

精准医学在儿外科出生缺陷疾病谱中的应用

张彦 夏慧敏

精准医学是以个性化医疗为基础,随着全基因组测序技术及生物信息学的交叉应用而发展起来的新兴医学模式。复杂遗传病的不同临床表型都是个体基因组、转录组或表观遗传学因素等综合作用的结果。精准医学根据导致疾病的潜在遗传因子对疾病进行诊断和分类,而非传统的根据症状和体征对疾病分类的方法,从而出现了罹患同种疾病的“个性化差异”模式。精准医学的重要内容是基因的检测,通过先进的基因检测技术,如全基因组测序、外显子测序和转录组测序等,发现与疾病发生相关的致病基因。因此,基因测序是精准医学的基础和灵魂。

一、DNA 测序技术的发展及新一代测序技术的兴起

脱氧核糖核酸 (deoxyribonucleic acid, DNA) 是遗传信息的基本载体,决定了生命的延续、个体生存,甚至物种的进化。新一代测序技术 (Next Generation Sequencing, NGS) 应运而生,高通量,测序速度快,成本低,生成海量数据。这些优势使得 NGS 近年来被广泛运用在基础生物医学研究领域,对一种疾病不同状态和过程进行精确亚分类,精确寻找疾病的原因和治疗的靶点,并为解释疾病的发病机制开辟了新路径。伴随测序技术的发展,测序技术的突破不仅推动了生物学的发展,还使得那些在基础生物学研究领域取得的成果,能够成功运用于临床,并将在疾病的预防控制,快速临床诊断,个性化治疗等方面起到重要的作用。

二、儿外科常见出生缺陷精准医疗的应用进展我国在出生缺陷的诊疗研究上有了一定的进展。近年来迅速发展起来的基因诊断技术,在出生缺陷与遗传病检测领域获得广泛应用。

1. 先天性心脏病:先天性心脏病 (congenital heart disease, CHD) 是引起我国新生儿死亡的主要原因之一^[1]。其病因复杂,绝大多数是遗传因素和

环境因素共同作用的结果。部分 CHD 临床亚型与染色体异常有关,如 22q11 缺失、21 号染色体三倍型等^[2]。此外,迄今为止超过 50 个基因与 CHD 密切相关。CHD 相关突变大部分落在一小段发育基因 (例如 NKX2-5, GATA4, 和 NOTCH1)^[2]。小鼠模型研究表明,超过 500 种基因可诱发心脏缺陷。可想而知人类 CHD 致病基因广泛存在。

我国目前对胎儿染色体非整倍体异常筛查的技术成熟度较高^[3],仍有较大部分 CHD 病人无法做到产前诊断。随着测序技术的飞速发展,从全基因组遗传变异层面,国内外在 CHD 遗传致病位点有大量前期研究,如 MTHFR c. 677 C > T、c. 1298 A > C、MTRR c. 66 A > G 等被证实^[4]。Zaidi 等^[5]利用 362 例复杂 CHD 病人和 264 例对照,通过全基因组测序技术证实了一系列与 His3Lys4 甲基化密切相关的新发突变。Turki 等^[6]利用外显子测序检测了 13 个胎儿心脏结构异常 (AVSD) 核心家系和 112 个散发对照,证实了高保守性 NR2F2 多个罕见错义突变,小鼠模型证实 NR2F2 与心脏发育过程的动脉生成有关;后续收集 CHD 家系证实了 NR2F2 区域更多的结构变异。尽管有上述各项成功的案例,并不是所有研究都能成功找到致病基因。Arrington 等^[7]采集有明确的常染色体显性遗传模式的 11 个不同 CHD 临床表型患者的四代家系,利用全外显子组测序技术挖掘其致病基因;然而出人意料的是,不同临床亚型的家族成员并未确认共同的致病位点。Martin 等^[8]利用全外显子测序技术检测共 17 人的先天性两叶型主动脉瓣 (BAV) 大家系,其中 BAV 病例 3 例,其他心血管畸形 3 例;然而并未有明确的致病变异确证。

上述案例显示仅有部分病例符合孟德尔遗传规律。遗传变异的频率和变异对疾病表型的贡献度或外显率成正比,高频率的遗传变异的效应值 (odds ratio, OR) 较低,因此对表型的贡献度或外显率 (penetrance) 较小;低频率的遗传变异的效应值较高,因此对疾病的贡献度或外显率较大^[10]。因此,国内外研究人员一致认为除致病突变以外,针对不

