

遗传基因检测

DOI: 10.13406/j.cnki.cyxh.002865

神经系统遗传病的基因检测策略

罗巍, 陈思

(浙江大学医学院附属第二医院神经内科, 杭州 310009)



【摘要】近年来,各种基因检测技术包括单基因测序、基因 Panel、全外显子组测序和全基因组测序等被广泛应用于神经系统遗传病的诊断中。基因检测技术的快速发展使临床医生能够探索表型更为复杂、遗传异质性更大的疾病,帮助他们疾病作出快速且精准的诊断。本文旨在帮助临床医生在种类繁多的基因检测技术中合适的方法辅助诊断神经系统遗传病。

【关键词】神经系统遗传病;基因检测;方法

【中图分类号】R741.04

【文献标志码】A

【收稿日期】2021-02-15

Strategies of molecular genetic testing for neurogenetic diseases

Luo Wei, Chen Si

(Department of Neurology, The Second Affiliated Hospital, Zhejiang University School of Medicine)

【Abstract】Nowadays various types of genetic tests including single gene testing, Panel testing, whole-exome sequencing and whole-genome sequencing, have been widely used in the diagnosis of neurogenetic diseases. With the advances in genetic testing, physicians are able to reach a rapid and precise diagnosis of diseases with vast genetic heterogeneity and complex phenotypes. This overview of genetic testing strategies aims to help physicians to select the test that is most appropriate for their practice setting.

【Key words】neurogenetic diseases; genetic testing; methods

神经系统遗传病种类繁多,且基因型-表型关系复杂,一个基因常与多种表型有关,如 *PMP22* 基因,其上发生片段缺失会导致遗传性压力易感性周围神经病,若该基因发生片段重复则又引起腓骨肌萎缩症^[1];而一个表型也常有多个致病基因,如多巴胺反应性肌张力障碍,其致病基因有 *GCH1*、*SRP* 和 *TH* 等,它们因在多巴胺合成通路的各个环节中发挥作用而致病^[2]。近年来,各种新兴基因检测技术帮助临床医生发现和识别大量与神经系统遗传病相关的致病基因和表型,提高

了人们对遗传病的认识,使临床诊断变得更加高效和精准。因此,在神经系统遗传病中合理应用基因检测尤为重要。它不仅可以帮助明确临床诊断,也为后续的治疗带来希望,同时还能够指导家庭优生优育。本文从各种基因检测方法的适用范围、表型对基因检测的影响、基因检测报告的解读以及目前基因检测中还存在的问题进行简要概述。

1 不同基因检测技术适用于不同的变异类型

1.1 单基因测序

单基因测序(single gene testing)可选择的检测手段主要有二代测序(sanger sequence)、多重连接探针扩增技术(multiplex ligation-dependent probe amplification, MLPA)、重复引物 PCR(repeat-primed polymerase chain reaction, RP-PCR)、长片段 PCR 技术(long-range PCR)及 Southern blot 等。二代测序可用于检测单个基因编码序列上发生的错义突变(mis-sense)、同义突变(synonymous)、无义突变(nonsense)、小片段的插入/缺失突变(small insertions/deletions)和剪接位点突变(splicing)^[3];MLPA 则用来检测基因上的拷贝数变异(copy

作者介绍:罗巍,主任医师,博士生导师,现任浙江大学医学院附属第二医院神经内科科室副主任,曾先后前往美国芝加哥大学医学中心、英国伦敦大学学院、法国 Grenoble 大学医院、美国西北大学访问。担任 SCI 收录期刊 *Am J Med Genet B Neuropsychiatr Genet* 编委、《中华医学遗传学杂志》通信编委。其团队长期聚焦于帕金森病等运动障碍疾病与神经系统遗传病的临床诊治和基础研究,近几年在原发性家族性脑钙化和家族性皮质肌阵挛性震颤癫痫这两种疾病中分别鉴定了一个新的致病基因。近 5 年来以通信作者发表 SCI 收录论文 20 余篇,包括 *Brain* 2 篇和 *Movement Disorders* 7 篇。研究方向:帕金森病等运动障碍疾病与神经系统遗传病。Email: Luoweirock@126.com。

优先出版: <https://kns.cnki.net/kcms/detail/50.1046.R.20210616.1103.004.html>

(2021-06-16)

number variation, CNV); RP-PCR 可以检测 DNA 上的重复序列发生不稳定性扩增,即动态突变(dynamic mutation)。当患者具有以下特征时,临床医生可优选单基因测序^[4]:①明确的家族史;②临床表型和其他检验、检查结果都指向某一种特定类型的疾病;③该疾病目前仅有 1 个致病基因或某个基因上的致病突变最为常见;④患者经济条件受限。

