

利用目标基因捕获结合第二代测序技术发现 腹主动脉瘤致病基因突变

郭俊, 蔡伦, 李小燕, 王绿娅, 杜杰

(首都医科大学附属北京安贞医院 北京市心肺血管疾病研究所 心血管重塑相关疾病教育部重点实验室, 北京市 100029)

[关键词] 腹主动脉瘤; 原纤维蛋白 1; 目标基因捕获; 第二代测序

[摘要] **目的** 建立目标基因捕获结合第二代测序技术对腹主动脉瘤患者原纤维蛋白 1(FBN1) 基因进行突变筛查, 探讨腹主动脉瘤与 FBN1 基因突变的关系。**方法** 提取 4 例腹主动脉瘤患者外周血全基因组 DNA, 利用 GenCap 目标基因捕获技术, 设计 FBN1 的 65 个外显子区域特异性捕获探针, 与基因组 DNA 文库进行杂交, 将目标基因组区域的 DNA 片段进行富集后, 再利用 Illumina HiSeq2000 第二代测序仪进行测序, 通过数据分析, 确定突变位点, 用 Sanger 测序法对突变位点进行验证。**结果** 设计合成的目标基因特异性捕获探针可有效地捕捉并富集基因组 DNA 的目标靶片段。4 例患者目标区域平均测序深度为 448.15 ~ 536.61, 99.5% ~ 99.7% 目标区域覆盖度。经过数据分析及 Sanger 测序验证发现 1 个新的错义突变 c. 2753 C > G (p. Pro918Arg), dbSNP137 数据库、千人基因组及内部 800 名正常汉族人数据库均无此突变。经 SIFT 预测为有害突变。**结论** 本研究所建立的 GenCap 目标基因捕获技术结合 Illumina HiSeq2000 第二代测序技术成功地发现了 FBN1 的新突变。该方法快速而有效, 对腹主动脉瘤分子病因学有更好的认识。

[中图分类号] R5

[文献标识码] A

Targeted Gene Capture and Next-generation Sequencing in Abdominal Aortic Aneurysm Identifies Mutations in Disease-causing Gene

GUO Jun, CAI Lun, LI Xiao-Yan, WANG Lv-Ya, and DU Jie

(Beijing Anzhen Hospital Affiliated to Capital Medical University & Beijing Institute of Heart, Lung and Blood Vessel Diseases & Key Laboratory of Remodeling-related Cardiovascular Diseases of Ministry of Education, Beijing 100029, China)

[KEY WORDS] Abdominal Aortic Aneurysm; Fibrillin-1; Targeted Gene Capture; Next-generation Sequencing

[ABSTRACT] **Aim** To develop an approach based on targeted gene capture and next-generation sequencing to determine the genetic defects in patients with abdominal aortic aneurysm precisely and effectively and confirm the role of fibrillin-1 (FBN1) in the pathogenesis of abdominal aortic aneurysm. **Methods** We analyzed four abdominal aortic aneurysm patients. Peripheral blood was collected and genomic DNA was isolated. FBN1 were selected by a gene capture strategy, using GenCap custom enrichment kit. The enrichment libraries were sequenced on Illumina HiSeq 2000 sequencer for paired read 90 bp to determine the mutation frequency in FBN1. Variants which have been reported in dbSNP137 or 1000 genome and in-house database were filtered. Selected variants were further validated by PCR and Sanger sequencing. **Results** For the samples subjected to targeted next-generation sequencing, the average sequencing depths on the targeted regions were yielded from 448.15 to 536.61. Meanwhile, coverage of targeted exons were ranged from 99.5% to 99.7%. We identified one missense mutation in FBN1: c. 2753 C > G (p. Pro918Arg), which was validated by Sanger sequencing. **Conclusions** Our results confirm the role of FBN1 in the pathogenesis of abdominal aortic aneurysm and demonstrated the robustness of targeted next-generation sequencing to precisely and rapidly determine genetic defects.