同 CHD 临床亚型存在一系列效应值中等的易感变异。2013 年 Cordell 等^[9]利用全基因组基因分型技术,基于 835 例法洛四联症患者和 5 159 例健康对照证实 12q24 和 13q32 为遗传易感基因;同年, Cordell 等^[10]利用相同的技术,基于 1 995 例复杂 CHD 患者(主要涵盖中隔、梗阻性和紫绀型缺陷)挖掘遗传易感基因,遗憾的是当各种临床亚型混在一起,并未有任何易感基因证实;而进一步按临床亚型聚类分析研究确证 4p16 与继发孔型房间隔缺损(ASD)易感,且在后续 417 例 ASD 病人和 2 520 例对照中进一步得到验证。

上述多个研究表明 CHD 不同临床表型的遗传机制确有不同,在后续研究中需进一步细化临床表型,相互佐证又相互区别地对待才能从中抽丝剥茧,确认真正的致病遗传机制。故而在后续临床转化研究中,应以临床标本为本,进一步细化研究。

本中心侧重于收集有明显临床特征的复杂型且预后较差的 CHD 病患,拟建立长期随访队列,对不同临床亚型的 CHD 进行聚类分析,利用 NGS,包括全基因组甲基化测序、转录组测序等挖掘致病因素和潜在治疗的靶点;并进一步根据预后个体差异,建立不同的预警、治疗方案对疾病进行有效诊断和干预。本中心科研人员前期利用新生儿心胸外科部分复杂 CHD 患者的共同临床特征:胸腺切除展开研究,回顾胸腺切除后对患儿免疫系统的影响,对术后不同时间段的病人进行外周血免疫细胞分选,选择恢复期特异富集免疫细胞,利用全基因组甲基化测序技术确定 CpG 点甲基化与基因表达调控的关系,拟建立疾病特有的甲基化图谱或生物标记物。并进一步在免疫细胞衰老程度、胸腺切除病人特异性标记物筛查等方向展开多方面探讨。希望能从契合临床的视角为 CHD 的研究添砖加瓦。

2. 消化道畸形:消化道畸形是新生儿期常见的畸形之一,常见病种包括:食管闭锁、肥厚性幽门狭窄、肠闭锁、先天性巨结肠、肛门闭锁、胆总管囊肿、胆道闭锁等。国外统计其发生率约为 3‰,国内新生儿科统计约占同期住院新生儿的 1‰。在临床诊疗技术得到大力发展的同时,科学家们仍继续在基础研究方面探索该类疾病的病因及发病机制,并寻求更先进可靠的诊疗方法。

以先天性巨结肠(Hirschsprung disease, HSCR)为例,其是新生儿期低位肠梗阻常见原因之一。普遍认为遗传因素在先天性巨结肠的发生发展中起着关键的作用。目前发现 RET 内多个位点突变可以

解释约 30%~50% 的家族性和 20% 的散发性先天性巨结肠发病原因。然而,目前发现先天性巨结肠的突变位点在不同家系中极为分散,无热点区域存在,极少数为大片段缺失,多为点突变,错义,无义,移码突变^[11-17]。已知 RET 有 4 种配体,分别是胶质细胞系源神经 营养因子(GDNF)、神经营养蛋白(NRTN)、artemin(ARTN)、persephin(PSPN),它们与 GDNF 家族受体结合后与 RET 结合形成复合物,激活 RET 信号通路。该通路对于肠神经系统的分化、发育和成熟至关重要^[18]。GDNF 突变常见于综合征型先天性巨结肠,突变率为 0.9%~5.5%。目前已经检测到多个功能突变与先天性巨结肠相关^[19]。NRTN 亦发现有限的功能突变与先天性巨结肠相关^[20,21]。

然而上述配体发现均存在外显率较低的现象,且往往与 RET 的功能突变分离。上述突变可能与 RET 的突变存在相互作用,能够影响先天性巨结肠的临床症状,而与其发病本身相关性较低。