一代测序适用于大多数基因测序,用于发现目标片段(<1 kb)上的点突变,曾被认为是过去十年临床基因检测的“金标准”^[5]。例如肝豆状核变性(Wilson disease, WD),它是一种常染色体隐性遗传的铜代谢疾病, *ATP7B* 是其唯一的致病基因,参与铜转运过程,78%的患者中存在 p.R778L、p.P992L 和 p.T935M 三个突变。因此,对临床上可疑的肝豆状核变性患者选用一代测序直接检测 *ATP7B* 基因不失为一种快速、敏感且价格低廉的基因检测方法^[6]。又如发作性运动诱发性运动障碍(paroxysmal kinesigenic dyskinesia, PKD),典型临床特征为儿童期起病的由随意运动诱发的短暂性单侧或双侧不自主运动,包括肌张力障碍、舞蹈样动作或手足徐动症等。曹立等^[7-8]发现约 1/3 的 PKD 患者携带常染色体显性遗传致病基因 *PRRT2* 上的突变,其中 c.649dupC (p.Arg217Profs*8) 引起的移码突变所占比例高达 76.47%,因此可直接对临床高度怀疑为 PKD 患者进行 *PRRT2* 基因的一代测序。此外,一代测序还多用于已明确致病突变的先证者家族成员的共分离验证。但是一代测序结果阴性只能排除目标区域的特定突变,对于大片段插入、拷贝数变异、多核苷酸重复扩增或非编码区序列突变等则不适用^[9]。另外,对于外显子数目较多的基因,使用该方法测序费时费力。

MLPA 技术多用于致病基因以拷贝数变异为主的疾病,例如腓骨肌萎缩症(Charcot-Marie-Tooth, CMT),它是最常见的遗传性周围神经病之一,其发病率约为 1/2 500。CMT 的经典表型发病年龄多在 10~20 岁,临床特征主要为肢体远端无力、感觉减退和足部畸形^[9]。目前报道和 CMT 相关的致病基因多达 100 余个,其中最常见的是染色体 17p11.2 区域上 1.5 Mb 范围的大片段重复,该区域包含编码周围髓鞘蛋白的 *PMP22* 基因。据文献报道 *PMP22* 基因突变占 CMT1 型的 70%,在所有 CMT 病例中占比高达 50%^[10]。樊东升等^[11]运用 MLPA 技术同时结合一代测序对 465 个中国 CMT 家系测序,发现 *PMP22* 基因重复占 29.5%,而 *PMP22* 基因点突变仅占 2.2%。由此可见,对于临床上可疑的 CMT 患者可以优先选择 MLPA 技术对 *PMP22* 基因进行检测,若首选全外显子组测序,虽可提高 CMT 相关的其他致病基因检测率,但因其检测技术的缺陷极易漏诊 *PMP22* 基因上的拷贝数变异。此外, X 连锁隐性遗传的杜氏肌营养不良(Duchenne muscular dystrophy, DMD)常在儿童早期发病,表现为近端肌无力、小腿肥大和肌酸激酶明显升高。它的致病基因为 *DMD*,其中由该基因上的大片段缺失导致的病例可达 65%~70%,多发生在

突变热点区域 45~53 号外显子^[12],因此临床上也常优先选择 MLPA。导致常染色体隐性遗传的脊髓性肌萎缩(spinal muscular atrophy, SMA)的致病基因 *SMN1*,其上约 96%的突变均为 7 号外显子的纯合缺失^[13];以及痉挛性截瘫(hereditary spastic paraplegia, HSP)最常见的致病基因 *SPG4*,其上 20%~25%的突变形式为拷贝数变异^[14]。所以对拷贝数变异为主的致病基因检测,临床上多选择方便且价格较低的 MLPA。

RP-PCR、长片段 PCR、Southern blot 及长读长测序技术(long-read sequencing, LRS)都可用于检测多核苷酸重复扩增突变。自 1991 年,首次发现 *FMR-1* 基因 5' 非翻译区(untranslated region, UTR)的 CGG 重复扩增可导致脆性 X 综合征(fragile X syndrome, FXS)以来,陆续发现多种疾病与多核苷酸重复突变相关^[15]。多核苷酸重复扩增突变可发生在编码区上,如脊髓延髓性肌萎缩(spinal and bulbar muscular atrophy, SBMA)由 *AR* 基因 1 号外显子上的 CAG 重复扩增致病^[16]、亨廷顿病(Huntington disease, HD)由 *HTT* 基因编码区的 CAG 重复扩增致病^[17];也可以出现在 UTR 区或内含子上,如强直性肌营养不良(myotonic dystrophy, DM)分为 DM1 型和 DM2 型,致病突变分别为 *DMPK* 基因 3'UTR 区的 CTG 重复^[18]和 *CNBP* 基因 1 号内含子 CCTG 重复^[19]。

