
博士学位论文

遗传性非综合征耳聋(DFNA52)候选基因的生物信息学 分析与突变检测及兜甲蛋白角化病的研究

作者姓名： 潘 琼
学科专业： 遗 传 学
学院(系、所)： 中国医学遗传学
国家重点实验室
指导教师： 夏家辉 院士

中 南 大 学
2007 年 12 月

分类号 VDC_____

密级_____

博士学位论文

遗传性非综合征耳聋 (DFNA52) 候选基因的生物信息学 分析与突变检测及兜甲蛋白角化病的研究

Bioinformatics Analysis and Mutation Screening of the
Candidate Genes for DFNA52 Locus & Loricrin Keratoderma
Caused by 730insG Mutation in Loricrin Gene

作者姓名： 潘 琼
学科专业： 遗 传 学
学院(系、所)： 中国医学遗传学
 国家重点实验室
指导教师： 夏家辉 院士
导师小组： 夏 昆 教授
 潘 乾 副教授

论文答辩日期_____

答辩委员会主席_____

中 南 大 学
2007 年 12 月

第一部分 遗传性非综合征耳聋 (DFNA52) 候选基因的生物信息学分析与突变检测

中 文 摘 要

耳聋(deafness)是各种不同程度听力损失(hearing loss)的总称。听力损失是最常见的感觉障碍。在世界范围内, 每1000名新生儿中就有1名先天性耳聋患儿, 50%患儿的耳聋与遗传因素有关。在遗传性耳聋中, 约70%为非综合征型耳聋(nonsyndromic hearing loss, NSHL)。NSHL具有高度的遗传异质性, 到目前为止, 已定位了100多个疾病相关基因位点, 其中40个疾病相关基因已被克隆。

本实验室收集到一个湖南怀化的遗传了5代的常染色体显性遗传的非综合征耳聋家系。通过全基因组扫描将该家系致病基因定位于5q31.1-q32, 两点连锁分析在D5S2017得到最大LOD值6.89, 精细定位及单体型分析提示该病基因位于D5S2056和D5S638之间8.8cM的区域。定位于5q31的两个耳聋基因*DIAPH1* (OMIM:602121)和*POU4F3* (OMIM:602121)经突变检测未发现突变。提示本研究发现了一个新的常染色体显性遗传进行性感神经性耳聋位点 (DFNA52)。DFNA52的定位区间与DFNA54和DFNB60定位区间有重叠, 很有可能在5q31另有一个或多个非综合征型耳聋基因。

已有研究显示导致常染色体显性遗传的非综合征型耳聋的基因是不同的, 许多基因参与了听觉形成过程, 很难根据一个基因的已知

功能或推测的功能来选择候选基因。根据 NCBI 数据库的注释, DFNA52 区间内共有 173 个 RefSeq Genes, 199 个 Know Genes。随着新的分子生物学技术的发展, 人类基因组计划的完成, 特异性耳蜗特异性 cDNA 文库数量的增加以及小鼠耳聋动物模型的建立, 以及可获得的生物信息数据库的激增, 为克隆新的遗传性耳聋基因提供了最可能的候选基因。我们应用生物信息学方法分析 DFNA52 在 D5S2056 和 D5S638 之间的候选基因, 通过分析 Morton 耳蜗 EST 数据库, 耳发育相关基因表达谱数据库, NEIBank 耳文库以及 Hereditary Hearing Impairment in Mice 和 Deaf Mouse Mutants 等相关数据库, 综合耳蜗基因表达谱, 基因功能, 耳聋临床表现特征, 小鼠耳聋致病基因以及识别疾病候选基因的生物信息学软件(如 G2D, DGP, PROSPECTR, GeneSeeker)的预测结果等资料, 建立了识别耳聋候选基因的策略, 在 DFNA52 区间内优先选择了 52 个候选基因, 包括两个小鼠耳聋基因 *NEUROG1*, *SMAD5*。

在对 52 个候选基因的编码区、剪切位点、部分内含子区和部分调控区的突变检测中, 15 个基因未发现任何序列变异, 在其余 37 个候选基因共发现 108 个 SNPs, 其中包括杂合变异和纯合变异。查询单核苷酸多态数据库 dbSNP, 已报道的 SNPs 有 94 个, 未收录的多态位点有 15 个。经再对家系内其他患者、家系内正常人及家系外正常对照的突变分析, 结果表明所有 SNPs 与疾病表型无共分离。尽管从基因表达和基因功能上来看 *NEUROG1*, *SMAD5*, *SLC23A1*, *PCDHB5*, *PCDHB11*, *PCDHAC1*, *PCDHAC2*, *PCDHB3* 和 *CTNNA1* 等