虽然 RET 信号通路被广泛认为是与先天性巨结肠发病最为密切的通路,然而由于先天性巨结肠发病率较低,已知致病突变多来源于独立家系报道,针对该通路中的致病遗传突变尚未有系统性的群体研究,缺乏系统性的筛查总结指导临床诊断、治疗。随着高通量基因分型技术的广泛应用,全基因组关联分析(GWAS)被逐渐运用在包括先天性巨结肠在内的各种复杂疾病易感基因的发掘。2008 年 Garcia-Barcelo 等首次证实 NRG1 为先天性巨结肠的易感基因^[22];2012 年 Tang 等发现 NRG3 中 11 个新发结构突变被证实与先天性巨结肠相关^[23]。证实在先天性巨结肠散发病例存在较家系中发现的罕见突变而言外显率较低的易感基因。2014 年, Kim 等又提出了 SLC6A20、RPRA、ABCC9 等多个潜在易感基因^[24]。2016 年,相同的研究团队基于 SLC6A20 的潜在易感位点进行精细定位提出 SLC6A20 多个潜在易感位点^[25]。

虽然分子遗传学取得了如上述巨大进展,但仍不能完全解释先天性巨结肠的起因,术后提示尚有多种未发现的遗传因素。特别是基于不同人群中的群体遗传易感机制筛查,寥寥无几,故而有待后续挖掘。但具体如何利用致病基因指导临床诊断、治疗,尚未有定论。本中心儿外科科研人员基于已知、新证、潜在先天性巨结肠遗传致病基因或区域,利用目标区域测序技术构建遗传诊断数据库;针对已知的致病突变携带者,我们准备根据其所携带的致病突

变分类类型进而进行更深入的致病基因遗传机制研究;针对携带未知致病突变的样本,我们将在后续挖掘全新致病基因中进一步筛选。此外,基于单中心收集巨结肠家系样本组,利用全基因组测序技术挖掘新发突变、罕见突变,证实全新先天性巨结肠遗传基因。基于上述工作,围绕已知致病基因和全新挖掘出的致病基因,拟制备先天性巨结肠遗传诊断芯片,为临床诊断提供更为有效的证据,并为针对先天性巨结肠病人的个性化医疗提供助益。

三、出生缺陷队列的建立及生物标本的存储

诚然,基于先天性心脏病、先天性巨结肠本中心展开了一系列遗传机制的临床转化研究。但出生缺陷的预防、诊断和治疗依然面临诸多挑战。众所周知,出生缺陷的发生是一个动态的过程,在此期间发生的许多独立事件都可能是致病的协同作用。单纯针对某一时间点的临床样本进行后续研究可能会有失偏颇。而队列研究可观察多个因素的多种临床效应,这正是队列法不可取代的用途,为疾病个体的精准化医疗提供了更为详实的病程记录。本中心基于先天性心脏病、先天性巨结肠已有前期研究基础的疾病建立大样本出生缺陷队列。目前已建立前瞻性出生队列公共平台^[26],队列中健康人群的已逐步生成多组学数据,包括基因组、转录组、蛋白组、代谢组和肠道菌群的宏基因组数据,可以综合、立体、量化反映个体健康状态,也可以成为各类疾病研究无可替代的黄金对照。给疾病的发生发展提供了重要的溯源研究资源,也为系统地、多时点地描述疾病的发展提供了可能,同时通过将生物样本库中的组学信息与流行病学数据、电子病历系统进行整合匹配,为先天性心脏病、先天性巨结肠提供更细致的定义疾病分类,更深入的阐释发病原因和更准确的预测疾病风险或治疗效果。

而如此大规模各异生物标本的储存是所有临床研究的根本和重中之重。标本类型包括全血、血清、血浆、单个核细胞、疾病组织、粪便等。本中心前期建立了生物样品采集、处理、贮存和解冻的标准化操作规程,为本项目的生物样本采集、贮存和管理提供基础。对临床生物资源库的出入库流程、样本捐献者的准备、临床样本的采集与运输、处理过程进行控制,以保证样本的合理合法使用、样本质量的稳定性和有效性,更好地为科学研究提供准确、可靠的样本储备。本中心基于 5 年来出生队列的建立总结了如下几条执行标准,以资借鉴。①设立临床生物资源库伦理委员会和临床生物资源库科学技术委员会,

