

人类遗传病的家系收集疾病基因定位克隆与疾病基因功能的研究

夏家辉

(中国医学遗传学国家重点实验室, 长沙 410078)

[摘要] 介绍了中国医学遗传学国家重点实验室在遗传病家系收集、疾病基因定位、疾病基因克隆和疾病基因功能研究方面的研究工作。用细胞遗传学G显带技术于1975年发现了一条与鼻咽癌相关的标记染色体t(1;3)(q44;p11);1981年将睾丸决定基因(TDF)定位于Yp11.32带;1991年以来收集遗传病家系345种共590个;1996年用显微切割、PCR、微克隆技术克隆了EXT2基因;1998年用基因家族-候选疾病基因克隆方法克隆了遗传性神经性耳聋基因GJB3;1999年用连锁分析和全基因组扫描将一种遗传性弥漫性浅表性光敏性汗孔角化症定位于12q23.2带,并在基因功能研究中发现了一个新的细胞内转运蛋白。

[关键词] 遗传病家系;基因定位和克隆;基因家族-候选疾病基因克隆;基因组扫描;基因功能研究

前言

1972年我们建立了国际上1971年发现的人类染色体G显带技术,1979年建立了国际上1977年发现的人类染色体高分辨技术,并结合国内的实验条件,在方法学上作了重大改进。1975年用G显

带技术发现了一条与鼻咽癌相关的标记染色体t(1;3)(q44;p11)(图1),1981年用高分辨技术在国际上最早将人类睾丸决定基因(TDF)定位于Yp11.32带(图2),引起了国内、外的重视。

为筹建中国医学遗传学国家重点实验室,在1984~1985年笔者对加拿大、美国15个相关实验

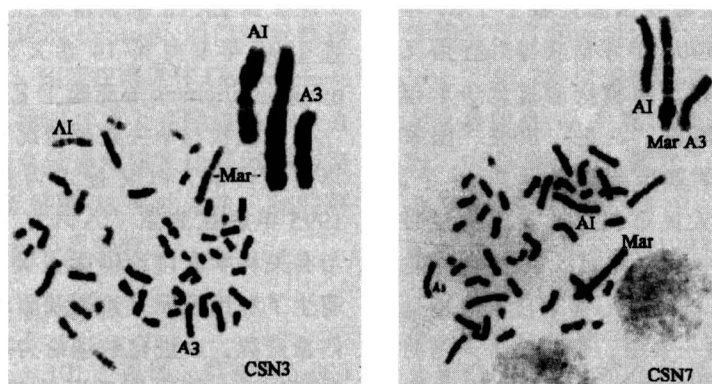


图1 1975年用G显带技术发现了一条与鼻咽癌相关的标记染色体t(1;3)(q44;p11)

Fig.1 A nasopharyngeal marker chromosome t(1;3)(q44;p11) was found by G-banding technique in 1975

[收稿日期] 2000-06-26

[作者简介] 夏家辉(1937-),男,湖南桃江县人,中国工程院院士,湖南医科大学人类与医学遗传学教授,博士生导师

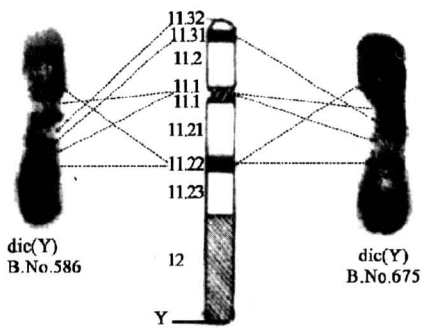


图2 1981年用高分辨技术在国际上最早将人类睾丸决定基因(TDF)定位于yp11.32带

Fig.2 TDF gene was firstly mapped to chromosome Yp11.32 by high-resolution banding technique in 1981

室进行了考察。考察结论是：在细胞遗传学方面，本实验室已处国际前沿；但在分子遗传学研究方面至少相差国外10年。从而使我们认识到，在分子遗传学方面如果沿用外国人现有的研究方法，是赶不上的，是不能完成国家的建室任务的；要赶，首先必须在方法学上有创新，同时实验室要在本学科领域代表国家在国际讲坛上占有一席之地，也必须在国际前沿最活跃的分子遗传学研究领域开展研究工作，在国际一流杂志上发表文章。为此，我们夜以继日地为在实验室内建立成套的分子遗传学实验技术，为了在中国本土上克隆遗传病的疾病基因而不懈探索。为建立重点实验室国家计委于1989年投入了 120×10^4 美元，1996年评估获得好评后又投入了 400×10^4 元人民币。实验室通过竞争承担了卫生部、国家教委、国家“八五”攻关、国家“八六三”、国家自然科学基金各类课题共29个，获得资助 560.1×10^4 元人民币，加上美国史克必成(SB)公司投入 236×10^4 美元，以上共计达 $3\,914.9 \times 10^4$ 元人民币。这些钱不管来自国内还是国外，都源于中华民族这一条根，都凝集了人民的血汗。作为代表中国的医学遗传学国家重点实验室，面对国际“人类基因组项目”的实施，每年近百个人类遗传病的疾病基因被克隆、被专利保护(到1998年11月20日国际上已克隆遗传病疾病基因890个)。中华民族占世界人口的1/5，是医疗、保健大国，却没有拿到一个遗传病的疾病基因，因此，“抢疾病基因”变成了实验室每时每刻的核心话题，几乎占据了本实验室骨干成员生活中的一切。

基础研究只有世界第一。1986年我们提出将1985年由美国Kary B. Mullis发现的PCR技术(1993年获诺贝尔化学奖)与染色体显微切割技术相结合，建立定点克隆基因的技术。由于国内研究经费申请困难，我们利用国外的条件，于1989年首先在日本建成了该技术，并用该技术先后在国内开展了睾丸决定基因(TDF)、遗传性多发性外生性骨疣遗传病(EXT)疾病基因的克隆。1990年当我们用显微切割、PCR、微克隆技术构建好Yp11.32带的DNA文库、克隆睾丸决定基因时，哈佛大学的Berta教授等宣布已克隆了此基因，并邀请我们于1991年10月在华盛顿召开的第8届国际人类遗传学会上发言，参加由他组织的专题讨论。在华盛顿会议上，笔者决定将家系收集与基因克隆工作马上转到外生性骨疣病的基因克隆上。在1991年9月至1995年10月的整整4年中，本实验室20多位研究生和技术人员，每周工作7天，每天工作12小时以上，为了解决实验中的一些技术难题，有时不分昼夜地吃住在实验室。遗憾的是，正值我们1995年克隆到位于染色体8q24位点与外生性骨疣病可能相关的2个克隆时，Nature Genetics在1995年10月第11卷发表论文，美国的Jung Ahn等已克隆到EXT1疾病基因；由于EXT病涉及到位于8号、11号、19号染色体上的三个基因，因此我们立即转到克隆位于11号染色体的EXT2基因上，更可惜的是当我们于1996年9月拿到EXT2基因准备投稿时，Nature Genetics在1996年9月第14卷又发表论文，美国的Dominique Stickens等克隆了EXT2基因。

