



Dystrophin基因3~5号外显子缺失连接片段的克隆和测序

抗肌萎缩蛋白(dystrophin)基因缺失突变是导致假肥大型肌营养不良症的最常见基因突变类型, 占所有突变类型的55%~65%。dystrophin基因缺失的机制至今未完全阐明, 为对dystrophin基因缺失机制作继续探讨研究, 我们继在国内首次直接用PCR法完成了45~54号外显子长片段基因缺失连接片段的克隆和测序后, 再次用该方法成功克隆3~5号外显子缺失的连接片段, 其中2号内含子上发生断裂的序列分析迄今国内外文献还未见有报道。

1 材料和方法

1.1 病例资料

男性杜氏肌营养不良症患者, 31岁, 有DMD家族病史, 经检测3~5号外显子缺失。

1.2 试剂

Taq Plus DNA聚合酶、dNTP购自上海生工生物工程技术有限公司, 扩增长片段的LA Taq DNA聚合酶购自大连宝生物工程有限公司, 西班牙Hispanagar公司Biowest Agarose进口分装琼脂糖, 其他试剂为国产分析纯。

1.3 引物设计

本实验涉及20对外显子引物, 其中18对引物参照Chamberlain等[1]和Beggs等[2]的文献设计, 另2对引物从<http://www.dmd.nl>上获取。

本实验涉及13对内含子引物, 其中在2号内含子分次间隔约30 kb 和3 kb 共设计11对引物, 在5号内含子间隔约3 kb 共设计2对引物。引物通过在线设计程序http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi完成设计, 每对引物均限定其PCR产物长度在301~400 bp范围。为提高克隆基因缺失连接片段的特异性, 经上述引物完成对断裂点的定位后在靠近2个内含子断裂点处再设计1对配对引物(D1-F: 5'-AAATCAAGAGAAGGTTAATGTGGAC-3' D1-R5'-CTTACCTATGACTATGGATGAGAGCA-3')以直接对连接片段进行长片段PCR扩增。以上所有引物均由上海英骏生物技术有限公司合成。

1.4 3~5号外显子缺失连接片段的克隆

常规酚-氯仿法抽提制备上述DMD患者的基因组DNA。

通过18对外显子引物进行PCR反应初步筛查, 确认患者3号、4号外显子缺失, 加用2号、5号外显子引物PCR反应后证实为3~5号外显子缺失。PCR反应条件: 96 °C, 2 min; 94 °C, 30 s, 退火(具体温度依据各引物T_m值确定), 30 s, 72 °C, 1 min, 35循环; 72 °C, 10 min。

用所设计的内含子引物定位断裂点的位置 若相邻2对内含子引物1对扩增出正常阳性结果而另1对扩增为阴性结果, 即可判断内含子上断裂点的大致位置。PCR反应条件同上。

以引物D1-F、D1-R直接对连接片段进行长片段PCR扩增 PCR反应条件: 96 °C, 2 min; 94 °C, 30 s, 54 °C, 30s, 72 °C, 2 min, 35循环; 72 °C, 10 min。

1.5 基因缺失连接片段的测序

将扩增成功的PCR产物纯化后测序。测序工作由上海生工生物工程技术有限公司完成。

1.6 连接片段和断裂点序列特点分析

将所测序列和正常内含子序列进行对比，分析5'和3'端断裂点两侧各100 bp序列的分子结构特点。内含子重复序列的分析采用RepeatMasker程序(<http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker>)，基质附着区(matrix attachment regions, MARs)分析采用MAR-Wiz程序(<http://www.futuresoft.org/MAR-Wiz>)。

2 结果

2.1 经以内含子引物进行PCR反应检测

5'端断裂点位于2号内含子距离第一个碱基153 kb左右的位点处，3'端断裂点位于5号内含子末端邻近6号外显子处。对基因缺失连接片段的长片段PCR反应扩增出长约2 kb的产物片段(图1)。

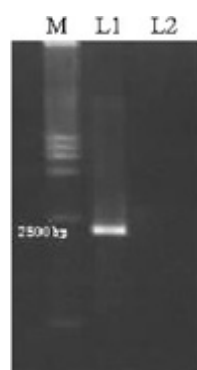


图1 连接片段的PCR扩增电泳图

Fig.1 Agarose gel electrophoresis of the PCR amplified product of the junction fragment

M: DL15000 Marker; L1: A PCR amplified product of about 2 kb from a DMD patient; L2: Normal control