基因是 DFNA52 位点区间内的最有可能与耳聋相关的基因，但突变检测结果却仅在这些基因中发现一些 SNPs，未找到致病性突变。这说明这些基因可能均不是 DFNA52 的致病基因，其致病性突变可能存在于基因的非编码区，或新的致病基因有待于发现。

关键词： 耳聋，非综合征耳聋，突变检测，候选基因，生物信息学

第二部分 LOR 基因 730insG 突变导致的兜甲蛋白角皮病

中 文 摘 要

兜甲蛋白角皮病(loricrin keratoderma)是一个新的疾病术语, 由 *LOR* 基因突变所引起的遗传性皮肤病现统一命名为兜甲蛋白角皮病。兜甲蛋白角皮病的表型具有多样性。临床上, 这些患者被诊断为伴随鱼鳞病的 Vohwinkel 综合征(残毁性掌跖角皮病, OMIM 604117), 进行性对称性红皮肤角化病 (progressive symmetric erythrokeratoderma, PSEK), 或出生时为火棉胶婴儿的先天性鱼鳞病样红皮病。

本研究收集到一个组织学特征高度相似的兜甲蛋白角化病家系。应用基因组扫描方法,利用 1 号染色体上的微卫星标记对该家系进行连锁分析。在 D1S2696 得到最大两点连锁 LOD 值 3.46($\theta=0$), 单体型分析将疾病基因定位在 1 号染色体 1p13.3-q21.3 的 D1S2726-D1S305 之间约 15cM 范围内, 表皮分化复合物基因(EDC)在此区间内。表皮分化复合物基因(EDC)与表皮的终末分化密切相关, 选择表皮分化复合物基因的成员丝聚合蛋白(Filaggrin)与兜甲蛋白(Loricrin)为候选基因, 因已报道 *FLG* 突变导致寻常性鱼鳞病, *LOR* 突变导致鱼鳞病型 Vohwinkel 综合征。

对候选基因 *FLG* 的部分编码区及外显子与内含子交界处进行突变检测, 但 *FLG* 测序结果未发现有致病性突变, 因此 *FLG* 基因不是

该家系的致病基因。*LOR* 基因突变检测结果发现在 *LOR* 基因的核苷酸 730 有单个核苷酸的插入(730insG)。*LOR* 羟基末端连续 6 个 G 碱基后单个 G 碱基插入 (730insG) 导致其后的甘氨酸环及高谷氨酰胺-赖氨酸区翻译编码框移位, 改变了其后的 84 个氨基酸密码子的编码, 并使终止密码子延迟出现, 导致突变型蛋白较野生型延长 22 个氨基酸。

查阅相关文献, 已报道在 6 个家系中发现了 *LOR* 基因 730insG 突变, 不同家系的兜甲蛋白角皮病患者的临床表型具有多样性。在本研究的鱼鳞病家系中 II3 和 III7 两个患者的小指上可见轻度的环状狭窄, 且 II3 出现趾断症(pseudoainhums), 自行离断(autoamputation), 其它患者表现为弥漫性鱼鳞病皮肤病。我们认为兜甲蛋白角皮病的两个特征是蜂窝状掌跖角化病和弥漫性鱼鳞病样皮肤病, 兜甲蛋白角皮病可描述为“一种由于兜甲蛋白突变导致的常染色体显性遗传的伴随不同临床表现的鱼鳞病样和蜂窝状掌跖角化的皮肤病”。

关键词: 兜甲蛋白, 兜甲蛋白角皮病, 鱼鳞病, 突变检测

PART I Bioinformatics Analysis and Mutation Screening of the Candidate genes for DFNA52 locus

ABSTRACT

Hearing impairment is one of the most common human sensory defects, affecting approximately one out of 1000 children at birth. In more than half of these cases, the defect can be attributed to a genetic cause. 80% of hereditary hearing impairment is categorized as non-syndromic, based on the absence of other symptoms. To date, more than 100 loci for non-syndromic sensorineural hearing loss have been mapped and 40 genes coding for a wide variety of proteins have been identified.