[收稿日期] 2014-04-14

[修回日期] 2014-11-06

[基金项目] 国家自然科学基金资助项目(81200074); 科研基地建设-心血管重大疾病防治协同创新中心资助项目(PXM2013_014226_07_000088)

[作者简介] 郭俊, 博士, 助理研究员, 研究方向为心血管分子病因学, E-mail 为 guojunbjmu@hotmail.com。蔡伦, 博士, 助理研究员, 研究方向为分子生物学。通讯作者杜杰, 博士, 教授, 研究方向为心血管分子病因学, E-mail 为 jiedubj@126.com。

腹主动脉瘤(abdominal aortic aneurysm, AAA)是受遗传与环境因素共同影响的复杂性疾病^[1]。有报道腹主动脉瘤的发生与马凡综合征的致病基因原纤维蛋白1(fibrillin-1, FBN1)基因相关^[2]。FBN1基因由65个外显子组成,并且突变的位点随机分布在整个FBN1基因中,没有明显的热点突变^[3],传统的Sanger测序法测序通量较低,难以满足测序需求,需要高通量的方法来进行致病基因的检测^[4]。新发展起来的目标基因捕获联合二代测序技术可以对多个目标区域同时进行测序,成本也比全外显子测序低^[5]。本研究采用目标基因捕获结合第二代测序技术分析了4例腹主动脉瘤患者FBN1基因,发现其中1例患者存在1个FBN1新的错义突变。

1 对象和方法

1.1 研究对象

选取2012~2013年就诊于北京安贞医院且临床诊断为腹主动脉瘤的患者4例,所有患者均知情同意。男性3例,女性1例,汉族,年龄61~76岁。对每例患者均进行较为详细的体格检查,包括骨骼系统、心血管系统、眼等。本研究得到北京市安贞医院伦理委员会批准。

1.2 基因组DNA的提取

取腹主动脉瘤患者外周静脉血2 mL(EDTA抗凝),用QIAamp全血DNA提取试剂盒(德国Qiagen公司)提取基因组DNA。

1.3 液相捕获目标基因测序

应用GenCap液相捕获目标基因技术(北京迈基诺公司)对FBN1外显子区域进行捕获^[6,7]。针对FBN1 65个外显子非重复区设计生物素化的单链捕获探针。将1 μg DNA文库与BL缓冲液和设计的探针混合,95℃加热7 min,65℃加热2 min,加入23 mL预热至65℃的HY缓冲液,65℃杂交22 h。用500 μL 1×结合缓冲液清洗50 μL MyOne磁珠(美国Life Technology公司)3次,重悬于80 μL 1×结合

缓冲液。加入64 μL 2×结合缓冲液至杂交混合物中,转移至含有80 μL MyOne磁珠的试管中。旋转混匀1 h。用WB1缓冲液室温清洗磁珠15 min,用WB3缓冲液65℃清洗3次,每次15 min。然后用洗脱缓冲液将结合的DNA洗脱。洗脱的DNA进行PCR反应,反应条件:98℃预变性30 s,98℃变性25 s,65℃退火30 s,72℃延伸30 s,共进行15个循环,最后72℃延伸5 min。根据产品说明书用SPRI beads(美国Beckman Coulter)纯化PCR产物。富集的文库用Illumina HiSeq 2000第二代测序仪进行双端测序,读长为90 bp。

1.4 数据分析

目标区域测序后原始数据经过index拆分后获得各个样本的原始序列,去除测序数据中的接头和低质量数据(质量值≤20),运用BWA软件比对到参考基因组上,用GATK软件对各个样本的比对数据进行多态性位点检测,对单核苷酸多态性及其插入缺失突变等数据进行统计和分析。同时进行测序深度、均一性、探针特异性等数据进行统计和分析。去除dbSNP137数据库、千人基因组和内部800正常汉族人数据库中出现的单核苷酸多态性及其插入缺失突变。同时,过滤掉SIFT预测对蛋白功能无影响的突变位点作为最后疾病相关的突变位点。

