

· 实验研究 ·

DMD 基因第 50 号和第 51 号内含子的 克隆和测序^①

盛文利^② 刘焯霖¹ 柴建华² 梁秀龄¹

(1 中山医科大学附属第一医院神经内科,广州,510080 2 复旦大学遗传所人类基因组实验室)

提 要 用 DMD cDNA8 探针从假肥大型肌营养不良(Duchenne muscular dystrophy,DMD)基因的 YAC 克隆的 cosmid 亚克隆库中筛选到 9 个阳性 cosmid 克隆,Southern 杂交鉴定 cosmid c0461 含 DMD 基因的第 51 号外显子,将该 cosmid 亚克隆于 pVC118 中,获得含 DMD 基因第 50 和 51 号内含子的 3.1 kb-*Hind* III 片段的亚克隆,将亚克隆的插入片段分别用 *Sau*3A I 完全和部分酶切克隆于 pUC118 中,Sanger 法双链测序测出质粒克隆的序列并重叠所测出的序列,确定该片段的长度为 3179 bp.与 Speer 测出的该段序列(3159 bp)比较相差 20 bp,有 33 处不同.并发现有重复顺序存在,它们之间可能会发生重组并可能导致第 51 号外显子的缺失.

主题词 基因;肌营养不良蛋白;外显子;内含子;碱基序列

中图分类号 R 746.2

自从 1987 年 Koenig^[1] 等人克隆了 14 kb 的假肥大型肌营养不良症(Duchenne muscular dystrophy,DMD)基因的 cDNA,并测出 DMD cDNA 的全顺序后,有关 DMD 基因的研究已取得不少突破性进展.已有的研究表明约 60%~70%的病人是由于 DMD 基因部分缺失或重复突变造成的,而缺失或重复突变主要集中于两个缺失热区(hot spot),即位于 DMD 基因的 5'端和 DMD 基因的中部^[2].虽然已取得了上述成就,但仍然不清楚 DMD 基因的缺失机制.有关 DMD 基因内含子的顺序资料甚少,而导致外显子缺失的断裂点都在内含子中.在基因缺失的病例中第 51 号外显子单独缺失的病例占有较高的比例.我们通过对含 DMD 基因第 50 号和第 51 号内含子的 3.1 kb-*Hind* III 片段全顺序的测定来研究 DMD 基因的缺失机制.该片段包含长 233 bp 的外显子序列和部

分第 50 号和第 51 号内含子序列,其长度为 3179 bp,与 Speer 测出的该段序列(3159 bp,from Genbank)比较相差 20 bp,且有 33 处不同,本文报道该片段的全顺序,并对测序的策略和技术路线进行了讨论.

1 材料和方法

1.1 随机引物同位素标记 DMD cDNA8 探针

用随机引物标记试剂盒(BRL 公司)进行同位素标记反应.在 1.5 mL Eppendorf 管中加入从低熔点胶回收的 DMD cDNA8 探针 25 ng,再加入 ddH₂O 17 μ L,煮沸变性 10 min,置冰浴 10 min,随后依次加入 1 mmol/L 的 dCTP,dGTP,dTTP 各 2 μ L,随机引物 15 μ L, α -³²P-dATP 4 μ L,Klenow 3U,反应总体积为 50 μ L,混匀,置 37 $^{\circ}$ C 温育 4 h~