近年来多核苷酸重复扩增已成为国内外研究热点,2018 年发现家族性皮质肌阵挛性癫痫(familial cortical myoclonic tremor with epilepsy, FCMTE)1 型由 *SAMD12* 基因内含子区域的五核苷酸重复(TTTCA)扩增致病^[20-21],后续多个研究团队相继报道不同基因(*STARD7*^[22]、*MARCH6*^[23]、*YEATS2*^[24]、*TNRC6A*^[20]及 *RAPGEF2*^[20])内含子上的 TTTCA 重复突变可引起不同的 FCMTE 亚型。2019 年英国团队鉴定出伴神经病变和前庭反射消失的小脑性共济失调综合征(cerebellar ataxia, neuropathy, vestibular areflexia syndrome, CANVAS)的致病突变为 *RFC1* 基因内含子区的 AAGGG 重复扩增,该疾病是一种罕见的晚发型共济失调^[25]。之后 Tian Y 等^[26]证实 *NOTCH2NL* 基因 5'UTR 区的 GGC 为神经元包涵体疾病(neuronal intranuclear inclusion disease, NIID)的致病基因,最近还发现该基因 GGC 重复与帕金森病(Parkinson's disease, PD)^[27]、阿尔茨海默病(Alzheimer's disease, AD)^[28]、肌萎缩侧索硬化(amyotrophic lateral sclerosis, ALS)^[29]、多系统萎缩(multiple system atrophy, MSA)^[30]及特发性震颤(essential tremor, ET)^[31]等疾病相关。2020 年 Deng J 等^[32]发现眼咽远端型肌病(oculopharyngodistal myopathy, OPDM)的另一致病基因 *GIPC1* 也为 5'UTR 区的 GGC 重复。随着基因检测技术的发展,尤其是长读长测序的应用,神经系统遗传病中关于多核苷酸重复扩增突变的研究进展突飞猛进。

1.2 基因 Panel

对于许多神经系统遗传病来说,临床表型相对复杂,而且导致这种疾病的致病基因往往不止一个,此时需要一个可

以检测多个基因的 Panel 来提高检测结果的阳性率。基于二代测序(next-generation sequencing)平台的基因 Panel 适用于以下情况^[4,33]:①疾病的遗传异质性较大,如 HSP、CMT、PD、原发性肌张力障碍(primary torsion dystonia,PTD)等;②临床表型相似的一组疾病,如肌病等;③临床表型相似的不同疾病,如遗传性脑白质病变,包括肾上腺脑白质营养不良(adrenoleukodystrophy)、遗传性弥漫性脑白质病变合并球状轴索(hereditary diffuse leukoencephalopathy with neuroaxonal spheroids,HDLS)、常染色体显性遗传成人型脑白质营养不良(autosomal dominant adult-onset demyelinating leukodystrophy,ADLD)、常染色体显性遗传性脑动脉病伴皮质下梗死和白质脑病(cerebral arteriopathy with subcortical infarcts and leukoencephalopathy,CADASIL)等。

基因 Panel 因为拥有较高的覆盖度、敏感性和特异性,与全外显子测序或全基因组测序相比其阳性率较高^[33]。不同基因检测公司对相同一组疾病制定的基因 Panel 并不完全相同,一味追求全面而纳入较多与特征疾病相距甚远的基因或外显率较低的基因,最终会检测到很多意义不明确(variant of uncertain significance,VUS)的变异,导致基因检测报告解读更加繁琐和复杂^[3]。因此,纳入检测的基因数目并非“越多越好”。譬如儿童期癫痫,含有 172 个相关基因 Panel 的诊断率约为 30%,其中 156 个基因未发现任何有意义的致病突变,说明诊断率不会随着纳入基因数目增多而呈线性增长^[34]。而有些基因检测公司则比较保守,只选择与特定疾病证据性较强的基因,又会导致其他相关基因的漏诊^[33]。因此,选择基因 Panel 测序时,作为临床医生需要仔细核对基因包中是否含有与关注表型相关的一类基因。