1.5 Sanger测序验证

根据需要测序的DNA片段合成引物,用PCR进行扩增,用ABI 3730xl测序仪(美国Applied Biosystems公司)以Sanger测序法进行测序,测序结果与目标区域捕获测序后的结果进行比对。

2 结果

2.1 目标基因捕获结合第二代测序数据分析

4例患者目标区域平均测序深度为448.15~536.61,99.5%~99.7%目标区域覆盖度,98.2%~99.2%目标区域>4×覆盖度,95.4%~97.9%目标区域>10×覆盖度,90.2%~95.2%目标区域>20×覆盖度(表1)。

表 1. 目标基因捕获结合第二代测序质量
Table 1. Quality of the targeted next-generation sequencing

样本	比对到人类基因组比例	目标区域覆盖度比例	目标区域平均测序深度	目标区域>4×覆盖度比例	目标区域>10×覆盖度比例	目标区域>20×覆盖度比例
1	99.75%	99.5%	449.97	98.4%	95.5%	90.8%
2	99.81%	99.7%	536.61	99.1%	97.7%	95.0%
3	99.80%	99.5%	456.56	98.2%	95.4%	90.2%
4	99.77%	99.7%	448.15	99.2%	97.9%	95.2%

2.2 FBN1 突变分析

4 例患者中有 1 例患者 FBN1 基因存在 1 个错义突变, c. 2753 C > G, 由第 918 位原密码子 CCA 编码的脯氨酸突变为 CGA 编码的精氨酸; 设计引物扩增后进行 Sanger 测序, 测序结果相同(图 1)。db-SNP137 数据库、千人基因组和内部 800 正常汉族人数据库中均无此突变。上述基因突变经 FBN1 基因 SNP 数据库 (<http://www.omim.org/allelicVariant/134797> 和 <http://www.umd.be/FBN1>) 检索, 为 FBN1 新的突变。

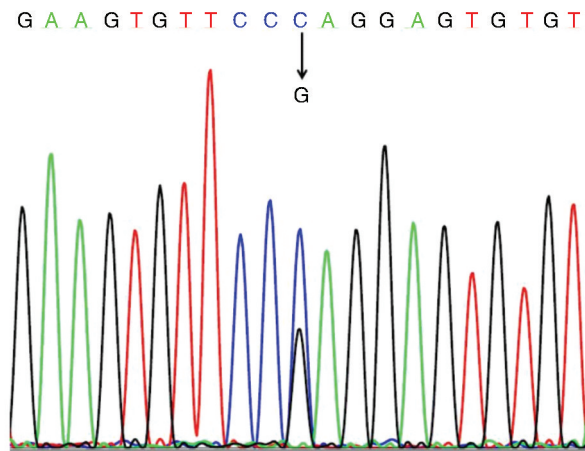


图 1. FBN1 基因突变位点 Sanger 测序图

Figure 1. Sanger sequencing analysis of FBN1 gene mutation

3 讨论

腹主动脉瘤的发生发展具有隐匿性, 仅在破裂前较短时间内或发生破裂后才出现症状, 并且一旦破裂具有极高死亡率。对腹主动脉瘤患者及早进行干预和治疗, 能够有效预防腹主动脉瘤破裂, 降低相关疾病的病死率。因为 FBN1 编码的外显子数量众多, 传统的对每个区域单独测序的方法不适用于临床。本研究中, 我们设计了针对 FBN1 的捕获芯片。利用 FBN1 外显子捕获结合第二代测序技术, 在 4 例腹主动脉瘤患者中发现 1 例患者携带 1 个 FBN1 新的错义突变。