① 国家 863 高科技发展计划及国家自然科学基金重点项目资助课题;② 第一作者,1964 年出生,男,博士,讲师

5 h,加入2 μ L 0.5 mol/L 的 EDTA(乙二胺四乙酸二钠)终止反应。

1.2 Southern 杂交

预杂交液为:7%的 SDS(十二烷基硫酸钠),0.5 mol/L 的 Na_3PO_4 ,1 mmol/L 的 EDTA,0.1 μ g/L 鲑精 DNA(先煮沸变性 10 min,冰浴 10 min),于 Hybaid 分子杂交仪中,65 $^\circ\text{C}$ 预杂交 4 h~6 h,然后加入同位素标记的变性的 DMD cDNA8 探针,在杂交仪中 65 $^\circ\text{C}$ 杂交 20 h.倒出杂交液加入洗膜液[2 \times SSC(氯化钠和柠檬酸钠的混合液),0.1% SDS]65 $^\circ\text{C}$ 洗膜 15 min.取出尼龙膜,用 0.5 \times SSC,0.1%SDS 室温洗膜直至背景为 10 counts~20 counts.置-20 $^\circ\text{C}$ 借助增感屏放射自显影 1 周。

1.3 双链质粒 DNA 的制备和纯化

用 Wizard Minipreps DNA Purification Systems(Promega 公司)制备和纯化双链模板.经纯化的模板可直接用于测序反应。

1.4 测序方法

按 Sanger 法双链测序.使用 USB 公司的测序试剂盒进行测序,测序反应按照说明书进行。

2 结果

2.1 含 DMD 基因第 51 号外显子的 cosmid 克隆

从 DMD 基因的 YAC 克隆的 cosmid 亚克隆库中筛选到 9 个阳性 cosmid 克隆,经 Southern 杂交鉴定 cosmid c0461 和 cosmid c0491 均含有 DMD 基因的第 51 号外显子,结果见图 1。

2.2 含 DMD 基因第 51 号外显子的 3.1 kb-*Hind* III 片段的亚克隆

将 cosmid c0461 亚克隆于 pUC118 中,杂交鉴定后证实 4 号和 46 号亚克隆均含 DMD 基因第 51 号外显子的 3.1 kb-*Hind* III 片段.结果见图 2。

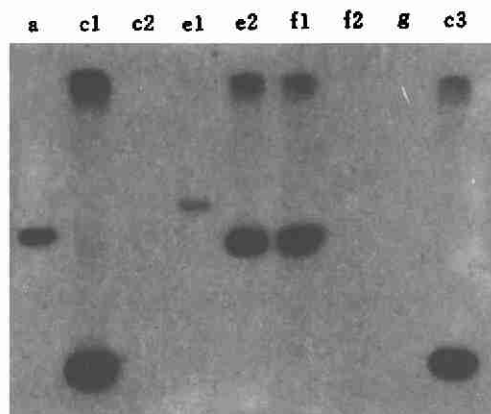


图1 含 DMD 基因第 51 号外显子的 cosmid 克隆
a, a1287; c1, c0461; c2, c0691; c3, c0491, e1, e0143; e2, e0691; f1, f0791; f2, f0891; g, g01493

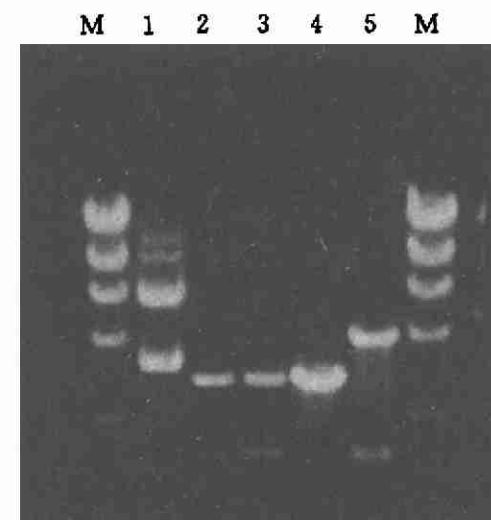


图2 阳性亚克隆的酶切电泳
M: Marker; 1: 质粒 DNA; 2和3: 4号亚克隆 *Hind* III + *Kpn* I 酶切; 4: pUC118-*Hind* III; 5: 4号亚克隆 *Kpn* I 酶切