在数次竞争失败的情况下，笔者开始考虑放弃以显微切割、PCR技术为主的路线，另辟蹊径。1995年底在北京“八六三”课题进展汇报会上，与有关科学家讨论国际人类基因组项目的进展中，萌生了利用国际“人类基因组计划”研究所积累的信息资源，以进化的理论为基础，在计算机上进行同源分析，筛选新基因的想法。经过近半年的摸索，终于在1996年7月成功地建立了基因家系——候选疾病基因克隆新方法，并用这一方法在不到两年的时间内抢先在国际上克隆了M6ba、Atrophin-1样基因、Ataxin-2样基因、DMGDHL1、GJB3、GJB5、MPZL1等7个与遗传疾病相关的基因，在国际基因库作了登记。1998年3月，我们克隆了间隙连接蛋白 $\beta-3$ 基因

(GJB3), 并用染色体原位杂交技术将其定位在 1 号染色体短臂 3 区 3 带到 3 区 5 带上。随后, 我们对实验室收集的 42 个相关疾病家系进行突变检测, 5 月 28 日终于从浙江和湖南两个神经性耳聋家系中发现了该基因突变, 从而确定了 GJB3 是决定人类遗传性神经性高频性耳聋的疾病基因。我们夜以继日地撰写论文向 Nature Genetics 作了投稿。编辑部收稿后复函通知: 我们正在审稿, 如果在审稿过程中, 国际上有同一基因的论文发表, 本刊就不再刊登了。经过三次修稿, 该刊于 1998 年 20 卷 12 期上发表我们的论文 Mutations in the gene encoding gap junction protein β -3 associated with autosomal dominant hearing impairment, 并给我们通知说, 编辑部将在 11 月 30 日美国东部时间 17 时在网上传布有关该研究的新闻, 并要求在他们发布新闻之前, 我们不要向外界作任何报道。编辑部 11 月 30 日的新闻称: 湖南医科大学中国医学遗传学国家重点实验室夏家辉教授等的这些发现为细胞通讯的重要性以及这些介导细胞通讯过程的亚单位是如何在不同的细胞类型中发挥作用提供了依据。在同期杂志上, 编辑部不但将该基因 Connexin connections 作为本期封面头条, 而且还刊登了英国诺丁汉大学 Karen P. Steel 教授以 "One connexin, two diseases" 为题的有关该基因研究的评论文章。Steel 教授在评论文章中指出: 夏教授和 Richard G 教授分别在听力减退和皮肤疾病中发现了 GJB3 突变这一结果, 为研究 Connexin 31 (见封面彩图) 蛋白不同区域的功能和由它组成的 Connexin 的门控机制提供了基础。论文发表后, 国际权威著作 OMIM Home 立即作了收录。这是在我国本土上克隆的第一个遗传病疾病基因, 是我国克隆遗传病疾病基因零的突破。经过 1972—1998 年整整 26 年的追求, 我们第一次圆了冲出亚洲、走向世界的梦。

1 家系收集与基因定位

人类约 $(10 \sim 15) \times 10^4$ 个基因, 其中约 5 000 ~ 10 000 个是疾病基因, 它们分布在 23 对染色体上, 由于 X 和 Y 染色体上的基因是不同的, 故人类有 24 个基因连锁群, 要从 10^5 个基因中找出某一个基因是某一种遗传病的疾病基因方法很多, 一般的程序是首先确定这个疾病基因在染色体上的位置, 即基因定位。基因定位的方法也很多, 最有效的方法有:

1.1 细胞遗传学定位法

即通过特殊的带有染色体异常的病例进行定位, 如笔者在 1981 年定位的睾丸决定基因 (TDF); 国际上定位的假性肥大型肌营养不良病基因 (DMD); 外生性骨疣病基因 (EXT1), 脆性 X 综合症基因 (Frax1) 等。

1.2 遗传病家系连锁分析法

我国人口占世界总人数的 1/5, 自然大家就会想到我们是人类遗传资源的大国, 其实我国遗传资源存在着一个抢救的问题。因为解放前我国的医疗水平低, 死亡率高, 一些具有遗传缺陷的病人死亡率更高; 解放后, 医疗条件虽有了改善, 寿命延长, 但随着一对夫妻生一个子女的计划生育政策的贯彻执行, 除了 30 岁左右的人有数个亲兄弟姐妹外, 20 岁左右的人已经多数没有亲兄弟姐妹了, 所以我国的遗传资源实际存在一个抢救的问题。为了抢救遗传病大家系, 本实验室自 1992 年受国家自然科学基金资助建立了 "中国人染色体突变细胞库"; 1994 年再次受国家自然科学基金资助建立了 "中国人突变细胞库", 1996 年受美国 Smithkline Beecham 公司资助, 扩建为 "中国遗传病家系收集数据库", 该库由 "中国人基因病家系库"、"DNA 样品库" 和 "中国人突变细胞库" 三部分组成。1997 年在 "八六三" 重大项目的资助下建立了 "人类孟德尔遗传疾病数据库管理系统", 其内容包括国际上所记载的遗传病种、分科、基因定位和克隆情况。并以上述各数据库为基础在中国医学遗传学国家重点实验室内成立了 "中国遗传病资源保藏中心" 和 "中国遗传病基因诊断与治疗研究中心"。至今全国有 18 个省、市向本库共申报遗传病 345 种, 家系 590 个, 建立了相应的细胞株 453 个, 保存 DNA 样品数 3327 个, 为遗传病疾病基因的克隆打下了基础。从 1993 年 5 月 23 日开始, 到 1999 年 9 月历时 7 年, 我们对一个分布在湖南望城、长沙、南县三个县的一种遗传性弥漫性浅表性光敏性汗孔角化症的家系进行了调查和血样收集。该家系共 8 代 538 人, 患者 75 人, 现存活第 4 至第 8 代, 共采集血标本 126 人, 其中病人 44 人, 用 382 个遗传标记作全基因组扫描, 用 Linkage (version 5.1) 软件作连锁分析, 在 D12S78 位点的 Lod 值达 20.53, 然后在 D12S78 两侧选择 16 个 Marker 作精细定位, 最后将该疾病基因定位在 D12S1727 和 D12S1605 之间的 9.6 cM 范围内(表