由 FBN1 基因编码的 FBN1 是一种 350 kDa 的糖蛋白, 为 10 ~ 12 nm 结缔组织微纤维的主要组成部分, 广泛分布于弹性和非弹性组织中。FBN1 有 43 个含有钙结合的保守序列的 cb EGF 结构域, 结合钙可以限制结构域弹性, 易化蛋白质间的相互作用, 也保护蛋白免于水解, 对于维持 FBN1 的稳定十分重要。这些钙结合 EGF 样基序长链结构中

还穿插其它功能区域, 包括 7 个与转化生长因子 1 (transforming growth factor-1, TGF-1) 结合蛋白同源的结构及富含半胱氨酸, 由 EGF 样结构和 TGF-1 样结构融合而成的杂交区域^[8]。钙结合区域的突变及半胱氨酸的置换可引起 FBN1 功能障碍^[9]。研究表明, 发生在 cb EGF 结构域的错义突变造成的单个氨基酸的改变能破坏钙结合部位或结构域的二级结构^[10], 还有的错义突变可以改变多肽的糖基化或蛋白水解位点, 影响到蛋白的正确折叠和稳定性的维持, 继而影响 FBN1 的运输, 延迟 FBN1 的分泌和增加 FBN1 对蛋白酶的敏感性^[8]。

与传统的 Sanger 测序相比, 目标基因捕获结合第二代测序技术快速、准确、高通量, 可以用在大样本量的疾病已知基因检测中, 使我们对腹主动脉瘤分子病因学有更好的认识。这种方法很可能变成诊断罕见单基因疾病的标准方法, 也为发现基因变异导致常见疾病的风险提供了重要的指导。

[参考文献]

- [1] Golledge J, Kuivaniemi H. Genetics of abdominal aortic aneurysm [J]. Curr Opin Cardiol, 2013, 28 (3): 290-296.
- [2] Biros E, Walker PJ, Nataatmadja M, et al. Downregulation of transforming growth factor, beta receptor 2 and Notch signaling pathway in human abdominal aortic aneurysm [J]. Atherosclerosis, 2012, 221 (2): 383-386.
- [3] Arslan-Kirchner M, Arbustini E, Boileau C, et al. Clinical utility gene card for: Marfan syndrome type 1 and related phenotypes [FBN1] [J]. Eur J Hum Genet, 2010, 18 (9).
- [4] Carvill GL, Heaven SB, Yendle SC, et al. Targeted resequencing in epileptic encephalopathies identifies de novo mutations in CHD2 and SYNGAP1 [J]. Nat Genet, 2013, 45 (7): 825-830.
- [5] Sikkema-Raddatz B, Johansson LF, de Boer EN, et al. Targeted next-generation sequencing can replace Sanger sequencing in clinical diagnostics [J]. Hum Mutat, 2013, 34 (7): 1 035-042.
- [6] He J, Wu J, Jiao Y, et al. IgH gene rearrangements as plasma biomarkers in Non-Hodgkin's lymphoma patients [J]. Oncotarget, 2011, 2 (3): 178-185.
- [7] Wu J, Matthaei H, Maitra A, et al. Recurrent GNAS mutations define an unexpected pathway for pancreatic cyst development [J]. Sci Transl Med, 2011, 3 (92): 92ra66.
- [8] Robinson PN, Godfrey M. The molecular genetics of Marfan syndrome and related microfibriilopathies [J]. J Med Genet, 2000, 37 (1): 9-25.
- [9] Dietz HC, Pyeritz RE. Mutations in the human gene for fibrillin-1 (FBN1) in the Marfan syndrome and related disorders [J]. Hum Mol Genet, 1995, 4 Spec No: 1 799-809.
- [10] Ng CM, Cheng A, Myers LA, et al. TGF-beta-dependent pathogenesis of mitral valve prolapse in a mouse model of Marfan syndrome [J]. J Clin Invest, 2004, 114 (11): 1 586-592.

(此文编辑 文玉珊)