1.3 全外显子组测序和全基因组测序

全外显子组测序(whole-exome sequencing,WES)和全基因组测序(whole-genome sequencing,WGS)既往通常用于基因 Panel 阴性或复杂表型的患者,不仅可以评估疾病所有相关的已知基因,同时随着新基因的和数据库的更新,为将来的再分析提供了机会,也是鉴定发现新基因的重要方法。随着二代测序技术的推广和降价,WES 和 WGS 会逐渐代替基因 Panel 和单基因测序,发挥更强大的检测潜能和再分析优势,提高检测阳性率。

据报道,对疑似孟德尔遗传病患者进行临床 WES 测序的诊断率约为 25%^[35-39],部分患者可能由于以下情况导致假阴性结果:①原始数据通过软件与参考序列进行比对、数据过滤、变异注释等任意一个环节的纰漏都会造成致病突变的漏检^[40-41];②如果检测实验室无法获得疾病表型的关键要素,就会降低致病突变被优先考虑的可能性;③部分数据库纳入的基因型-表型、变异-表型资料不完善,需要通过搜索引擎进行特定检索^[42];④每年大约有 250 个新的基因和 9 200 个变异被证实与疾病关联,数据库在不断更新^[43]。因此,为了最大

限度地提高 WES 的诊断率,临床医生应对临床表型提供详细且全面的描述,供检测公司筛选目标基因。另外,还要对疾病有一定的文献知识储备和信息检索能力,并要求公司定期对检测结果阴性的病例进行数据再分析。

WGS 几乎可以检测所有基因组变异类型(如点突变、动态突变、拷贝数变异及深部内含子变异等),简化了基因评估过程,越来越多的证据表明它在临床诊断和基因发现中具有广泛的应用前景^[44]。据报道,WGS 在儿童遗传性疾病中的诊断率可达 41%,明显高于染色体微阵列分析、基因 Panel 等传统基因检测方法(24%; $P=0.01$)。与靶向基因检测技术相比,WGS 因其较高的诊断率甚至可作为首选检测方法^[45]。由于大多数 CNV 的断裂点位于内含子上,WES 发现的外显子缺失极有可能包含一大段内含子序列,所以 WES 分析 CNV 的准确性有限,但是 WGS 通过生物信息学分析更易发现这种形式的 CNV^[46-47]。如 WGS 发现一名智力障碍患者的 *TENM3* 基因在 4 号染色体上部分重复,该基因发生的 CNV 却不足以解释患者的临床表型,进一步生物信息学分析证明重复的 *TENM3* 基因发生易位插入位于 X 染色体的 *IQSEC2* 基因上,*IQSEC2* 基因的连续性受到破坏才得以解释患者智力障碍的原因^[48]。另有研究表明,在 50 个智力发育迟滞患者中,WGS 对 CNV 的检测率可达 18%,发现被染色体微阵列、MLPA 漏诊的 3 个外显子缺失^[49]。目前非扩增 WGS(PCR-free WGS)受基因组高 GC 含量区域的影响较小,整体覆盖度和均一性更好,尽管平均测序深度低但具有更为广泛的外显子覆盖度,随着测序价格的下降有望在临床得到广泛应用^[50]。

2 临床表型对基因检测技术起至关重要的作用

选择合适的基因检测技术,不仅要了解不同疾病致病基因的变异类型,更重要的是进行完整而全面的临床检查,从而确定临床表型。例如本研究团队曾在临床工作中遇到 1 名 35 岁男性患者,因“双手抖动伴行走不稳 10 年,近半年出现口齿不清”在外院诊断为“肝豆状核变性”,通过 WES 分析在 *ATP7B* 基因上发现一处杂合点突变 c.A2785G(p.I929V),但意义不明。重新对其进行临床评估后发现先证者还有癫痫病史,“双手抖动”实际为肌阵挛样表现,妹妹也有类似疾病史,但父母正常,查眼底镜并未发现 K-F 环,血生化未见异常,因此修正诊断为进行性肌阵挛癫痫(progressive myoclonic epilepsy,PME)。公司重分析 WES 数据后在唾液酸贮积症(sialidosis I)的致病基因 *NEU1* 上发现一处纯合突变 c.A544G(p.S182G),对先证者父母及妹妹进行共分离验证,结果显示父母均为杂合突变,妹妹和先证者一样均为纯合突变。由此可见,临床表型决定基因检测的方向,有助于提高基因检测的阳性率,从而降低成本并缩短诊断时间。

