院·四川省人民医院(电子科技大学附属医院)

医学遗传中心, 四川 成都 610072

【摘要】 性发育异常(disorders of sex development, DSD)是一大类性腺结构及功能异常的先天性疾病,其主要表现为生殖器外观、性腺结构、性染色体核型不一致等。DSD 临床症状高度异质性,从生殖器外观和性腺功能的微小变化到严重的非典型性腺和生殖器都有可能发生。DSD 的病因非常复杂,仅通过临床表现、影像学 and 内分泌检查等手段,很难对 DSD 进行明确的诊断与鉴别诊断,这对患者的性别选择、性心理发展以及个体化的治疗影响很大。遗传学诊断技术的发展开启了 DSD 患者诊断的新时代,已经彻底改变了 DSD 诊断过程。本文综述了 DSD 遗传学诊断及相关研究进展,为这类疾病的预防、个体化治疗、改善疾病预后以及遗传咨询提供参考。

【关键词】 性发育异常;遗传学诊断;基因

【中图分类号】 R729

【文献标志码】 B

【文章编号】 1672-6170(2024)04-0195-05

性发育异常(disorders of sex development, DSD)疾病是指外生殖器、性腺结构以及性染色体核型不一致的一类先天性疾病。其主要的临床表现为:内、外生殖器模糊不清,患者可同时有男性与女性的两性特征,症状严重者甚至无法区别男女性别^[1]。DSD 具有高度的临床异质性,表现为多种不同的临床表型。多数 DSD 患者出生时即表现出外生殖器的异常,如具有女性外生殖器的患者阴蒂的

增大;具有典型男性外生殖器的患者表现为阴茎的短小,尿道口的异常,隐睾等;更有甚者外生殖器严重异常,无法区分其性别。部分 DSD 患者出生时无明显外生殖器官异常,直到青春期才表现为青春期延迟或缺乏。还有一部分患者因成年后不孕不育就诊,经过检查才临床诊断为 DSD^[2]。在活产儿中,生殖器官模糊的患儿的发生率约为 1:2000~1:4500,且不同人种之间差异很大^[1]。由于 DSD 病

- [27] Gouriou C, Chambaz M, Ropert A, et al. A systematic literature review on solitary rectal ulcer syndrome: is there a therapeutic consensus in 2018 [J]. *Int J Colorectal Dis*, 2018, 33 (12): 1647-1655.
- [28] Zergani FJ, Shaisthe AA, Hajiani E, et al. Evaluation of argon plasma coagulation in healing of a solitary rectal ulcer in comparison with conventional therapy: a randomised controlled trial [J]. *Prz Gastroenterol*, 2017, 12(2): 128-134.
- [29] Shah A, Bohra S, Desai S. Argon plasma coagulation-an effective treatment for solitary rectal ulcer syndrome: A single-center experience from western India [J]. *Indian J Gastroenterol*, 2021, 40(1): 35-40.
- [30] Rau BK, Harikrishnan KM, Krishna S. Laser therapy of solitary rectal ulcers: a new concept [J]. *Ann Acad Med Singap*, 1994, 23 (1): 27-28.
- [31] Keshtgar AS, Ward HC, Sanei A, et al. Botulinum toxin, a new treatment modality for chronic idiopathic constipation in children: long-term follow-up of a double-blind randomized trial [J]. *J Pediatr Surg*, 2007, 42(4): 672-680.
- [32] Keshtgar AS. Solitary rectal ulcer syndrome in children [J]. *Eur J Gastroenterol Hepatol*, 2008, 20(2): 89-92.
- [33] Shiratori Y, Ikeya T, Fukuda K. Endoscopic band ligation for persistent mucosal prolapse syndrome [J]. *Gastrointest Endosc*,

2023, 98(5): 870-871.

- [34] Tsuji T, Ootani S, Aoki H, et al. Two cases of the rectal mucosal prolapse syndrome in which endoscopic submucosal dissection was effective. [J]. *Nihon Syoukaki Naishikyoku Gakkai Zasshi (Gastroenterol Endosc)*, 2013, 55: 250-256.
- [35] Ohta H, Takamura K, Fujita T, et al. A case of colitis cystica profunda: the usefulness of a therapeutic trial by endoscopic submucosal dissection [J]. *Nihon Syoukaki Naishikyoku Gakkai Zasshi (Gastroenterol Endosc)*, 2015, 57: 2469-2475.
- [36] Hayasaka J, Hoteya S, Tomizawa K, et al. The long-term efficacy of endoscopic submucosal dissection in the treatment of symptomatic mucosal prolapse syndrome [J]. *Intern Med*, 2021, 60 (7): 1005-1009.
- [37] Hayasaka J, Hoteya S, Ochiai Y, et al. Endoscopic submucosal dissection improves bloody stool associated with polypoid type mucosal prolapse syndrome: A case series [J]. *Intern Med*, 2022, 61(21): 3211-3215.
- [38] Abdi S, Tavakolikia N, Yamini M, et al. Solitary rectal ulcer syndrome: addition of rectal therapies to biofeedback is more effective than biofeedback alone [J]. *Gastroenterol Hepatol Bed Bench*, 2019, 12(3): 197-202.