we mapped a new novel DFNA locus(DFNA52; OMIM 607683) to a 8.8 cM region of chromosome 5q31.1-q32 by a genome-wide scan. The phenotype was characterized by bilateral, postlingual, progressive, and sensorineural nonsyndromic hearing impairment involving all frequencies. Fine mapping indicated that the disease gene was located within an 8.8-cM region between markers D5S2056 and D5S638, with a maximum two-point logarithm of differences (LOD) score of 6.89 ($\theta=0$) at D5S2017. No mutations could be detected in either the POU4F3 (OMIM:602460) or DIAPH1 (OMIM:602121) genes, both of which map to 5q31. What's more, the genetic interval of DFNA52 overlaps with the

regions for DFNA54 and DFNB60; it is likely that there is another gene in 5q31 associated with NSHHI.

Of the 152 named genes and 26 uncharacterized cDNAs in DFNA52 interval. Because those genes responsible for ADNSSHL are diverse including members of the gap junction family, cytoskeletal proteins, transcription proteins, ion channels, motor molecule, and genes of unknown function, it is difficult to select candidate gene in large numbers of genes. Candidate gene approaches are based on the study of genes selected through knowledge of the gene expression, function, the disease phenotype, genetic studies in mouse model(phenotypes, expression), and also from literature, also use some bioinformatics methods for identifying candidate disease genes, such as G2D, DGP, PROSPECTR, GeneSeeker. 52 genes were considered to be potential functional candidate genes for DFNA52. We selected 52 genes as candidate genes from DFNA52 interval, including the mouse deafness gene (NEUROG1, SMAD5) .

In total, Mutation detection for 52 candidate genes located in the critical region of DFNA52 was performed in two patients from pedigree. No disease-causing mutation segregating with the disease in coding and splice site regions of these genes was found, except for some single nucleotide polymorphisms (SNPs) in these patients. 108 SNPs were identified, of which 15 were novel. Eleven of the 108 SNPs changed the

encoded amino acid. Although these genes(Neurog1, SLC23A1, PCDHB5, PCDHAC1, PCDHAC2, PCDHB3 and CTNNA1 et.al)are plausible candidate genes in the region of DFNA52 locus due to their gene functions and expression profiles, but the results of mutation analysis found no pathogenic mutation. As a rule,these findings suggest those genes are unlikely to be involved in DFNA52.Maybe it is in untranslated region of the pathogenic mutation in deafness family, or it was a new gene which is responsible for the disease.

KEY WORDS: deafness, candidate gene, mutation detection, bioinformatics

PART II An Ichthyosis vulgaris phenotype of Loricrin keratoderma caused by 730insG mutation in Loricrin gene

ABSTRACT

Mutations in the loricrin gene were found as a cause of Vohwinkel syndrome with ichthyosis (OMIM 604117), progressive symmetric erythrokeratoderma (PSEK) or congenital ichthyosiform erythroderma born as a collodion baby. These disorders which mutations are insertions of a single nucleotide in LOR resulting in a frameshift have been collectively termed as loricrin keratoderma (LK).

We collected a Chinese Vohwinkel syndrome with ichthyosis family with highly similar histological features. Linkage analysis was performed by using STR markers in chromosome 1. A maximum two-point Lod score of 3.46($\theta=0$) was obtained at D1S2696. Haplotype analysis placed the critical region in a 15-CM interval defined by D1S2726 and D1S305, and localized the disease locus to chromosome 1 p13.3-q21.3.

To identify the disease gene, we screened for mutation in Profilaggrin(FLG) and Loricrin(LOR) gene. Because profilaggrin and loricrin are major component of the epidermal cornified cell envelope, and mutations in FLG, LOR gene are responsible for ichthyosis vulgaris, ichthyotic variant of Vohwinkel syndrome. No mutations were identified

in the non-filaggrin repeat coding regions and the promoter. So, the FLG gene may not undertake the pathology of this ichthyosis family. However, in the LOR gene after nucleotide 730 (730insG), an insertion of a G residue was found in an area of six normally occurring G residues(codons 230-231) of the loricrin gene, introducing a frameshift at codon 232 that altered the terminal 84 amino acids and produced a delayed termination codon that extended the mutant protein by 22 amino acids.