负责入库项目的审批。②医护人员和实验室人员指导样本捐献者如何正确留取样本。③门诊采血人员和病房护理人员负责临床样本的采集,特殊样本由临床医生采集。④样本运送人员负责按时按标准运送样本。⑤临床生物资源库工作人员负责核收样本捐献人员签署的《知情同意书》并扫描存档,定期检查,如发现缺漏应及时与入库申请者沟通反馈。⑥临床生物资源库工作人员负责样本核收和出入库。

此外,出生缺陷队列的临床资料采集也是后续大数据分析重要的参考因素;而临床数据采集主要通过本中心的临床大数据中心集成分散存储在不同诊疗系统中的临床表型、实验室检测、影像学检查、诊疗记录数据的迁移集成,清洗转换,元数据抽取,实现异构整合;利用海量数据统一存储模型,搭建高性能、高可靠、易扩展、易使用的专病主题数据仓库,作为后续研究的数据库基础。基于出生缺陷专病主题数据仓库及临床指南、专家共识、标准等已有的知识型文本,分别通过文本实体识别及语义提取、生物医学数据挖掘技术成功构建出生缺陷医学知识网络和知识图谱;并通过进一步集成利用深层索引、相关性挖掘、重要性标注、新颖度分析等挖掘工具,实现了出生缺陷精准医学知识的抽提、注释、聚类、关联及分析,构建出生缺陷相关精准医学知识库。有助于更深入地诠释重大出生缺陷发生的遗传背景,阐述基因调控和功能改变在重大出生缺陷发生发展中的作用,加速认识遗传、环境、遗传-环境交互作用对重大出生缺陷发生的关键作用,从而采取有针对性的预防措施,为出生缺陷防控尤其是出生缺陷一级预防提供更新的思路。

四、展望

精准医疗是以个体化医疗为基础,伴随着基因组测序技术快速进步,以及生物信息与大数据科学的交叉应用而发展起来的。近年来,二代测序技术由于其快速发展和成本的降低,正在占据越来越多的市场份额。通过新一代测序生成不同层面的遗传数据、结合临床表型对解析不同出生缺陷个体都具有重要作用,它将大大促进我们对出生缺陷疾病谱的遗传发病机制的理解。对于每一个患者家庭也是福音,它将能以每一个家庭为单位,鉴定其家族性的致病基因,为优生优育提供导向。而在人群层面上,对于出生缺陷疾病谱的进一步理解,能够切实指导临床治疗,为特异的靶向药物提供可能。然而以 CHD 为例,仅考虑遗传因素,仅能解释 8% 左右的发病原因;大部分原因是在一定的遗传背景下结合环

境因素最终致病。故而发病机制的复杂性和不确定性限制了对其风险预测、以及临床轻重症预警及个体精准化治疗的实施。当前急需对 CHD 特征进行系统、长期跟踪,利用母婴和出生队列寻找与疾病发病风险以及个体差异相关的遗传、表观遗传和环境影响因素,并通过形成预防和诊疗指南的形式,提高新生儿和儿童的生活质量,节约社会医疗成本。

此外,完善的队列后续追踪研究可以为疾病预后整合个体的遗传数据提供指南。以先天性巨结肠为例,手术治疗虽能挽救患者的生命,但术后长期存在明显的并发症。超过 50% 的短段型先天性巨结肠患者术后长期存在便秘、大便失禁和小肠结肠炎等。长段型的患者长期预后则更差,大便失禁、反复发作的小肠结肠炎、代谢紊乱、发育迟缓等发生率更高,约 20% 需要行永久性回肠造瘘术。我们前期已基于遗传数据证实术后肠炎特异性遗传变异(未发表数据),基于队列的长期随访,相信能够对先天性巨结肠术后恢复的临床指导有部分助益;能为不同临床表型提出遗传标记物,乃至靶向药物。总而言之,精准医学要求既要了解疾病的本质,即遗传变异,又要在结合不同个体生活环境、自然属性的基础上,实现个体化及疾病针对性的治疗。这一切都依赖于人类基因组不同层面上的研究。在过去的十年间,疾病个体的遗传背景研究的重要性越来越受到人们重视,相信在未来,对于疾病本身的精准医学解析一定能为医生提供更好的预防、诊断和治疗依据。