(收稿日期:2024-02-06;修回日期:2024-03-20)

(本文编辑:林 赞)

因复杂且伴有高度的临床异质性,因此要想对 DSD 患者进行精准诊断,进而帮助患者性别选择、个体化治疗、改善预后及预防疾病,遗传学诊断显得非常关键。

1 DSD 的分类及遗传病因

依照欧洲 (ESPE) 和美国儿科内分泌学会 (LWPES) 的分类法,把 DSD 按染色体核型分为以下三类:性染色体 DSD、46, XX DSD、46, XY DSD^[3]。性染色体 DSD 是由于性染色体非整倍体异常或结构异常导致。DSD 中约 15% 左右患者为性染色体 DSD。常见的性染色体 DSD 核型主要包括性染色体非整倍异常,如 45, X、47, XXY; 性染色体嵌合体,比如 46, XX/46, XY 嵌合体、45, X/46, XY 嵌合体等。其中 45, X 主要导致 Turner 综合征; 46, XX/46, XY 嵌合导致卵睾型 DSD; 47, XXY 与 Klinefelter 症候群有关,而混合型性腺发育不良患者主要核型为 45, X/46, XY 嵌合^[2, 4~5]。

46, XY DSD 的患者为正常男性 46, XY 核型。该类疾病是 DSD 中最常见的一个类别,约占 DSD 病例的 50%。它可以分为性腺(睾丸)发育不良,雄激素合成异常或功能异常及其他类型。性腺发育不良的患儿性腺表现出不同程度的睾丸发育不良,如苗勒管的残留、外生殖器的雌性化。还能表现为睾丸退化综合征、部分和完全性腺发育不良等。雄激素合成异常或功能异常的 DSD 患者以男性化不全为主要表现,通常无苗勒管结构残留。雄激素受体、LHCGR、甾体激素合成通路酶和睾酮转化途径酶等基因变异是导致该类疾病的遗传病因。目前已发现至少 46 个与 46, XY DSD 相关致病基因^[2, 4, 5]。

46, XX DSD 患者的染色体核型为正常女性核型即其染色体核型为 46, XX, 约 35% 的 DSD 患者为该类型。46, XX DSD 又可分为雄激素过多、性腺发育异常和单纯生殖道发育异常三大类。大多患者在出生后由于外生殖器外观异常就诊,也有部分患者因青春期后的男性化或青春发育延迟就诊。目前已发现 61 个 46, XX DSD 致病基因^[2, 4, 5]。

2 DSD 遗传学诊断的常用技术

2.1 染色体核型分析和荧光原位杂交技术 (fluorescence in situ hybridization, FISH)

DSD 诊断依赖于临床表型、内分泌激素和核型分析。外周血染色体核型分析可用于性染色体非整倍体异

常、性染色体嵌合体、5Mb 以上的染色体结构异常。因此,染色体核型分析方法将首先用于确定 DSD 的分型^[4]。与染色体核型分析技术比较,FISH 检测不需要细胞培养,能对间期细胞进行检测,通常 24 小时内可完成对性染色体的检测。而染色体核型需要细胞培养,检测时间需要 1~2 周。因此运用 FISH 技术对 X、Y 染色体以及 SRY 基因进行检测可在较短时间内将 DSD 患者分型为 46, XY DSD、46, XX DSD 或性染色体 DSD^[6]。

2.2 拷贝数变异测序技术 (copy Number Variation sequencing, CNV-seq) 和染色体微阵列分析技术 (Chromosomal microarrays, CMA)