1,图3)。1999年12月1日我们向国际皮肤病杂志 The Journal of Investigative Dermatology 作了投稿,

这是国际上首次对这种病作的定位,(现已发表在该刊114卷,2000年第6期)。

表1 用12号染色体长臂q23.2带到长臂q24.1带的16个遗传标记对汗孔角化症疾病基因所作的连锁分析
Table 1 Two-Point LOD scores between the disease gene and 16 markers of chromosome 12q23.2 - q24.1

Marker	LOD score at $\theta = \alpha$							Genetic Distance ^b	Z_{\max}	θ_{\max}
	0.00	0.001	0.01	0.05	0.10	0.15	0.20			
D12s1671	−∞	6.24	8.06	8.60	8.10	7.30	6.36	0.5	8.63	0.04
D12s306	−∞	9.34	10.16	10.05	9.30	8.38	7.35	1.4	10.27	0.02
D12s1727	−∞	−1.78	1.11	2.71	2.98	2.86	2.57	1.3	2.98	0.10
D12s1607	4.94	4.93	4.84	4.44	3.93	3.41	2.88	2.5	4.94	0.00
D12s318	5.97	5.96	5.86	5.41	4.83	4.22	3.60	1.4	5.97	0.00
D12s78	20.53	20.50	20.20	18.84	17.08	15.22	13.27	0.0	20.53	0.00
D12s338	10.92	10.90	10.77	10.10	9.19	8.19	7.13	3.3	10.92	0.00
D12s1342	12.92	12.90	12.70	11.79	10.61	9.39	8.13	0.0	12.92	0.00
D12s330	11.68	11.65	11.46	10.61	9.51	8.37	7.18	0.5	11.68	0.00
d12s1613	6.91	6.90	6.79	6.27	5.57	4.83	4.07	0.6	6.91	0.00
d12s1605	−∞	5.67	6.57	6.76	6.35	5.76	5.07	1.2	6.80	0.03
d12s84	−∞	9.03	11.11	11.93	11.37	10.38	9.17	1.6	11.93	0.05
d12s1583	−∞	8.57	11.35	12.39	11.90	10.94	9.74	3.4	12.39	0.05
D12s1341	−∞	−1.52	4.28	7.52	8.05	7.74	7.05	0.7	8.05	0.10
D12s354	−∞	1.98	6.80	9.32	9.41	8.77	7.83	1.5	9.50	0.08
D12s369	−∞	2.64	3.86	4.60	4.51	4.15	3.66	−	4.62	0.07

a-LOD值计算的条件设定为常染色体显性遗传模式,外显率为100%;b-与下一个Marker的遗传学距离(cM),根据Genethon的人类遗传连锁图确定(1996年)

2 疾病基因克隆

通过基因定位可将某一疾病基因定位到某号染色体或某号染色体的某一区带(细胞遗传学概念,高分辨染色体技术可显出1000条带,平均每条带约含 300×10^4 个核苷酸,约100~150个基因),或某一区带的十几个或几个分摩内(分子遗传学概念,人类染色体为3300分摩,每个分摩约含 90×10^4 核苷酸,约30~50个基因)。然后采用各种分子遗传学和计算机相结合的方法,将这个定位区内可能相关的基因分离出来,再用这个基因对家系中的患者与正常者逐个进行基因检测。如果发现某一突变仅存在于患者中,则可初步确定这个基因可能是导致这种病的疾病基因。最后,再作相应病人或动物相关组织的基因表达,若这个基因在这种组织中表达,则可认为找到了该病的疾病基因。

克隆疾病基因的方法很多,而且随着国际人类基因组计划研究的进展,新的方法在不断地出现,

我们实验室从1989年至今先后采用了两种方法:

2.1 显微切割、PCR、探针池、微克隆方法 (图4,5)

8q24.1带外生性骨疣疾病基因的克隆过程如表2所示。

2.2 利用计算机完成的基因家系—候选疾病基因克隆的方法(图6)

利用国际人类基因组研究所提供的资料,根据生物进化论的观点,首先在计算机上完成基因的模拟克隆,然后用实验证明并克隆出该基因。这一方法的理论基础是人类是由古老的原始的低等生物进化来的,进化的基础是基因突变,现存的物种是基因突变加自然选择(即适者生存,不适者死亡)的结果。该方法的切入点有两个:从一个已知的人类的疾病基因出发寻找新的疾病基因;从一个已知的动物(鼠或果蝇等)的疾病或功能基因出发寻找人的疾病基因。