2.2 PCR产物测序

经<http://www.ncbi.nlm.nih.gov/genome/seq/HsBlast.html>作相似性比较证实为dystrophin基因片段，缺失总长度约49 kb。重复序列分析本例基因缺失的5'端断裂点位于2号内含子153106~153387位的SINE/Alu序列内，3'端断裂点在单一序列。MARs分析3'端断裂点在MAR下游4720 bp，5'端断裂点不靠近内含子的MAR区域。5'端断裂点附近在2号内含子正链153424、153428、153432位点以及反链16868、17006位点有拓扑异构酶II酶切位点；3'端断裂点附近在5号内含子正链5726位点以及反链891、1005位点有拓扑异构酶II酶切位点。同时发现3'端断裂点下游15 bp处有TTTAAA序列。分析本例基因缺失相连接的序列无广泛同源性，连接片段插入了26 bp序列并在断裂点周围形成3个13 bp的短序列重复(GGCTTATATTTAA)连接断裂点两端。另外在5'端断裂点上游20 bp位置发现一处A→G点突变(图2、3)。

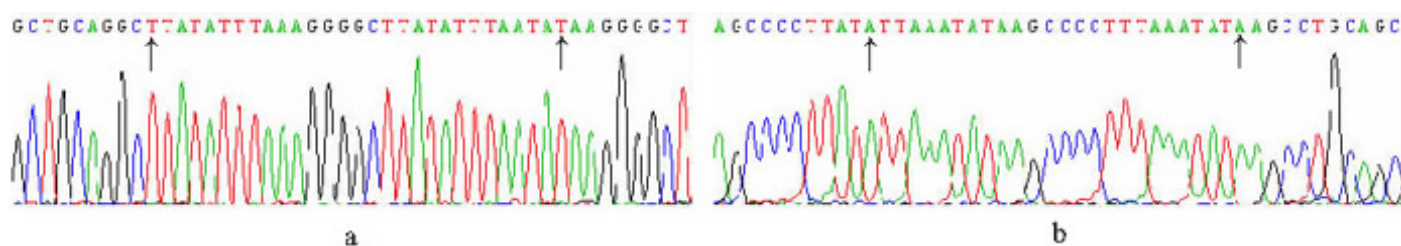


图2 DMD患者3~5号外显子缺失连接片段测序图

Fig.2 Sequence of deletion-junction fragment from a DMD patient with exons 3-5 deletion

a is the forward sequence, and b the reverse sequence. The nucleotides between the arrows indicate the deletion-junction fragment sequence.

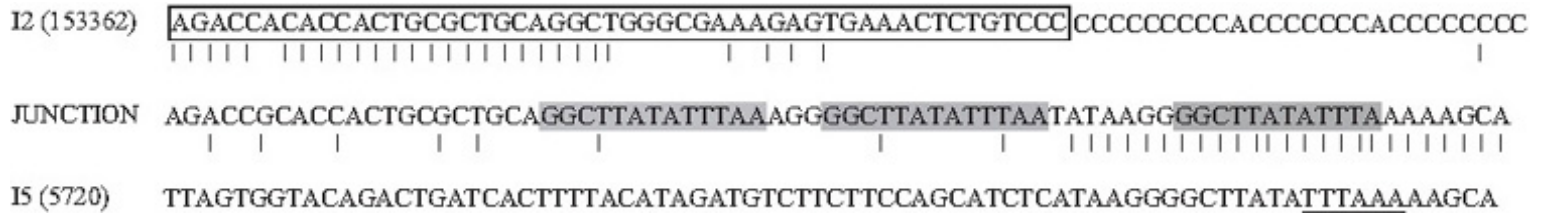


图3 连接片断和相应正常内含子的序列

Fig.3 Nucleotide sequence spanning the deletion junction with corresponding normal intron sequences

Numbers in brackets indicate the breakpoint position with respect to the first nucleotide of the intron (I2, I5). The sequence in the box is SINE/Alu repeated element. The shaded nucleotides are 13-bp duplications. The sequence TTTAAA is underlined. The nucleotides represented in bold is A→G point mutation.