3 基因检测结果的解读

研究表明,每个人基因组中存在数百个功能缺失变异(loss-of-function variants, LOF)和数千个 VUS^[51-52],因此每例 WES 通过数据过滤仍有 150~500 个非同义突变或剪接突变等^[4],健康个体也会携带 40~110 种人类基因变异数据库(human gene mutation database, HGMD)中被判定为致病突变的变异^[53],导致真正的致病突变与罕见非致病性变异之间的界限变得难以捉摸。WES 或 WGS 在提高诊断阳性率的同时相比基因 Panel 产生了更多的 VUS^[5],为实验室筛选变异的优先级带来了巨大挑战。目前对基因检测结果中发现的变异进行致病性等级评判时,基因检测实验室主要遵循美国医学遗传学与基因组学学会(The American College of Medical Geneticists, ACMG)制定的序列变异解读指南^[54]。

有时会在患有孟德尔遗传病的患者中发现某个致病基因上存在罕见变异,但它却不足以解释患者的临床表型。文献报道过 1 例智力障碍、神经发育迟滞、小头畸形伴脑裂畸形的婴儿,其父母为近亲婚配,妹妹类似表型但无脑裂畸形。WES 在 *MCPH1* 基因上发现 1 处纯合移码突变 c.421_422insA(p.Q141fs),但该位点不能解释 2 个患儿表型上的差异。之后通过再分析,在先证者中又发现基因 *ALG8* 和 *CLN5* 各有 1 处纯合突变,它们与脑结构发育异常相关,从而更好地解释了先证者的临床表型^[55]。基因背景中的其他罕见变异位点对患者的临床表型也起着至关重要的作用,家系内共分离时常遇到与先证者携带相同突变的人并没有出现临床症状,研究表明除了基因外显率不全,也可能是先证者因携带更多其他基因上的罕见变异造成的突变负荷(mutational burden)较高所致^[56]。另外,本研究团队曾遇到 1 例 CADASIL 患者,WES 检测报告提示在 *NOTCH3* 基因上存在 1 对复合杂合变异 c.2656C>T(p.R886C)和 c.6202G>A(p.G2068R),查文献可知只有位于表皮生长因子样蛋白(epidermal growth factor-like protein, EGFr)功能区改变半胱氨酸蛋白数目的突变(即 c.2656C>T)才具有较高致病性^[57],所以和疾病相关基因上的变异还需通过文献、共分离验证或生物学实验等进一步明确其致病性。因此,临床医生对基因检测结果的解读不能完全依赖实验室分析,需要有一定的遗传学背景,对检测报告进行分析甄别,为患者提供更准确的遗传咨询。

4 讨论

随着高通量测序时代的到来,临床医生对神经系统遗传病的诊断需求也在不断提升,选择合适的基因检测技术有助于明确诊断,减少患者不必要且痛苦的就医过程和各种多余而昂贵的临床检查^[58];明确诊断后,临床医生可以帮助患者

制定个性化的治疗方案,对家中尚未发病的亲属提早进行干预、指导就业、规划生活等。对有生育要求的年轻患者,基因检测结果可作为产前诊断的依据,阻断疾病代代相传。

如今,二代测序技术的广泛应用使神经系统遗传病得到迅速发展,许许多多的致病基因被鉴定发现,临床表型也被进一步细分,但目前的基因检测还面临众多挑战。首先,不同实验室的测序平台对基因的覆盖度差别较大,数据分析跟不上测序技术的发展且缺乏一致性的共识^[59]:①不是所有的目标序列都有较高的覆盖度,若测序芯片恰好在某个致病基因上的覆盖度较低,则影响致病突变的有效检出。②变异致病性预测软件(如 Mutation Taster、PolyPhen-2)准确性并非 100%,有时会将致病性突变预测为良性或多态。③一些致病突变频率在某个人群中较高,按照默认的参考突变频率会被过滤(如遗传性血色素沉着症的致病基因 *HFE*,该基因上的致病突变 p.Cys282Tyr 在北美人群中频率高达 11%,若不考虑人群,只按频率>5%的变异都被过滤的规则将会漏诊)^[3,33],所以基因检测技术有时会因为测序平台和分析的不完善而漏掉真正的致病基因。其次,遗传异质性使基因型-表型相关性研究更为复杂^[60]:①携带致病基因的个体可能不表现出相应的疾病表型,即基因外显不全,在显性遗传模式的致病基因中更常见。因此,有些显性遗传模式疾病的家系内会出现隔代遗传(携带致病基因的下一代未表现出临床症状,即无症状携带者),对明确家系遗传方式带来了困难(尤其当家系较小时,更容易出现隔代遗传的现象)。②携带相同致病基因的患者会表现出不同程度的临床表型,主要与动态突变的重复次数、同一个基因的不同突变类型等有关,所以对基因上的突变类型进行功能研究对基因型-表型相关性分析至关重要。③一个基因同时会影响 2 个或多个不同表型特征(即基因多效性),比如, *VCP* 基因编码的蛋白质参与多种细胞活动,与包涵体肌病、佩吉特骨病和额颞叶痴呆等多系统疾病有关^[61]。因此,与临床表型相关基因进行优先级排序时,基因多效性使其变得尤为棘手。最后,基因检测结果的分析和解读也存在诸多问题,比如,对于基因检测中意外发现的已知明确致病性突变或者预期会导致疾病的突变,是否有必要告知患者等。