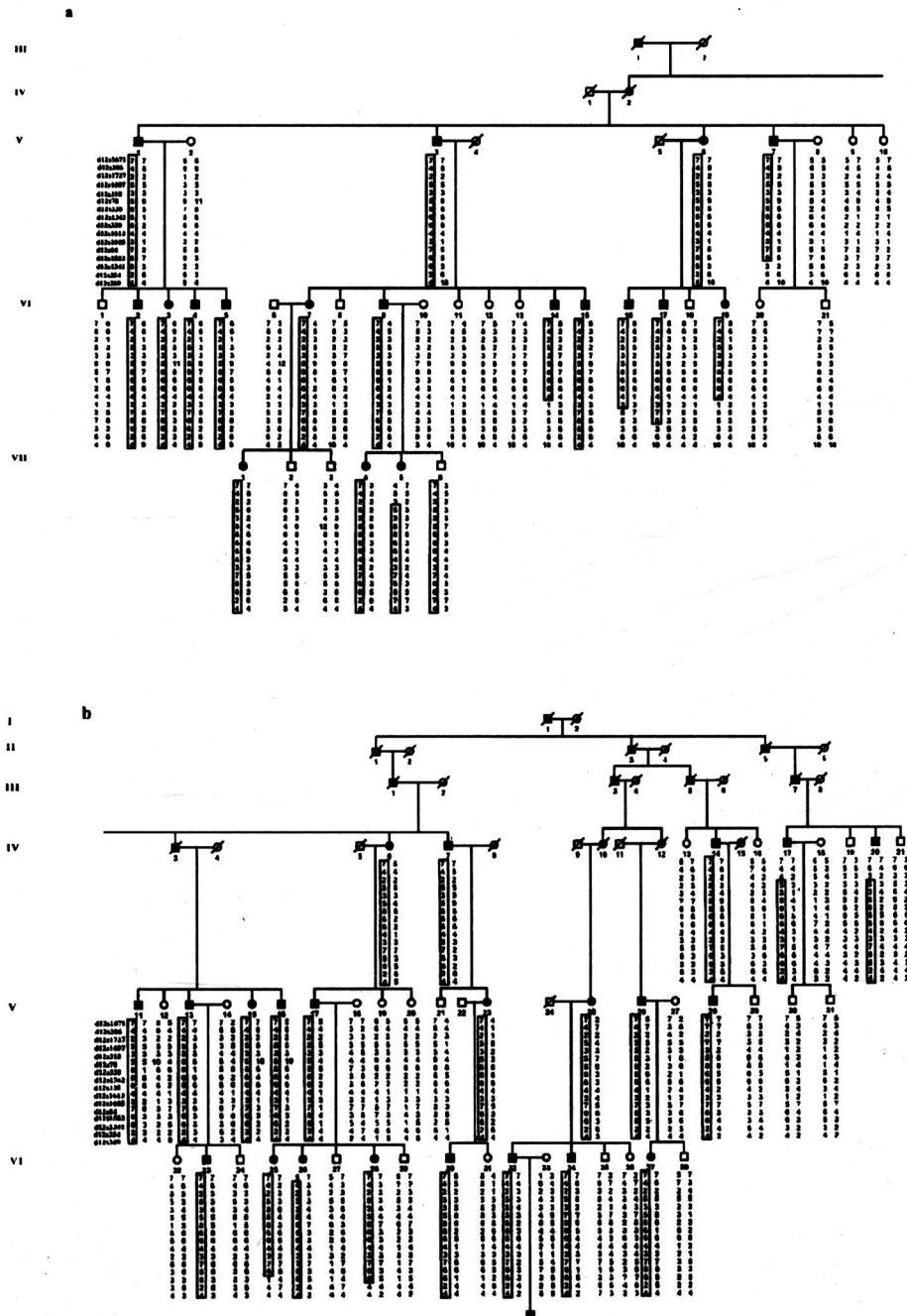


图 3 弥漫性浅表性光敏性汗孔角化症部分家系的连锁分析图

Fig.3 Haplotype analysis of the family with DSAP

用计算机将一个已知的基因与 180 多万个已完成测序的基因表达片段 (EST) 比较两者在 DNA 序列上的类似性。在同一物种内 (如人、鼠、果蝇), 如果发现某一 EST 片段在 200bp (base pair) 左右的核苷酸组成上与一个已知的基因有 60% ~

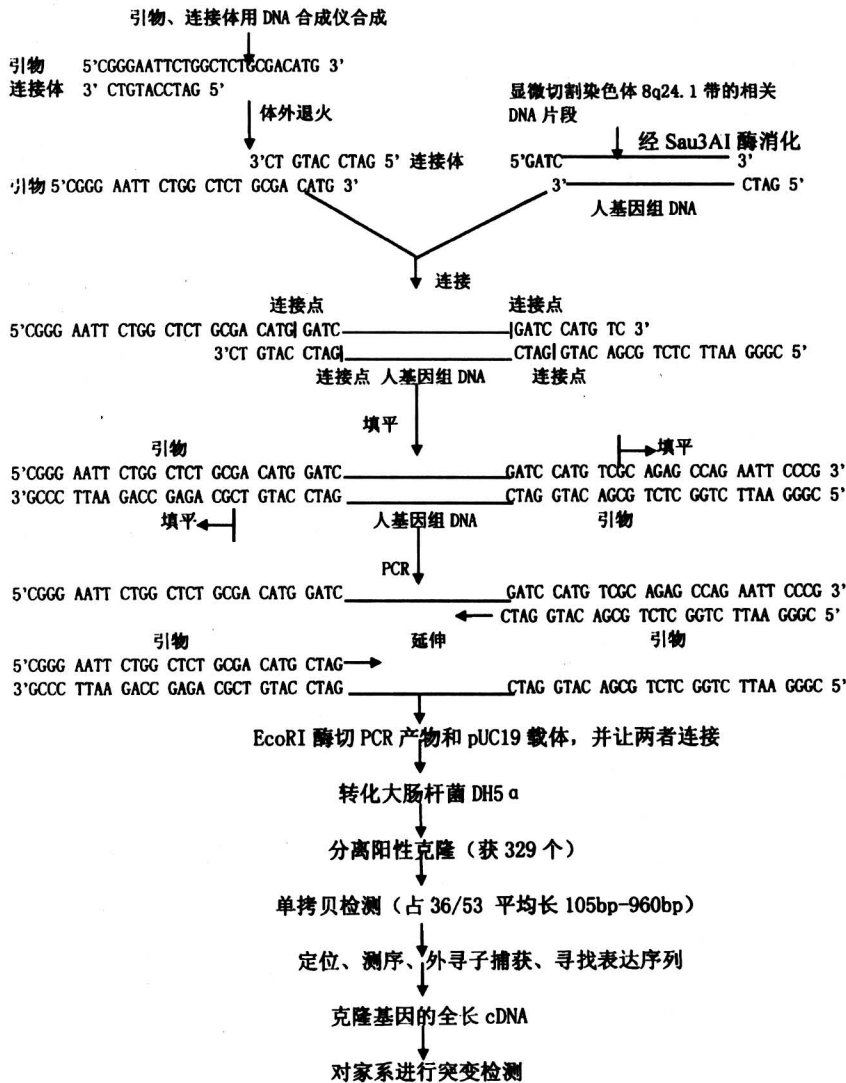
70% 的相似性, 则可认为找到了一个同类的新基因 (即同一基因家族中的一个新成员)。在不同物种如鼠与人之间, 鼠的某一疾病或功能基因在 200bp 左右的范围内与人的某一 EST 片段核苷酸序列有 80% ~ 90% 的同源性, 则可认为找到了一个人的新