2.3 含 DMD 基因第 51 号外显子的 3.1 kb-*Hind* III 片段全顺序

Sanger 法双链测序测出该片段的全顺序,确定该片段的长度为 3179 bp,其中长 233 bp 的第 51 号外显子的序列与 Koenig 测出的序列相同.部分序列的照片见图 3.我们所测出的该片段的全顺序如图 4。

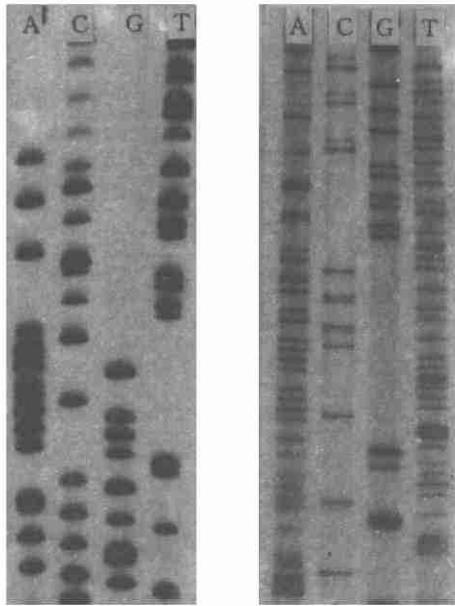


图3 第50号(左)和第51号(右)内含子部分序列

3 讨论

3.1 研究第51号外显子缺失的意义

DMD 基因长约2 300 kb,含79个外显子,外显子的平均长度为200 bp,范围为32 bp~500 bp.内含子的平均长度为35 kb,范围为1 kb~180 kb^[3].由于外显子的序列已测出,而导致 DMD 基因外显子缺失的断裂点均位于内含子中,故对 DMD 基因内含子序列的测定是弄清 DMD 基因缺失的关键所在.但有关 DMD 基因的内含子的序列资料甚少.我们测定了 DMD 基因第51号外显子及两侧的第50号和第51号内含子的部分核苷酸顺序,籍此来研究该外显子的缺失机制.并推测 DMD 基因其它的外显子的缺失机制.之所以选择第51号外显子来研究 DMD 基因的疾病机制,是因为第51号外显子位于 DMD 基因的缺失热区,且编码肌营养蛋白(dystrophin)的绞链3和第3疏水区.这两个结构在维护细胞膜的稳定性方面有重要的作用^[4].故研究 DMD 基因第51号外显子的缺失有重要的医学意义.我们测出的 DNA 片

段的长度为3 179 bp,包含了第51号外显子的全部和第50和第51号内含子的部分,在第50号和第51号内含子中发现36个短段串联重复顺序,22个正向和反向重复顺序和8个同源顺序,提出重复顺序之间的重组可能是导致第51号外显子缺失的假设(有关论文另发).推测在 DMD 基因的其它内含子中也有类似的重复顺序,由于此重复顺序之间的重组导致了两断裂点之间的外显子的缺失.

有必要说明的是,我们测出的第50号和第51号内含子的序列与 Speer 测出并送 Genebank 的序列有33处不同,且 Speer 测出该片段的长度为3 159 bp,我们所测出的该片段的长度为3 179 bp,相差20 bp.但我们在重叠克隆时前后进行过3次测序,并对含该片段的另一亚克隆(即第46号亚克隆)进行过测序,测出的第4号和第46号亚克隆的序列相同,更为重要的是 Speer 测出的序列并未在正式的刊物上发表,故这两段序列有一定的可比性.