现已证明拷贝数变异 (Copy number variations, CNVs) 是 DSD 的遗传病因之一,特别是合并发育迟缓的 DSD 综合征,比如泄殖腔畸形或 Mayer-Rokitansky-Küster-Hauser 综合征 (MRKH 综合征)^[7, 8]。某些参与性腺发育调控的基因具有剂量依赖性,如 FGF9、SOX3 与 SOX9 重复可导致 46, XX DSD; DAX1 和 WNT4 重复可导致 46, XY DSD。而 ATRX、DMRT1、EMX2 与 WT1 缺失可导致 46, XY DSD^[9]。CMA 与 CNV-seq 是检测 CNVs 的临床适宜技术。尽管 CMA 与 CNV-seq 能探测基因组约 25~50 kb 的 CNVs,甚至能检测更小的 CNVs,但是临床实验室一般仅报告大于 100 kb 的 CNVs。一些生物信息学软件可对大于 1 kb 的 CNV 进行分析及功能注释,运用此类软件在 52 个 DSD 病例中鉴定出 300 多个 CNV,这些 CNV 涉及到 68 个 DSD 致病基因^[9]。CNV-seq 和 CMA 也能够检出低比率嵌合体,在理想的情况下, CNV-seq 甚至可检出低至 5% 的染色体非整倍体嵌合^[10]。

2.3 第二代测序技术 (Next generation sequencing, NGS)

NGS 能一次对大量核酸分子进行大规模平行测序,一次测序反应可生成大量测序数据。与第一代 Sanger 测序比较,NGS 测序具有更快的速度、更高的通量。在生成相同单位测序数据量方面,NGS 有更低廉的成本优势。NGS 在疾病临床诊断的应用主要有:①全基因组测序 (Whole genome sequencing, WGS): 对人类基因组的全部序列进行检测,包括编码和非编码 DNA 以及线粒体基因组。②全外显子组测序 (Whole exome sequencing, WES): 运用目标序列捕获技术对超过 20,000 个人类基因外显子及其比邻 20bp 内含子区域进行捕获,构建高通量测序文库并进行测序。在人类遗传疾病中,约 85% 的致病突变存在于外显子区。相较 WGS 而言,它具有成本低廉、数据处理更方便的优势。③靶向序列测序技术 (Targeted panel

sequencing, TPS): 针对某种特定类型疾病(如 DSD)设计可捕获所有已知的 DSD 致病基因外显子区域探针,并可增加能捕获位于内含子区致病变异的探针,对捕获序列进行高通量测序。相对于 WES, TPS 的捕获效率更高,对于单样本来说,测序成本更低。同时通过增加测序深度可降低变异检测错误率,提高低比例嵌合体检出率。TPS 仅针对 DSD 致病基因,数据分析周期更短。但是 TPS 也存在局限性, TPS 检测目标基因均为现阶段已知的致病基因,每年均会有新的疾病致病基因被鉴定,需要定期更新 TPS 检测的致病基因。

在 80 例临床患者中进行了 30 个 DSD 致病基因的 TPS 检测研究显示, 25 例确诊了遗传病因。在 AMH、AMHR2、AR、HSD17B3、HSD3B2、MAMLD1、NR5A1、SRD5A2 和 WT1 等基因中检出致病变异。46, XY DSD(25/73)的诊断率为 34.2%^[11]。而 141 个 DSD 致病基因 TPS 检测结果显示, 50 个 46, XY DSD 患者中有 16 例患者检测出致病变异或可疑致病变异, 诊断率约为 32%。其中 SRD5A2 基因变异检出率最高, 其次为 AR 基因^[12]。应用 2742 个致病基因的 TPS 对 125 例 DSD 患者检测的研究显示, 46, XY DSD 患者诊断率为 46.9%, 显著高于 46, XX DSD 约 10.3% 的诊断率^[13]。在 14 个 TPS 检测阴性样本中, WES 仅检出一个临床意义不明的变异, 提示 WES 并没有显著提高诊断率。总体而言, TPS 或 WES 对 DSD 患者的诊断率 20.7% ~ 38.6%。WES 和 TPS 在 DSD 检测中有各自的优点。对于经过 TPS 检测未确诊的 DSD 患者, 如果有新的致病基因被发现, 存在疾病漏诊可能。因此, 需要更新 TPS 可检测的 DSD 基因。而经过 WES 检测未确诊的 DSD 患者, 则可利用 WES 测序数据重分析的方法发现 DSD 新致病基因^[14]。现有的研究结果显示 NGS 方法可提高 DSD 患者的诊断率并对 DSD 患者治疗有指导意义, 建议将 NGS 作为临床有效且具有成本效益的一线基因检测方法^[11, 15, 16]。