基因，而且这种基因可能在人类中引起与老鼠相似的疾病或具有相似的功能；如果在人 EST 中找到的某一基因片段与老鼠仅 50% 的同源性，就可以认为可能找到了一个人的与老鼠的某一疾病或功能的同一类的基因。

不管是走路线 1 还是 2，均应继续用计算机寻找与之相连接的片段进行拼接，有时可以用计算机将一个未知基因完全拼接出来，然后用分子遗传学的方法作 1~2 次实验检测，运气好则可在 3~5 天内克隆到一个相关的基因。

表 2 显微切割、PCR、探针池及微克隆方法克隆 8q24.1 带外生性骨疣病疾病基因

Table 2 Exostosis associated gene cloning at chromosome 8q24.1 by microdissection, PCR, probe pool and microcloning



2.3 从人类已知的基因克隆新基因—GJB3 (CX31)基因的克隆

2.3.1 EST (Expressed Sequence Tag, 表达序列标签) 数据库的分析 将人的 GJA1 (Gap junction protein alpha - 1, CX43)、GJA4 (Gap junction protein alpha - 4, CX37)、GJA5 (Gap junction protein alpha - 5, CX40)、GJA7 (Gap junction

protein alpha - 7, CX45)、GJB2 (Gap junction protein beta - 2, CX26) 基因的编码区对 NCBI (National Center for Biotechnology Information, 美国国家生物技术信息中心) 的人 EST 数据库进行 BLAST (Basic Local Alignment Search Tool) 分析, 筛选出在 100bp 以上序列中 60% ~ 75% 一致的 EST 共 27 个, 排除已克隆基因后, 得到 6 个可能

代表人类新基因的 EST, 分别为 AA078777、AA079696、AA235826、AA400343、AA400185、AA505369; AA078777 与 GJB2 在 172bp 一致性 68%, 与 GJA5 在 184bp 一致性为 64%, 与 GJA4 在 182bp 一致性为 69%。AA079696 与 GJB2 在 114bp 一致性为 64%。AA078777 和 AA079696 构成的重叠群 (Contig) 与大鼠的 Gjb3 (Gap junction protein beta - 3, CX31) 在 645bp 一致性为 83%, 与小鼠 Gjb3 (Gap junction protein beta - 3, CX31) 在 645bp 一致性为 82% (图 7)。

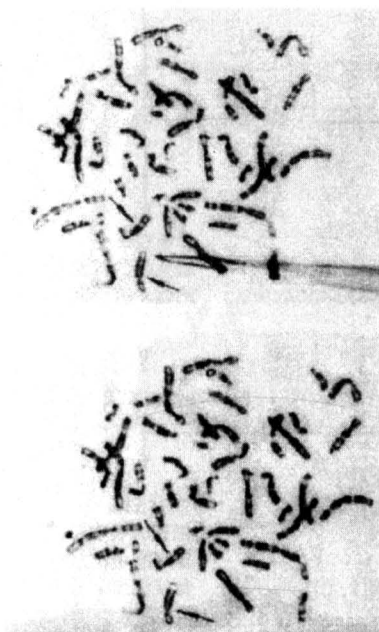


图 4 人染色体 8q24.1 的显微切割图

Fig.4 Microdissection of chromosome 8q24.1

2.3.2 全长 cDNA 的克隆 已有研究表明间隙连接蛋白在编码区均为单外显子, 故以基因组 DNA 为模板, 在 AA078777 和 AA079696 构成的重叠群设计引物 af/ar 和 bf/br 扩增, 产物测序、拼接得到 650bp 的序列, 与重叠群的一致性达到 95%, 并校正了 EST 序列。该重叠群经比较分析, 推测 5' 有起始密码子 ATG, 于是在它的 3' 设计引物 bc, bf 和 bc 进行 3' RACE (Rapid Amplification cDNA End, cDNA 末端的快速扩增)。在人的胎盘 cDNA (Complementary DNA) 文库中获得 400bp 的新序列, 与 650bp 的序列拼接成 1059bp 的序列。用 GCG (Genetics Computer Group) 软件 Frame 分析, 其中最大的 ORF (Open Reading Frame, 开放阅读框) 为 813bp, 编码 270 个氨基酸, 如图 7 所

示。在 1059bp 序列的 5' 和 3' UTR (Untranslated Region, 非翻译区) 设计引物扩增, 产物测序证实了拼接序列和阅读框的正确性。该 ORF 与小鼠 *Gjb3* 在 813bp 的一致性为 83.4%, 与大鼠 *Gjb3* 在 813bp 的一致性为 84.6%。推测的蛋白质分子量为 31KD, 与小鼠 *Gjb3* 在 270aa (Amino Acid) 的一致性为 82.6%, 与大鼠 *Gjb3* 在 270aa 的一致性为 83%。故认为它是鼠 *Gjb3* 的同源基因, 命名为人类 *GJB3*。

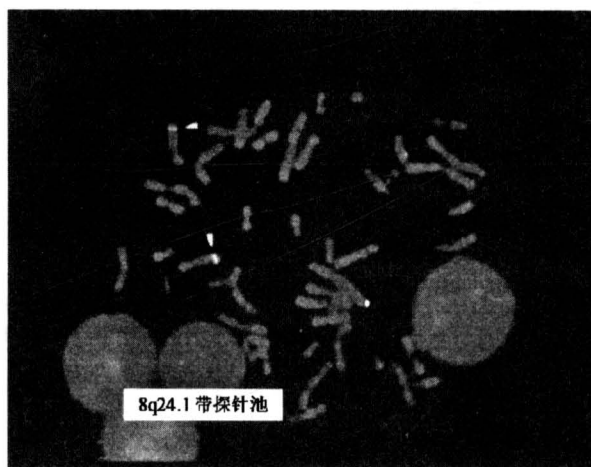


图 5 人染色体 8q24.1 显微切割的 PCR 产物的原位杂交图

Fig.5 FISH mapping of PCR products of microdissected human chromosome 8q24.1

2.3.3 *GJB3* 的定位 在 NCBI 的 STS (Sequence Tagged Site, 序列标记位点) 和 HTGS (High Throughput Genomic Sequence, 基因组大规模测序) 数据库中查询, 无定位的 STS 和大规模测序与 *GJB3* 相匹配。在 *GJB3* 的 3' UTR 设计引物扩增, 得到的产物作为探针, 筛选人基因组 DNA 文库, 得到插入子在 13kb 以上的克隆, 测序证实此克隆包含 *GJB3*。荧光原位杂交 (FISH) 显示 *GJB3* 基因定位于 1p34 (图 8)。