参考文献

- 1 盛伟,黄国英. 先天性心脏病分子遗传学精准医学研究[J]. Chinese Journal of Practical Pediatrics, 2016, 31 (8): 568-571. DOI:10. 7504/ek2016080603.
Sheng W, Huang GY. Application of precision medicine in the study of molecular genetics of congenital heart disease mutations and phenotype of Hirschsprung's disease[J]. Chinese Journal of Practical Pediatrics, 2016, 31 (8): 568 - 571. DOI:10. 7504/ek2016080603.
- 2 Swerdlow DI, Humphries SE. Genetics of CHD in 2016: Common and rare genetic variants and risk of CHD[J]. at Rev Cardiol, 2017, 14(2): 73-74. DOI:10. 1038/ nrcardio. 2016. 209
- 3 Trevisan P, Zen TD, Rosa RF, et al. Chromosomal abnormalities in patients with congenital heart disease[J]. Arq Bras Cardiol, 2013, 101 (6): 495 - 501. DOI: 10. 5935/ abc. 20130204.
- 4 Xuan C, Li H, Zhao JX, et al. Association between MTHFR

- polymorphisms and congenital heart disease: a meta-analysis based on 9,329 cases and 15,076 controls[J]. Sci Rep, 2014, 4: 7311. DOI:10. 1038/srep07311.
- 5 Zaidi S, Choi M, Wakimoto H, et al. De novo mutations in histone-modifying genes in congenital heart disease [J]. Nature, 2013, 498 (7453): 220 - 223. DOI: 10. 1038/ nature12141.
- 6 Al Turki S, Manickaraj AK, Mercer CL, et al. Rare variants in NR2F2 cause congenital heart defects in humans [J]. Am J Hum Genet, 2014, 94 (4): 574-585. DOI:10. 1016/ j. ajhg. 2014. 03. 007.
- 7 Arrington CB, Bleyl SB, Matsunami N, et al. Exome analysis of a family with pleiotropic congenital heart disease [J]. Circ Cardiovasc Genet, 2012, 5 (2): 175 - 182. DOI: 10. 1161/CIRC GENETICS. 111. 961797.
- 8 Martin LJ, Pilipenko V, Kaufman KM, et al. Whole exome sequencing for familial bicuspid aortic valve identifies putative variants[J]. Circ Cardiovasc Genet, 2014, 7(5): 677-683. DOI:10. 1161/CIRCGENETICS. 114. 000526.
- 9 Cordell HJ, Topf A, Mamasoula C, et al. Genome-wide association study identifies loci on 12q24 and 13q32 associated with tetralogy of Fallot[J]. Hum Mol Genet, 2013, 22 (7): 1473-1481. DOI:10. 1093/hmg/dd552.
- 10 Cordell HJ, Bentham J, Topf A, et al. Genome-wide association study of multiple congenital heart disease phenotypes identifies a susceptibility locus for atrial septal defect at chromosome 4p16[J]. Nat Genet, 2013, 45(7): 822-824. DOI:10. 1038/ng. 2637.
- 11 Tou J, Wang L, Liu L, et al. Genetic variants in RET and risk of Hirschsprung's disease in Southeastern Chinese: a haplotype-based analysis[J]. BMC Med Genet, 2011, 12: 32. DOI:10. 1186/1471-2350-12-32.
- 12 Fitze G, Cramer J, Ziegler A, et al. Association between c135G/A genotype and RET proto-oncogene germline mutations and phenotype of Hirschsprung's disease[J]. Lancet, 2002, 359 (9313): 1200-1205. DOI:10. 1016/ S0140 - 6736(02)08218 - 1.
- 13 Vaclavikova E, Kavalcova L, Skaba R, et al. Hirschsprung's disease and medullary thyroid carcinoma: 15-year experience with molecular genetic screening of the RET proto-oncogene[J]. Pediatr Surg Int, 2012, 28(2): 123 - 128. DOI:10. 1007/s00383 - 011 - 2993 - 2.
- 14 Nunez-Torres R, Fernandez RM, Acosta MJ, et al. Comprehensive analysis of RET common and rare variants in a series of Spanish Hirschsprung patients confirms a synergistic effect of both kinds of events[J]. BMC Med Genet, 2011, 12: 138. DOI:10. 1186/1471-2350-12-138.
- 15 Ruiz-Ferrer M, Fernandez RM, Antinolo (下转第 331 页)