WGS、WES 和 TPS 都是建立在短读测序(short reads sequencing, SRS)原理的基础上。与 WES 和 TPS 比较, WGS 无捕获环节, WGS 能更均匀地覆盖那些高 GC 含量的外显子区域。在一项 WES 和 WGS 的测序质量对比研究中, 对单核苷酸变异(SNV)和小插入/缺失(indels)进行了检测, 结果显示: WGS 的覆盖深度、变异位点测序质量都比 WES 更均匀。研究结果显示, 共有 656 个基因编码区的 SNV 能被 WGS 检测而不能被 WES 检测, 约占所有编码区 SNV 的 3%。远远高于被 WES 检测而被 WGS 遗漏的编码区 SNV 数量(26 个变异)。对于

indels, WES(44%)和 WGS(46%)的假阳性变异比例相似。结果显示 WGS 在检测编码区内致病变异, 特别是 SNV 方面比 WES 更强大^[17]。因此, 预计 WES 会漏诊很多致病变异。在 ClinVar 数据库中, 有 89% 的致病变异都位于编码区, 编码区只占基因组的 2%。WGS 不仅能检测编码区变异, 还能检测到编码区外的变异。WGS 已经被证明能提高儿科罕见病的诊断率以及改进临床治疗效果^[18, 19]。WGS 检测覆盖了 DSD 致病基因的启动子、增强子及内含子区域, 预计 WGS 也可提高 DSD 患者的诊断率^[20]。然而, 与 WES 相比, WGS 检测费用仍然较高, 并且在非编码区变异位点致病性的解释上存在一定困难。虽然 WGS 具有明显的技术优势, 但是在 DSD 基因诊断中 WGS 应用仍然较少。

3 改善 DSD 病因遗传学诊断新技术

目前临床应用的 DSD 遗传学检测技术主要能检测 DSD 致病基因编码区的 SNVs、indels 及 CNVs。基因调控区域和内含子区域的变异、复杂基因组结构变异和表观遗传学改变等与 DSD 发病有关。如果针对这些 DSD 致病机制进行病因学检测, 可提高 DSD 的遗传学诊断率。

3.1 第三代测序技术(The third generation sequencing, TGS) 与 NGS 序列读长普遍为 35 ~ 700 bp, TGS 读长一般在 10 kb 以上, 因此 TGS 也被称为“长读测序”(long-read sequencing, LRS)。按照 TGS 的原理, TGS 主要分成单分子实时测序(single-molecule real-time sequencing, SMRT)及纳米孔测序(oxford nanopore sequencing technology, ONT)技术^[21]。利用 TGS 的长读长及对基因组复杂区域测序的技术优势, TGS 可应用于从头组装人类基因组, 也可用于其它物种如水稻、小麦、猪、鸡、牛、羊等动植物基因组的从头构建^[22]。

结构变异(structural variation, SV)是指基因组中超过 50 个碱基的变异。主要有易位、倒位、插入、缺失、可移动元件插入(Mobile Element Insertion, MEI)、串联重复与 CNV 等。SV 在遗传疾病中起重要作用^[23]。两个单倍体人类基因组 SMRT 测序结果显示, 在千人基因组计划中, 估计约有 89% 的 SV 被遗漏^[23]。TGS 显示出较好的发现 SV 的技术优势。巴尔得-别德尔综合征(Bardet-Biedle syndrome, BBS)是一种涉及性发育异常的常染色体隐性遗传罕见疾病。其主要临床表现有: 肥胖、智力落后、性腺功能低下、视网膜色素变性。可伴有多种畸形, 如多指(趾)、肾结构异常等^[24]。在临床被诊断为 BBS 的男性患儿中, CMA 和 NGS 检测到 BBS9 外显子 1 ~ 4 纯合缺失, 该缺失位于 7 号染色体 p14.2 ~

p21.1 的 17.5 Mb 杂合性缺失区域。通过 TGS 技术确定了先证者及其父母存在 72.8 kb 缺失,并精确定位了断裂点,显示出长读测序在精确鉴定 SV 核苷酸断点的技术优势。

大约 90% 的先天性肾上腺增生症 (CAH) 病例是由 CYP21A2 的变异导致 21-羟化酶功能的缺陷^[25]。CYP21A2 与假基因 CYP21A2P 串联排列在 RCCX 模块的遗传单元中,其中还包含 TNXA/B、C4A/B 和 RP 基因^[26]。CYP21A2 基因与 CYP21A2P 假基因的核酸序列具有高度同源性,而且真基因与假基因可发生复杂的基因重排。当前临床广泛应用的 CYP21A2 基因检测方法,如一代测序结合多重连接探针扩增技术 (Multiplex ligation-dependent probe amplification, MLPA) 仅能检测 CYP21A2 基因点突变及外显子拷贝数变异,无法检测复杂基因重排。这种方法存在一定的假阴性和假阳性率。对于这种具有复杂结构的基因变异,以短读长为基础的 NGS 也面临着类似的问题。应用 TGS 技术可通过一次实验即可完成 5 个 CAH 致病基因的全面的检测和分析。能准确地检测 CYP21A2 真、假基因变异、基因外显子拷贝数变异、基因重排导致的嵌合基因亚型及断裂位置。此外, TGS 无需其它家系成员样本,只需检测患者就可以确定突变等位基因的顺、反式结构。临床研究结果表明,该方法检测费用接近其它临床检测技术,在新生儿的基因筛查、携带者和疾病诊断等方面具有很好的临床应用前景^[27]。