3.2 测序策略

“读出”和“读懂”人类基因组的长约30亿个核苷酸顺序是人类基因组计划的两个主要任务^[5].测序本身是一个花费巨大的工程,如何在尽可能少花钱的情况下获得尽可能多的序列资料是我们进行测序必须面对的问题.在测序之前进行测序策略和技术路线的设计是解决这个问题的较好办法.现有的测序策略主要有两种:即鸟枪法和定向法.这两种方法的缺点是:虽然可通过控制条件将待测 DNA 片段打断为大小适合测序的片段,但待测 DNA 片段被打断后由于片段的末端为非粘性末端,故这种片段难以克隆,虽然可通过修补末端来提高转化效率,但繁琐且克隆效率仍然较低,大大降低了工作效率,而且要测定待测片段长度约5倍~10倍的序列方能重叠所有的克隆,故工作量和测序成本均非常大^[6].我们是采用将待测 DNA 片段用 *Sau3A I* 完全酶切,随后用 *BamH I* 酶切的载体与 *Sau3A I* 消化的待测 DNA 连接和转

```

1    AAGCTTGAGA GACAAGAAAC ATTCTTCCAT TCTACTCATC TTCTTCTCTA ATGAGGAGAC
51   AACCTTAAAA GCACAGTTAC ATAGCCATAA AAATTAATGA TTGGCTACCT CAGAATGAAA
121  ATTCAATGTC TCATTTTTTTT TTAATATTCT TAGAATCGTT CACTGGTTGT CCAGTGTGAG
181  TCTCCTGTTG AGATGCTCTTT TGCAGCTTTC CTTGAAACCT TTCATTCCAA ACTACATAGT
241  CCAATAATTT TGCCACCAAT CTTCTGGTTA TATTATGCTC TTGAGTCTGT TGTCTATAAA
301  CTTGATTAGG CATTCCGGCC CCTCACCCT CACCTCTGAT AACCCAGCTG TGTGTTGGTA
361  TTTAGTATCA ATTACACCA GCAAGTTCAG CCCTCTTCAA TCAATATAGG GCCACACAGG
421  ACTTTTTTGA CTGACTACTC CCCAAGTATT TCACATTTTG GAGGCCTTAT CTCCAGTTTC
481  TCACCACAGT TGTCATCAC TGTGTTTCTT ACTAGCCAGG CGTTTATAAA AACAGTAATA
541  CCTAACACTA TTGATCACCT ACTATAGTGT CAGGCGCTGT AATAATATTA TTGTGATGAT
601  GATGATTATG CTGCTCTTCT CCGATTGTCA TACGTGTATT GCCTGTACTA CTCACTGAAT
661  CTACACAAGT GCCCTTATGA CATTTACCCT GTTATTATTC CTCTTTTAAG GTAAAATCAT
721  GAAAAAGCTC TCCCACTTTG CCTTGCTTAC TGCTTATTGC TACTGTCTGAA CAATGTGTAG
781  AACTGAAACT TAGAGAGGTT ATGTGGCTTT ACCAAGGTCC CAGAGTTCCT AGGGCAGAGA
841  ACAGGATTGT CTACCAGACA TTTTAATTCT AGTACTATGC ATCTTAACCA TTACCATAGG
901  CTGACTTACT CTACAGTGT CCAACTATT CATATTAAGA TTTATTTAAT GACTTTGAAA
961  CGTATTCAT CTCTAAATAG AAAAACTACT AACTCGCATT TTTAAGAAAA TATTGTATCT
1021 TGGTTTTTCT TCACTGCTGG CCAGTTTACT AACAATCTGA AATAAAAAGA AAAAAATATG
1081 ATAAACTGTG CCCAGTATAA AATACAGAGC TAAGACAAGA ACGTGTTC A TTGGCTTTGA
1141 TTTCCCTAGG GTCCAGCTTC AAATTAATTT ACTTCTATTC AAGGGAATTT TAAATCAGAA
1201 AGAAGGATCT AGGATCTTAT CCCATCTTGT TTTGCCTTTG TTTTCTTCTG AATAAAAAAA
1261 AAATAAGTAA AATTTATTTT CCTGGCAAGG TCTGAAAAC TTTGTTTTTC TTTACCACTT
1321 CCACAATGTA TATGATTGTT ACTGAGAAGG CTTATTTAAC TTAAGTTACT TGTCAGGCA