此外,要知道基因检测是帮助临床医生明确临床诊断的一种辅助工具,过度滥用不仅不会提高诊断率还将给患者带来不必要的心理负担和经济负担。

参 考 文 献

- [1] Li J, Parker B, Martyn C, et al. The PMP22 gene and its related diseases[J]. Mol Neurobiol, 2013, 47(2): 673-698.
- [2] Wijemanne S, Jankovic J. Dopa-responsive dystonia—clinical and genetic heterogeneity[J]. Nat Rev Neurol, 2015, 11(7): 414-424.

- [3] Fogel BL. Genetic and genomic testing for neurologic disease in clinical practice[J]. *Handb Clin Neurol*, 2018, 147(1): 11–22.
- [4] Xue Y, Ankala A, Wilcox WR, et al. Solving the molecular diagnostic testing conundrum for Mendelian disorders in the era of next-generation sequencing: single-gene, gene panel, or exome/genome sequencing[J]. *Genet Med*, 2015, 17(6): 444–451.
- [5] Katsanis SH, Katsanis N. Molecular genetic testing and the future of clinical genomics[J]. *Nat Rev Genet*, 2013, 14(6): 415–426.
- [6] Chang IJ, Hahn SH. The genetics of Wilson disease[J]. *Handb Clin Neurol*, 2017, 142: 19–34.
- [7] Meneret A, Gaudebout C, Riant F, et al. PRRT2 mutations and paroxysmal disorders[J]. *Eur J Neurol*, 2013, 20(6): 872–878.
- [8] Huang XJ, Wang SG, Guo XN, et al. The phenotypic and genetic spectrum of paroxysmal kinesigenic dyskinesia in China[J]. *Mov Disord*, 2020, 35(8): 1428–1437.
- [9] Pipis M, Rossor AM, Laura M, et al. Next-generation sequencing in Charcot–Marie–Tooth disease: opportunities and challenges[J]. *Nat Rev Neurol*, 2019, 15(11): 644–656.
- [10] Rossor AM, Polke JM, Houlden H, et al. Clinical implications of genetic advances in Charcot–Marie–Tooth disease[J]. *Nat Rev Neurol*, 2013, 9(10): 562–571.
- [11] Liu X, Duan X, Zhang Y, et al. Clinical and genetic diversity of PMP22 mutations in a large cohort of Chinese patients with Charcot–Marie–Tooth disease[J]. *Front Neurol*, 2020, 11: 630.
- [12] Yiu EM, Kornberg AJ. Duchenne muscular dystrophy[J]. *J Paediatr Child Health*, 2015, 51(8): 759–764.
- [13] Zhu SY, Xiong F, Chen YJ, et al. Molecular characterization of SMN copy number derived from carrier screening and from core families with SMA in a Chinese population[J]. *Eur J Hum Genet*, 2010, 18(9): 978–984.
- [14] Tesson C, Koht J, Stevanin G. Delving into the complexity of hereditary spastic paraplegias: how unexpected phenotypes and inheritance modes are revolutionizing their nosology[J]. *Hum Genet*, 2015, 134(6): 511–538.
- [15] Verkerk AJ, Pieretti M, Sutcliffe JS, et al. Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome[J]. *Cell*, 1991, 65(5): 905–914.
- [16] La Spada AR, Wilson EM, Lubahn DB, et al. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy[J]. *Nature*, 1991, 352(6330): 77–79.
- [17] Read AP. Huntington's disease: testing the test[J]. *Nat Genet*, 1993, 4(4): 329–330.
- [18] Brook JD, McCurrach ME, Harley HG, et al. Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member[J]. *Cell*, 1992, 69(2): 385.