2.3.4 突变分析 人类间隙连接蛋白基因家系成员的突变导致多种遗传病, 如腓骨肌萎缩症 (Charcot - Marie - Tooth disease, CMT)、耳聋、白内障、心脏畸形。在 OMIM (Online Mendelian Inheritance in Man) 中查询定位在 1p34 的疾病有常染色体显性遗传的感觉神经性耳聋 (DFNA2) 和 CMT2A 型的腓骨肌萎缩症。我们在收集的神经性耳聋和腓骨肌萎缩症的病人中检测突变, 在两个

高频听力减退的浙江家系 A 和湖南家系 B 中分别发现错义和无义突变的杂合子 (图 9)。并用实验证明 Gjb3 在老鼠内耳中有强表达 (图 10)。选择

了非相关的 150 名个体进行检测, 未检测到错义和无义突变。

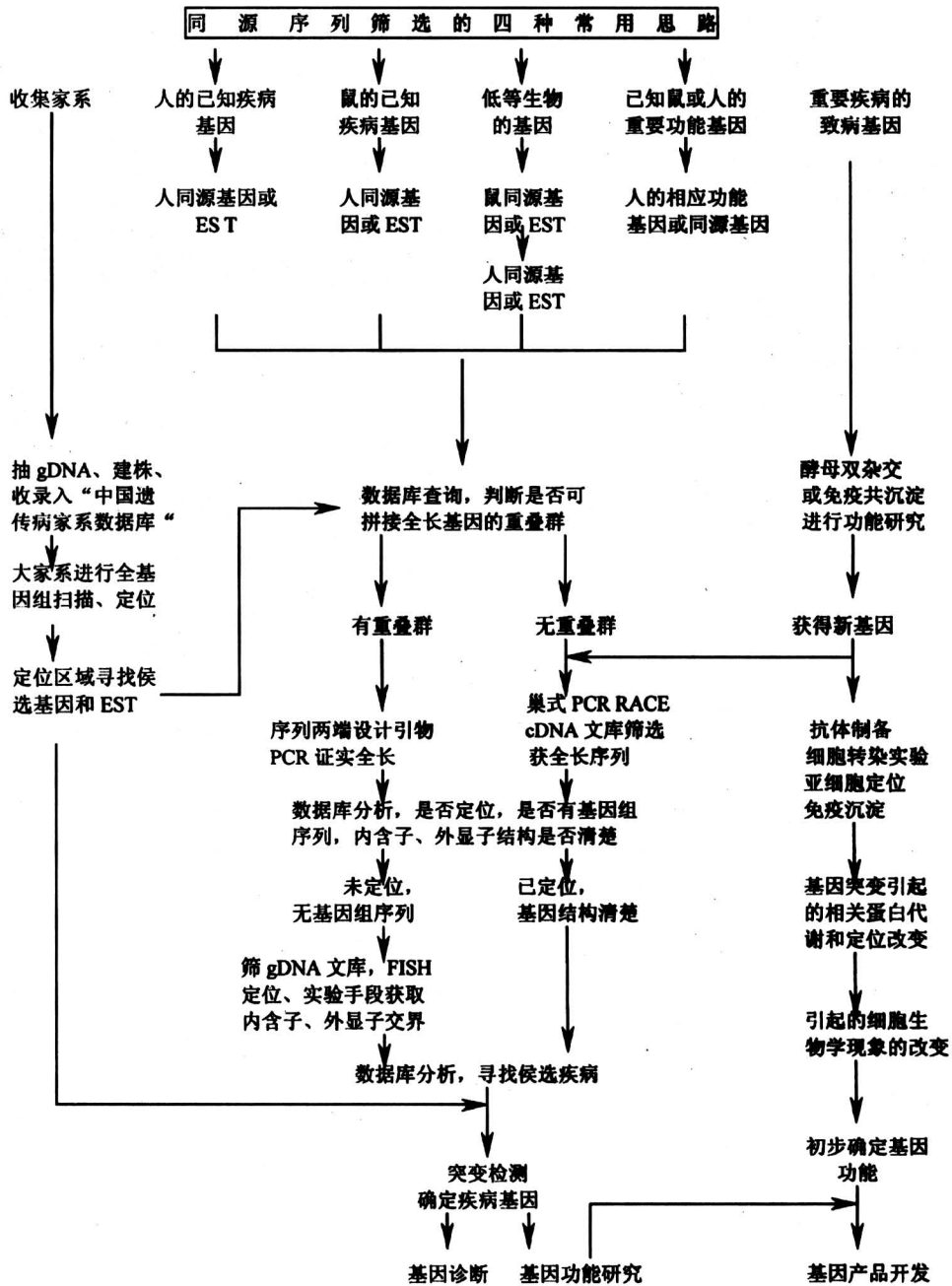


图 6 基因家系—候选疾病基因克隆示意图

Fig.6 Strategy of Gene family-Candidate disease gene cloning

2.4 从老鼠的疾病基因克隆人的同源基因 (CAC-NG3 基因的克隆)

2.4.1 EST 数据库分析 最新文献报道小鼠神经

元电压门控钙通道 γ -2 亚基基因 (*Cacng2*) 的内含子 2 插入一个约 6kb 的早期转座子导致头部摇摆和共济失调的失神性癫痫的 Stargazer 小鼠。在