1381 TGAGAAATGAG CAAAATCGTT TTTTAAAAAA TTGTTAAATG TATATTAATG AAAAGTTTGA
1441 ATCTTTTCAT TTTCTACCAT GTATTGCTAA ACAAAGTATC CACATTGTTA GAAAAAGATA
1501 TATAATGTCA TGAATAAGAG TTTGGCTCAA AGTTGTTACT CTTCAATTAA ATTTGACTTA
1561 TTGTTATTGA AATTTGGCTCT TTAGCTTGTG TTTCTAATTT TFCTTTTTCT TCTTTTTTCC
1621 TTTTTGCAAA AACCCAAAT ATTTTAGCTC CTACTCAGC TGTACTCTG GTGACACAAC
1681 CTGTGGTTAC TAAGGAACT GCCATCTCCA AACTAGAAAT GCCATCTTCC TTGATGTTGG
1741 AGGTACCTGC TCTGGCAGAT TCAACCCGGC TTGGACAGAA CTTACCGACT GGCTTCTCT
1801 GCTTGATCAA GTTATAAAAT CACAGAGGGG TGATGGTGGG TGACCTTGAG GATATCAACG
1861 AGATGATCAT CAAGCAGAAA AATGATAAA GTTGGAGAAA GTTGGCAGAA GTTTTTCTTT
1921 AAAATGAAGA TTTTCCACCA ATCACTTTAC TCTCCTAGAC CATTTCCAC CAGTTCCTAG
1981 GCAACTGTTT CTCTCTCAGC AAACACATTA CTCTCACTAT TCAGCCTAAG TATAATCAAG
2041 GATATAAAT AATGCAAATA ACAAAGTAG CCATACATTA AAAAGGAAAT ATACAAAAAA
2101 AAAAAAATAA AAAAGCAGAA ACCCTTCAAG AATAGTTGTC TCAGTTAAAT TTACTAAACA
2161 ACCTGGTATT TTA AAAATCT ATTTTATACC AAATAAGTCA CTCAACTGAG CTATTTACAT
2221 TTA AACTGT TGT TTTGGAC TACGCAGCCC AACATATTGC AGAATCAAA ATAAATAGTCT
2281 GGGAAATGTT ATTA TCCACT CTTCTAAGTT GTCTGTGCCA ATTTGCCTTC TCCAATGATA
2341 AGGATAATTG AAAGAGAGCT ATA AACTTAAA AAGAGAAAAA TAACAAAAACA TAAGATATTT
2401 AAAATTACCC TAGATCTAAA GTTGGCATT ATGCAATGCC ATGTTCAAAT GAACATGTTT
2461 TTAATACAAA TAGTGCAATT TTCCTGTCAG TCCTCAGTGT AATCCATTTG GTA AATATC
2521 GTACATCAAC TAGAAAACAT AGAATACATT GATGTAATA TGGTTTACCT AGCTAGATCA
2581 ATATACTATA TATCTTTTAT ATTTGTGAAT GATTAAGAAA AAATAATGTT GGAATTGTTA
2641 TACATTAAG TTTTTTCACT TGTAACAGCT TPCAAGCCT TCTAAAGAAA TACAAGTTG
2701 TGCTGAAGGT ATTTAGGTAT TAAAGTACTA CCTTTTGA AAACAAGA TGAGGCAGAC
2761 AGAGTAAGGG GAATTTCTT GTAAAATAA CTTCAACCA TCCATAGAAT AAAAGTAATT
2821 TGATAGTAAA CAACCTGCAT TTA AAGGCCT TGAGCTTGAA TACAGAAGAC CTGAATTCAG
2881 TGCCATTGTC AAATGATGAT TGTGGTCAAG CCATCTCTGG ATCTTCTGTT CCTATCTGA
2941 GTACAGAGCA TACTGAGTAC ACATTCACAT TCACAATATA GTTATGGATA TGGATGTATA
3001 TAAATATATG TAAATACTAC ATATATGTAC CAAAATTTGT TTTACTFTG CTTTAAAAAA
3061 AGTAATTATA GCCATGACAT TTTTTTCAGAA AAAGTAACTG AGGAGGGCTC ATAGATGTCA
3121 AATCCCAGT AAGTAGCAGA ACAAGGATTC AAATCCAAGT CCATTTGATT CCTAAGCTT