3 讨论

总长约2500 kb 的dystrophin基因是人类最大基因，现有的资料显示dystrophin基因缺失的断裂点都位于内含子上，且可遍布其内含子序列的各个位点，因此准确寻找到断裂点并克隆基因缺失连接片段并非易事。本例的5' 端断裂点位于2号内含子，2号内含子是dystrophin基因上长度仅次于44号和1号内含子的庞大内含子，长约170 kb。研究发现内含子的稳定性主要取决于其内在的分子结构[3]，但随机的长度因素对内含子的缺失率也有重要的影响，因此如何定位dystrophin基因庞大内含子上的断裂点对研究者一直都具有挑战性。本例是我们第2例以PCR步移法定位基因缺失的断裂点后，成功地用断裂点附近上下游的一对引物完成对连接片段的扩增和测序。

分析结果进一步发现5' 端断裂点位于2号内含子上的一个Alu序列内，Alu作为重复序列中的重要一种可以形成发夹环结构，从而使DNA双链易于发生断裂重组。同时观察到两端断裂点附近均有较强的拓扑异构酶II酶切位点关联，有资料显示与强拓扑异构酶II酶切位点相关联可能是涉及DNA双链断裂的因素[4]。此外本例基因断裂重组后的连接片段上游20 bp发现了A→G点突变，以往的文献报告有5例连接片段附近发生了点突变，这种错误的发生符合基因非同源重组的特征，它提示断裂点末端在连接前可能需进行加工，影响范围可以达到20 bp或以上。

在3' 端断裂点附近发现有6 bp的TTTAAA序列，该序列被认为可以使DNA弯曲从而使基因内部发生重组[5]，然而Sironi等[6]通过对dystrophin基因全长的弯曲位点分析认为仅有不到15%的断裂位点与DNA的弯曲特性相关，因此不认为DNA的弯曲特性是dystrophin基因发生断裂缺失的一个主要因素。

我们注意到本例连接片段插入长达26 bp的序列并在断裂点周围形成3个短序列重复连接断裂点两端。插入的26 bp序列经<http://www.ncbi.nlm.nih.gov/BLAST>作相似性分析未发现与数据库中任何序列有完全性同源，该序列是现有文献资料中dystrophin基因非同源末端连接所插入的3个最长序列之一，并且形成短序列重复数量最多。形成长度不等的序列重复是非同源末端连接的一种方式，该修复机制是显现细胞积极适应DNA双链断裂的主要方式之一[7]。DNA双链分子断端捕获的序列可以有各种各样的来源，包括微小卫星、反转录转座成分、拼接内含子的反向转录成分以及外源性DNA序列。依据本例在庞大的2号内含子上出现基因断裂缺失并插入小的序列重复，同时结合既往的研究资料，我们支持这种猜测即一些大的内含子的原始序列易于发生

DNA双链分子断裂、重组和插入DNA片断，插入的序列又更容易发生进一步的片断插入和重组，这样就形成了这些内含子区域的特性，即长度增加和缺失频率增加[4]。

(责任编辑：吴锦雅)

参考文献：

- [1]Chamberlain JS, Gibbs RA, Ranier JE, et al. Multiplex PCR for the diagnosis of Duchenne muscular dystrophy[A]. In: PCR protocols: a guide to methods and applications[M]. San Diego: Academic Press, 1990, 272-81.
- [2]Beggs AH, Koenig M, Boyce FM, et al. Detection of 98% of DMD/ BMD gene deletions by polymerase chain reaction[J]. Hum Genet, 1990, 86(1): 45-8.
- [3]潘速跃, 谢咏梅, 张 成, 等. 44号内含子并非是Dystrophin基因中央缺失热区最不稳定的内含子[J]. 中华医学遗传学杂志, 2001, 18(3): 191-4.
- [4]Toffolatti L, Cardazzo B, Nobile C, et al. Investigating the mechanism of chromosomal deletion: characterization of 39 deletion breakpoints in introns 47 and 48 of the human dystrophin gene[J]. Genomics, 2002, 80(5): 523-30.
- [5]Nobile C, Toffolatti L, Rizzi F, et al. Analysis of 22 deletion break- points in dystrophin intron 49[J]. Hum Genet, 2002, 110(5):418-21.
- [6]Sironi M, Pozzoli U, Cagliani R, et al. Relevance of sequence and structure elements for deletion events in the dystrophin gene major hot-spot[J]. Hum Genet, 2003, 112(3): 272-88.
- [7]Sargent RG, Brenneman MA, Wilson JH. Repair of site-specific double-strand breaks in a mammalian chromosome by homologous and illegitimate recombination[J]. Mol Cell Biol, 1997, 17(1):267- 77.