- [19] Liquori CL, Ricker K, Moseley ML, et al. Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9[J]. *Science*, 2001, 293(5531): 864–867.
- [20] Ishiura H, Doi K, Mitsui J, et al. Expansions of intronic TTTCA and TTTTA repeats in benign adult familial myoclonic epilepsy[J]. *Nat Genet*, 2018, 50(4): 581–590.
- [21] Cen Z, Jiang Z, Chen Y, et al. Intronic pentanucleotide TTTCA repeat insertion in the SAMD12 gene causes familial cortical myoclonic tremor with epilepsy type 1[J]. *Brain*, 2018, 141(8): 2280–2288.
- [22] Corbett MA, Kroes T, Veneziano L, et al. Intronic ATTTTC repeat expansions in STARD7 in familial adult myoclonic epilepsy linked to chromosome 2[J]. *Nat Commun*, 2019, 10(1): 4920.
- [23] Florian RT, Kraft F, Leitao E, et al. Unstable TTTTA/TTTCA expansions in MARCH6 are associated with familial adult myoclonic epilepsy type 3[J]. *Nat Commun*, 2019, 10(1): 4919.
- [24] Yeetong P, Pongpanich M, Srichomthong C, et al. TTTCA repeat insertions in an intron of YEATS2 in benign adult familial myoclonic epilepsy type 4[J]. *Brain*, 2019, 142(11): 3360–3366.
- [25] Cortese A, Simone R, Sullivan R, et al. Biallelic expansion of an intronic repeat in RFC1 is a common cause of late-onset ataxia[J]. *Nat Genet*, 2019, 51(4): 649–658.
- [26] Tian Y, Wang JL, Huang W, et al. Expansion of human-specific GGC repeat in neuronal intranuclear inclusion disease-related disorders[J]. *Am J Hum Genet*, 2019, 105(1): 166–176.
- [27] Ma D, Tan YJ, Ng ASL, et al. Association of NOTCH2NLC repeat expansions with Parkinson disease[J]. *JAMA Neurol*, 2020, 77(12): 1–5.
- [28] Jiao B, Zhou L, Zhou Y, et al. Identification of expanded repeats in NOTCH2NLC in neurodegenerative dementias[J]. *Neurobiol Aging*, 2020, 89: 142e1–142e7.
- [29] Yuan Y, Liu Z, Hou X, et al. Identification of GGC repeat expansion in the NOTCH2NLC gene in amyotrophic lateral sclerosis[J]. *Neurology*, 2020, 95(24): e3394–e3405.
- [30] Fang P, Yu Y, Yao S, et al. Repeat expansion scanning of the NOTCH2NLC gene in patients with multiple system atrophy[J]. *Ann Clin Transl Neurol*, 2020, 7(4): 517–526.
- [31] Ng ASL, Lim WK, Xu Z, et al. NOTCH2NLC GGC repeat expansions are associated with sporadic essential tremor: variable disease expressivity on long-term follow-up[J]. *Ann Neurol*, 2020, 88(3): 614–618.
- [32] Deng J, Yu J, Li P, et al. Expansion of GGC repeat in GIPC1 is associated with oculopharyngodistal myopathy[J]. *Am J Hum Genet*, 2020, 106(6): 793–804.
- [33] Adams DR, Eng CM. Next-generation sequencing to diagnose suspected genetic disorders[J]. *N Engl J Med*, 2018, 379(14): 1353–1362.
- [34] Rim JH, Kim SH, Hwang IS, et al. Efficient strategy for the molecular diagnosis of intractable early-onset epilepsy using targeted gene sequencing[J]. *BMC Med Genomics*, 2018, 11(1): 6.