```

图4 DMD 基因第51号外显子3.1 kb-Hind III 片段全顺序图

化,对所获的亚克隆进行序列的测定,由于 *Sau3A I* 是4碱基识别酶,平均每256个碱基就有一个识别位点,故可以将待测 DNA 片段消化

为大小适合测序的片段由于用 *Sau3A I* 酶切后产生的是粘性末端,可与 *BamH I* 酶切的载体进行连接和转化;故克隆的效率较高。然后再用

Sau3A I 对待测序 DNA 片段进行不完全酶切, 不完全酶切后可获得一组相互有重叠的片段, 再亚克隆于 pUC 载体中, 对所获得的亚克隆进行两端测序, 最后将序列资料输入计算机重叠各亚克隆就可获得待测片段的全顺序, 按这种策略来测序只需要测定待测片段约两倍的序列就可获得待测片段的全顺序, 所以这是一省时、省力的方法。

3.3 测序的技术路线

我们用 DMD cDNA8 探针从 DMD 的 YAC 克隆的 cosmid 亚克隆库中筛选到 9 个阳性 cosmid 克隆, 进一步鉴定后确认 cosmid c0461 和 cosmid c0491 含有第 51 号外显子, 将 cosmid c0461 DNA 用 *Hind III* 酶切后亚克隆于 pUC118 中, 点杂交鉴定后确认 4 号和 46 号亚克隆含有 DMD 基因第 51 号外显子的 3.1 kb-*Hind III* 片段, 对 4 号克隆的插入片段从低熔点胶中回收后用 *Sau3A I* 完全和不完全酶切后再克隆于 pUC 载体中, 对所获得的亚克隆进行测序, 将序列资料输入计算机重叠各克隆后即获得 3.1 kb-*Hind III* 片段的全顺序。我们所使用的这条从 YAC→cosmid→pUC 的技术路线是目前应用较广泛的技术路线, 从目前已报道的大规模测序的结果来看, 这条技术路线是行之有效的^[7]。

3.4 测序方法

对 DNA 一级结构的测定仍然沿用 Sanger^[8] 的加减法和 Maxam-Gilbert^[9] 的化学法, 但 Sanger 法应用最为广泛。Sanger 法测序最初所使用的模板是从重组 M13 噬菌体颗粒中分离获得的单链 DNA, 80 年代中期有人用经热变性或碱变性的双链模板测序获得成功。时至今日, 双链测序已得到广泛应用^[10], 过去人们一直认为采用 CsCl-溴化乙锭梯度平衡离心法纯化的质粒 DNA 是双链测序的最佳模板, 但 CsCl-溴化乙锭梯度平衡离心法纯化的质粒 DNA 均需花费大量的人力和物力。并认为双链测序比单链测序成功率要低, 可读出的长度要比单链测序的短, 背景也比单链测序的高。我们是用 Promega

公司的 Wizard Minipreps DNA Purification Systems 来纯化双链模板的, 用该方法纯化的质粒 DNA 进行双链测序同样可获得满意的效果。从图 3 的测序照片可以看出使用该方法纯化的模板 DNA 进行的双链测序反应可读性较强, 背景较低, 每个测序反应可读出约 400 bp 左右的序列, 所以是一较好的方法。