3.2 全基因组光学图谱 (Optical genome mapping, OGM) OGM 是一种新型超长单分子检测技术,其通过对 DNA 分子特异性识别序列 CTTAAG 进行标记,结合微流控、高分辨率显微镜和自动图像分析技术,实现对完整 DNA 分子上标记信号模式的直接可视化分析。其基本技术原理与染色体核型分析类似,但 OGM 标记信号更多,因此检测分辨率更高 (500 bp ~ 5 kb)。一项多中心联合研究表明,在 85 份临床样本的检测中,该技术检出 99 个染色体的结构异常 (包括:非整倍体、重复、缺失、易位、倒位、插入、复杂重排等),精准定位染色体发生断裂的具体位置,并能确定染色体重复或插入片段的方向,检出率达到了 100%^[28]。但 OGM 无法对缺乏参考序列和标记的区域 (如着丝粒,端粒,异染色质区域) 进行检测,对某些变异类型 (如罗伯逊易位等) 可能产生漏检^[28]。

OGM 与染色体分析技术 (如核型、CNV-seq 及 CMA) 相比,显著提高了 SV 检测的分辨率,特别是提高了平衡性 SV 的检出。已有研究团队把 OGM

用于 DSD 的诊断中^[29]。WGS 和 OGM 联合应用几乎能检测所有类型的基因组变异:INDELs、SNV 和 SV。两种技术的联合应用极大地提高了 SV 检出率,同时也将识别 SV 断裂点的分辨率提至 < 140bp^[5]。预计 WGS 和 OGM 联合应用能提高 DSD 和其他遗传病的诊断率。

3.3 转录组测序技术 (Transcriptome Sequencing)

转录组测序是通过高通量测序定量或定性地检测细胞或组织中部分或全部的 mRNA、small RNA 以及 no-codingRNA。因此,该技术也叫做 RNA 测序 (RNA-Seq)。该技术能对人类基因转录本的结构和表达量进行分析,发现罕见的、甚至未知的转录本,识别基因可变剪切位点^[30]。在所有非编码突变中,约 9% ~ 30% 的变异会对 RNA 的加工和表达产生影响从而导致疾病发生。利用转录组测序技术,研究非编码区及基因编码区变异位点对 RNA 剪切与表达的影响,可发现一些尚未确诊的遗传性疾病的遗传病因,提高遗传病的诊断率。RNA-seq 还能用于疾病等位基因表达分析,有助于理解基因型与疾病易感性、疾病表型之间的关系等^[30]。

在 DSD 的遗传学诊断中,常常会检出临床意义不明的变异 (VUS)。通过 RNA-seq 分析特定基因的 RNA 表达水平或相应转录本,可评估 VUS 的功能进而评估其致病性。但仅有约 50% 的 DSD 致病基因在白细胞中表达,不是所有的患者均能提供致病基因表达的靶组织 (如皮肤成纤维细胞、性腺组织) 进行 RNA-seq。因此基于 RNA-seq 的 DSD 诊断应用受到检测靶组织可用性的限制。此外,不同组织器官、细胞类型、不同发育阶段基因表达均有不同,基因表达会随组织器官及细胞内外的刺激而发生改变。特定组织器官及细胞的转录组结果并不一定能够反应疾病的基因表达真实情况,因此 RNA-seq 并不能适用于所有疾病^[31]。

4 小结

遗传学诊断技术的发展开启了 DSD 患者诊断的新时代,已经彻底改变了 DSD 诊断过程。目前,仅有约 35% ~ 45% 的 DSD 能明确遗传病因。临床实践中检出大量的临床意义不明变异位点与疾病关系有待确认,仍有一些未知 DSD 致病基因有待发现。多种技术综合应用、多组学大数据整合分析将有助于发现新的 DSD 致病基因,更深刻地理解 DSD 的发病机制。综合应用基因组、转录组和表观遗传学等多种技术可提高遗传诊断的敏感性与特异性,缩短疾病诊断周期,为 DSD 患者提供更精准的治疗指导,这将是未来 DSD 遗传诊断的重要发展方向。