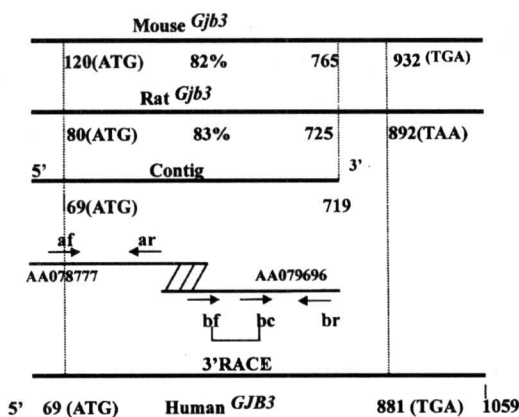


图 7 示 *GJB3* 基因的计算机克隆
Fig.7 Electronic cloning of *GJB3* gene

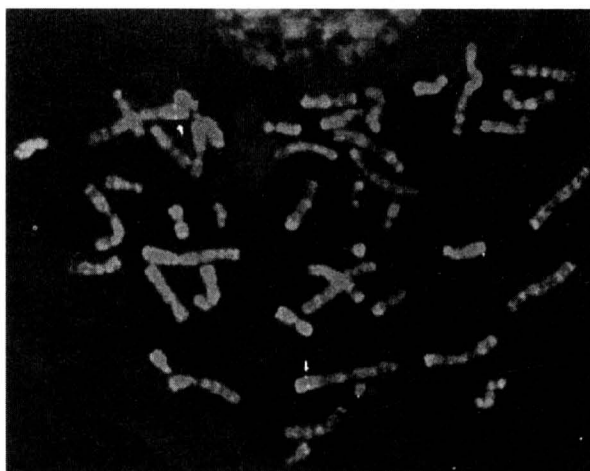


图 8 示 *GJB3* 基因定位
Fig.8 Gene mapping of *GJB3*

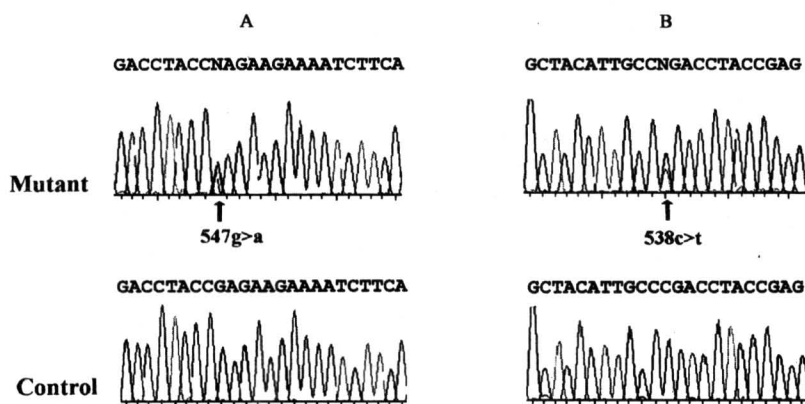


图 9 示在两个家系中发现 *GJB3* 基因突变 (mutant) 与正常对照 (control) 的测序图谱
Fig.9 Mutations of *GJB3* in two families: Part A and B show sequencing results of mutants and controls in Zhejiang and Hunan family respectively

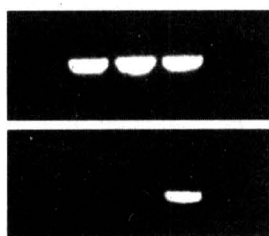


图 10 示 *Gjb3* 在小鼠内耳中的表达
Fig.10 Expression analysis of *Gjb3* in mouse inner ear

NCBI 的 Entrez 和 PubMed 中查询, 人的 *CACNG2* 基因未克隆。将小鼠 *Cacng2* 基因的编码区在

NCBI 的人 EST 数据库进行 BLAST 分析, 得到三个未克隆相应基因的 EST (W29095、H04905、H11833)。W29095 与 *Cacng2* 在 435bp 的一致性为 75%, H04905 与 *Cacng2* 在 337bp 的一致性为 77%; H11833 与 *Cacng2* 在 339bp 的一致性为 76%; H04905、H11833 与 W29095 的一致性达 90% 以上 (图 11)。

2.4.2 全长 cDNA 的克隆 在 W29095 设计一对引物 Caf/Car 在各种 cDNA 文库中扩增, 人脑前叶皮质 cDNA 文库的产物测序证实为 W29095 的序列。在 W29095 的 5' 和 3' 设计引物 CaA/CaB 和 CaC/CaD 做 5' 和 3' RACE。在人脑前叶皮质 cDNA 文库中的巢式 PCR 获得 620bp 的序列, 3'

RACE 获得 600bp 的序列，这两个序列与 W29095 拼接成一个 1545bp 的 cDNA 序列，GCG 软件 Frame 分析其中包含一个 948bp 的阅读框 (ORF)，编码 315 个氨基酸，如图 1 所示。在 1545bp 序列的 5' 和 3' UTR 区设计引物在人脑前叶皮质 cDNA 文库扩增，产物测序证实了拼接序列和阅读框的正确性。该 ORF 与小鼠的 *Cacng2* 在 948bp 一致性为 73%，推测的蛋白质在 323aa 一致性为 74.3%。该基因经 HUGO 命名委员会命名为 CACNG3。

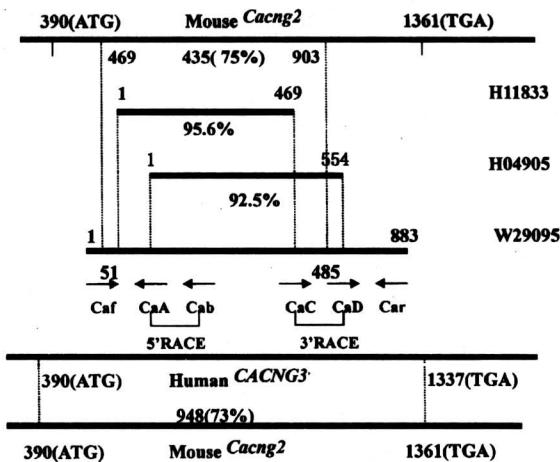


图 11 示 CACNG3 基因的计算机克隆

Fig.11 Electronic cloning of CACNG3 gene

2.4.3 基因的结构和定位 CACNG3 基因与 HT-GS 数据库做 BLAST 分析，它与定位于 16p12~p13.1 的大规模测序的 BAC 克隆 AC004125 序列 100% 一致，从而将 CACNG3 基因定位于 16p12-p13.1，并获得该基因的基因组结构，其编码区由 4 个外显子构成。

2.4.4 突变检测 定位于 16p12~p13.1 的疾病有常染色体显性遗传婴幼儿良性惊厥伴舞蹈手足徐动症和常染色体隐性遗传视网膜色素变性。在部分癫痫和视网膜色素变性家系中进行了突变分析，至今尚未检测到突变。