参 考 文 献

- 1 Koenig M, Hoffman EP, Bertelson G, *et al.* Complete cloning of the duchenne muscular dystrophy and preliminary genomic organization of the DMD gene in normal and affected individuals. *Cell*, 1987, 50:509
- 2 Den Dunnen JT, Grootsholten PM, Bakker E, *et al.* Topography of the duchenne muscular dystrophy gene; FIGE and cDNA analysis of 194 cases reveals 115 deletions and 13 duplications. *Am J Hum Genet*, 1989, 45:835
- 3 Roberts RG, Coffey AJ, Borrow M, *et al.* Exon structure of the human dystrophin gene. *Genomics*, 1993, 16:536
- 4 柴建华. DMD 基因的结构及其变异. *科学通报*, 1993, 38:2113
- 5 柴建华. 人类基因组研究将有大的发展. *国际学术动态*, 1994, 57(2):58
- 6 Maniatis T, Fritsch EF, Sambrook J, *et al.* *Molecular Cloning, a laboratory manual*. 2nd ed. New York: Cold Spring Harbor Laboratory Press, 1989. 610
- 7 Martin-Gallardo A. Automated DNA sequencing and analysis of 106 kilobases from human chromosome 19q13.3. *Nature Genetics*, 1992, 1:34
- 8 Sanger F, Coulson AR, Barrel BG, *et al.* A rapid method for determining sequences in primed synthesis with DNA polymerase. *J Mol Biol*, 1975, 94:441
- 9 Maxam AM, Gilbert W. A new method for sequencing DNA. *Proc Natl Acad Sci*, 1977, 74:560
- 10 Zhang H, Scholl R, Browse J, *et al.* Double strand ed DNA sequencing as a choice for DNA sequencing. *Nucleic Acid Research*, 1980, 16:

THE CLONING AND SEQUENCING THE 50TH AND 51ST INTRONS OF DMD GENE

Sheng Wenli Liu Zhuolin Chai Jianhua Liang Xiuling

(Department of Neurology, First Affiliated Hospital, Sun Yat-sen University of Medical Sciences, Guangzhou, 510080)

With DMD cDNA8 probe, the DMD-cosmid library from DMD-YAC clones was screened and 9 positive DMD-cosmid clones were obtained. The cosmid c0461 was confirmed to contain the 51st exon by Southern hybridization. The cosmid was subcloned into pUC118 and the subclone containing the 3.1 kb-*Hind* III fragment was obtained. The insert fragments were digested and partially digested with *Sau*3A I, then were subcloned into pUC118. The plasmid subclones were sequenced with double strand sequencing of Sanger's method. It was determined that the fragment was 3 179 bp long. Compared with the sequences by Speer's, it was differ at 33 sites and 20 bp longer than Speer's. It was possible that the recombination of repeat sequences was the reason of 51st exon deletion.

Subject headings genes; dystrophin; exons; introns; base-sequences

~~~~~

·新成果·

## 白内障的防治研究

课题负责人 李绍珍

白内障是当今最主要的致盲眼病。本研究分别将人透明和白内障晶体中 r 晶体蛋白、糖的含量测定与醛糖还原酶的活性及黄酮类化合物对其抑制作用研究相结合,将丙二醛和超氧化物歧化酶的测定与近紫外线对胚晶体蛋白的光氧化作用研究相结合,将人血清抗晶体蛋白抗体测定与机体对晶体蛋白的免疫反应研究相结合,从能量代谢、细胞呼吸、氧化和自由基防御以及体液和细胞免疫等方面探索白内障的发病机理;通过晶体上皮细胞的培养与抑制,探索后发性白内障的发生与预防;用视诱发电位观察正常婴幼儿与先天性白内障的视功能发育以探索弱视的预防。通过对我国部分地区的流行病学调查,证实了白内障是这些地区盲目的最主要原因,找出相关危险因素。在我国率先进行复杂性白内障的囊外摘除、二期和门诊人工晶体植入术、预先 YAG 激光前囊膜切开术、手术并发症防治的研究,提高了术后的视觉质量。研究达国际先进水平,1995年获卫生部科技进步三等奖,1996年获国家科技进步二等奖。

(陈丽芳)