- [35] Yang Y, Muzny DM, Xia F, et al. Molecular findings among patients referred for clinical whole-exome sequencing[J]. *JAMA*, 2014, 312(18):1870–1879.
- [36] Yang Y, Muzny DM, Reid JG, et al. Clinical whole-exome sequencing for the diagnosis of mendelian disorders[J]. *N Engl J Med*, 2013, 369(16):1502–1511.
- [37] Lee H, Deignan JL, Dorrani N, et al. Clinical exome sequencing for genetic identification of rare Mendelian disorders[J]. *JAMA*, 2014, 312(18):1880–1887.
- [38] Iglesias A, Anyane-Yeboah K, Wynn J, et al. The usefulness of whole-exome sequencing in routine clinical practice[J]. *Genet Med*, 2014, 16(12):922–931.
- [39] Retterer K, Juusola J, Cho MT, et al. Clinical application of whole-exome sequencing across clinical indications[J]. *Genet Med*, 2016, 18(7):696–704.
- [40] Cornish A, Guda C. A comparison of variant calling pipelines using genome in a bottle as a reference[J]. *Biomed Res Int*, 2015, 2015:456479.
- [41] McCarthy DJ, Humburg P, Kanapin A, et al. Choice of transcripts and software has a large effect on variant annotation[J]. *Genome Med*, 2014, 6(3):26.
- [42] Baumgartner WA Jr, Cohen KB, Fox LM, et al. Manual curation is not sufficient for annotation of genomic databases[J]. *Bioinformatics*, 2007, 23(13):i41–i48.
- [43] Wenger AM, Guturu H, Bernstein JA, et al. Systematic reanalysis of clinical exome data yields additional diagnoses; implications for providers[J]. *Genet Med*, 2017, 19(2):209–214.
- [44] Stavropoulos DJ, Merico D, Jobling R, et al. Whole genome sequencing expands diagnostic utility and improves clinical management in pediatric medicine[J]. *NPJ Genom Med*, 2016, 1:15012.
- [45] Lionel AC, Costain G, Monfared N, et al. Improved diagnostic yield compared with targeted gene sequencing panels suggests a role for whole-genome sequencing as a first-tier genetic test[J]. *Genet Med*, 2018, 20(4):435–443.
- [46] Campbell IM, Gambin T, Dittwald P, et al. Human endogenous retroviral elements promote genome instability via non-allelic homologous recombination[J]. *BMC Biol*, 2014, 12:74.
- [47] Conrad DF, Bird C, Blackburne B, et al. Mutation spectrum revealed by breakpoint sequencing of human germline CNVs[J]. *Nat Genet*, 2010, 42(5):385–391.
- [48] Hehir-Kwa JY, Pfundt R, Veltman JA. Exome sequencing and whole genome sequencing for the detection of copy number variation[J]. *Expert Rev Mol Diagn*, 2015, 15(8):1023–1032.
- [49] Gilissen C, Hehir-Kwa JY, Thung DT, et al. Genome sequencing identifies major causes of severe intellectual disability[J]. *Nature*, 2014, 511(7509):344–347.
- [50] Meienberg J, Bruggmann R, Oexle K, et al. Clinical sequencing: is WGS the better WES?[J]. *Hum Genet*, 2016, 135(3):359–362.
- [51] MacArthur DG, Tyler-Smith C. Loss-of-function variants in the genomes of healthy humans[J]. *Hum Mol Genet*, 2010, 19(R2):R125–R130.
- [52] MacArthur DG, Balasubramanian S, Frankish A, et al. A systematic survey of loss-of-function variants in human protein-coding genes[J]. *Science*, 2012, 335(6070):823–828.
- [53] Xue Y, Chen Y, Ayub Q, et al. Deleterious- and disease-allele prevalence in healthy individuals; insights from current predictions, mutation databases, and population-scale resequencing[J]. *Am J Hum Genet*, 2012, 91(6):1022–1032.
- [54] Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants; a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology[J]. *Genet Med*, 2015, 17(5):405–424.
- [55] Karaca E, Posey JE, Coban Akdemir Z, et al. Phenotypic expansion illuminates multilocus pathogenic variation[J]. *Genet Med*, 2018, 20(12):1528–1537.
- [56] Pizzo L, Jensen M, Polyak A, et al. Rare variants in the genetic background modulate cognitive and developmental phenotypes in individuals carrying disease-associated variants[J]. *Genet Med*, 2019, 21(4):816–825.
- [57] Rutten JW, Haan J, Terwindt GM, et al. Interpretation of NOTCH3 mutations in the diagnosis of CADASIL[J]. *Expert Rev Mol Diagn*, 2014, 14(5):593–603.
- [58] Di Fonzo A, Monfrini E, Erro R. Genetics of movement disorders and the practicing clinician; who and what to test for?[J]. *Curr Neurol Neurosci Rep*, 2018, 18(7):37.
- [59] Gotway G, Crossley E, Kozlitina J, et al. Clinical exome studies have inconsistent coverage[J]. *Clin Chem*, 2020, 66(1):199–206.
- [60] Magrinelli F, Balint B, Bhatia KP. Challenges in clinicogenetic correlations: one gene-many phenotypes[J]. *Mov Disord Clin Pract*, 2021, 8(3):299–310.
- [61] Watts GDJ, Wymer J, Kovach MJ, et al. Inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia is caused by mutant valosin-containing protein[J]. *Nat Genet*, 2004, 36(4):377–381.

(责任编辑:冉明会)