3 疾病基因功能研究

遗传病分两类，一类是染色体病，一类是基因病。染色体病是由于染色体的数目或结构异常所引起的疾病，至今已记载 100 余种，发病率达 0.7%。由于某一染色体或染色体某一片段的增加或减少涉及到成千上万个基因的增加或减少，其影

响重则导致胚胎不能发育，引起流产，轻则出生智力低下与畸形儿。基因病又分两类，由多个基因异常所引起的病称多基因病，如某些高血压、糖尿病等。由单个基因异常所引起的疾病称为单基因病，至今已记载 1647 种，已克隆疾病基因 975 个。多数单基因病患者在婴儿期、儿童期，甚至中壮年期均很正常，到 5 岁、10 岁、20 岁、30 岁甚至到 50 多岁左右才发病，而且有一个逐步加重和延续十几年或几十年最终导致死亡的病程。为什么到一定的年龄后才发病，这里涉及到某一基因的表达物的积累或这个基因的表达与其它基因间复杂的相互作用的问题。研究和解决这一问题是我们开展遗传病的诊断、预防和治疗必须回答的。对遗传病的诊断只要我们研究清楚了某一种遗传病的具体染色体畸变或基因突变，我们就可以通过染色体的分析或基因突变的分析来确诊。如染色体病，我们可以在产前宫内进行诊断，并用人工引产的方法阻止这类患儿的出生，如 21 三体、22 三体、18 三体及 9p 部分三体等；但基因病就有点不好办，对于较小年龄发病的疾病（进行性肌营养不良，DMD，5 岁左右发病）我们也可以采用产前诊断，引产阻止这类患儿出生；但对那些发病较晚而且较轻的疾病采取引产的办法就有可能扼杀大批有用之才，甚至天才。怎么办呢?! 可行的办法是通过基因治疗，导入一个正常的基因（如血友病等）一次性治愈此病，或用药物治疗延缓或阻止发病，或减轻症状。特别是药物治疗，它的根基就是基因功能的研究。研究基因表达的规律、基因与基因间的相互作用如何引起细胞或组织逐步损伤以至最终导致细胞或组织病变的，明白这其中的关系，找出延缓或阻断发病的药物，这已成为我们面临的基因功能组学研究的一个极其重要的方面。

3.1 遗传性神经性耳聋基因 GJB3 的功能研究

GJB3 是本实验室于 1998 年克隆的一个遗传性耳聋的疾病基因，是一个 connexin 家系的基因。利用酵母双杂交 (Yeast - Two Hybrid) 的方法，我们发现 P11 蛋白与 GJB3 有显著的相互作用，并利用免疫共沉淀在体外进一步证实了这一相互作用。

3.2 帕金森氏病 (Parkinson' s Disease, PD) 研究

帕金森氏病是一种重要的神经退行性疾病，是由于脑黑质的神经元凋亡而导致的。Parkin 是 1998 年克隆的一个 PD 的疾病基因，我们的研究发

现一个酮戊二酸载体 (PB16) 与之有相互作用, 而线粒体功能失调正好与 PD 的发生有关, 我们对于 PB16 的研究有可能阐明其引起 PD 的分子基础。

3.3 早老性痴呆 (Alzheimer' s Disease, AD) 和 APC 基因的研究

早老性痴呆也是一种重要的神经退行性疾病。张灼华教授的研究首次发现该病的一个疾病基因 Presenilin - 1 的一种功能, 即突变的 Presenilin - 1 失去正常的调节信息传递作用, 使 β -catenin 失稳定而加速神经元的死亡。我们现正在研究 APC 基因与 AD 的关系。利用 Yeast - Two Hybrid, 我们已经找到一个与之相互作用的新基因 CH, 并克隆了一个基因家系 ABP (CH 是其中一个成员)。进一步的研究发现 ABP 与细胞内蛋白质转运有关, 对这些基因在细胞内的基本功能及其调节的方式的研究, 发现了一个新的细胞内蛋白调节的通道, 为在分子水平上研究肿瘤抑制作用和神经传导过程找到了一新的切入点。

参考文献

- [1] Xia Jiahui. Mutation in the gene encoding gap junction protein β -3 associated with atosomal dominant hearing impairment [J]. Nature Genetics, 1998, 20 (12): 370~373
- [2] Gabriele Richard, Lisa E. Smith. Mutation in the human connexin gene GJB3 cause erythrokeratoderma variabilis [J]. Nature Genetics, 1998, 20 (12): 366~369
- [3] Xia Jiahui. Identification of a locus for disseminated superficial actinic porokeratosis (DSAP) at Chromosome 12q23.2 - 24.1 [J]. Journal of Investigative Dermatology, 2000, 114 (6): 1071~1074
- [4] Deng Hanxiang. Molecular cloning of a candidate gene for hereditary multiple exostoses type II [J]. Progress in Natural Science, 1996, 6 (6): 692~699
- [5] Xu Lei, Xia Jiahui. Mutation analysis of hereditary multiple exostoses in the Chinese Hum Genet. 1999, 105: 45~50

Study on Family Collection, Gene Mapping, Gene Identification and Gene Function of Human Genetic Diseases

Xia Jiahui

(National Laboratory of Medical Genetics of China, Changsha 410078, China)

[Abstract] In this paper, the study on family collection, gene mapping, gene identification and gene function of human genetic diseases carried out in the National Laboratory of Medical Genetics of China, were described in detail. Using G-banding technique a marker chromosome t (1; 3) (q44; p11) associated with nasopharyngeal cancer was found in 1975 at first, and human TDF gene was mapped to chromosome Yp11. 32 in 1981. Since 1991, 590 families with 345 kind of genetic diseases were collected. In 1998, GJB3, a human genetic neurological deafness gene, was identified using a novel strategy of "Gene Family-Candidate Disease Gene Cloning", and the paper was published in Nature Genetics (20: 370). In 1999, by linkage analysis and Genome Wide Scanning, a locus responsible for disseminated superficial actinic porokeratosis (DSAP) was identified at Chromosome 12q23. 2 - 24. 1; and at the same year a novel protein trafficking gene was also cloned from gene function study.

[Key words] genetic disease; gene mapping and cloning; gene family-candidate disease gene cloning; genome wide scan; gene function study