In the ichthyosis family, mild constricting bands encircling fingers are seen in II3,III7 and pseudoainhums, autoamputation are seen in II3. Other patients all have honeycomb palmoplantar keratoderma and diffuse ichthyosiform dermatosis. we suggest that loricrin keratoderma should be described as ‘autosomal dominant honeycomb palmoplantar keratoderma with ichthyosis heterogeneous in clinical appearance which caused by loricrin mutations’.

KEY WORDS: loricrin, Loricrin keratoderma, Vohwinkel syndrome with ichthyosis , Mutation detection

目 录

中 文 摘 要.....	IV
ABSTRACT.....	IX
英文缩略词表.....	XV
主要实验仪器、试剂与耗材.....	XVI
第一部分 遗传性非综合征耳聋 (DFNA52) 候选基因的生物信息学分析与突变检测.....	1
前 言.....	1
第一章 生物信息学方法分析 DFNA52 的候选基因.....	14
1.1. 材料与方法.....	14
1.2. 结果.....	23
1.3. 分析与讨论.....	26
第二章 DFNA52 的候选基因的突变检测.....	28
2.1. 材料与方法.....	28
2.2. 结果.....	30
2.3. 分析与讨论.....	35
第一部分 结论.....	37
参考文献.....	38
第二部分 LOR 基因 730insG 突变导致的兜甲蛋白角化病.....	47
前 言.....	47
3.1. 材料与方法.....	52
3.2. 结果.....	58
3.3. 讨 论.....	63
第二部分 结 论.....	66
参考文献.....	67
综述.....	70
参考文献.....	86
附录.....	94
致 谢.....	101
攻读博士学位期间主要的研究成果.....	102

英文缩略词表

SHL	Syndromic Hearing Loss	综合征型耳聋
NSHL	Non-syndromic Hearing Loss	非综合征型耳聋
DFNA	Autosomal Dominant nonsyndromic hearing impairment	常染色体显性遗传非综合征耳聋
DFNB	Autosomal recessive nonsyndromic hearing impairment	常染色体隐性遗传非综合征耳聋
SNPs	Single Nucleotide Polymorphisms	单核苷酸多态
3'UTR	3' Untranslated Region	3'非翻译区
5'UTR	5' Untranslated Region	5'非翻译区
Blast	Basic Local Alignment Search Tool	基本局部比对搜索工具
BLAT	BLAST-Like Alignment Tool	Blast 样比对工具
cds	Coding sequence	编码区
dbSNP	The Single Nucleotide Polymorphism Database	NCBI 单核苷酸多态数据库
gDNA	Genomic DNA	基因组 DNA
NCBI	National Center for Biotechnology Information	美国国立生物技术信息中心
OMIM	Online Mendelian Inheritance in Man	在线人类孟德尔遗传数据库
PCR	polymerase chain reaction	聚合酶链反应
RefSeq	NCBI Reference Sequence	NCBI 的参考序列数据库
rSNP	Regulatory single nucleotide polymorphisms	调控性单核苷酸多态
SAP	Shrimp Alkaline Phosphatase	虾碱性磷酸酶
IV	Ichthyosis vulgaris	寻常型鱼鳞病
IVS	ichthyotic variant of Vohwinkel's syndrome	鱼鳞病型 Vohwinkel 综合征
PSEK	progressive symmetric erythrodermatoderma	进行性对称性红皮肤角化病
FLG	Filaggrin	丝聚合蛋白
LOR	loricrin	兜甲蛋白
LK	loricrin keratoderma	兜甲蛋白角皮病

主要实验仪器、试剂与耗材

主要实验仪器

420 型恒温空气摇床	Forma 公司
9600 PCR 仪	Perkin Elmer 公司
9700 PCR 仪	Perkin Elmer 公司
ABI 377 型 DNA 测序仪	Perkin Elmer 公司
Centrifuge5415D	Eppendorf 公司
CO ₂ 培养箱	Forma 公司
DU®640 紫外分光光度仪	Beckman 公司
Eppendorf 台式低温微量离心机	Eppendorf 公司
Gel Doc1000 凝胶成像分析系统	Bio-Rad 公司
GS-700 扫描仪	Bio-Rad 公司
J6-HC 高速离心机	Beckman 公司
PTC-200 PCR 仪	MJ Rsearch 公司
Rotator	Appligene 公司
超净工作台	苏州净化工程公司
低电压电源电泳仪	Appligene 公司
隔水式电热恒温培养箱	上海跃进医疗器械厂
水平电泳仪	Bio-Rad 公司
移液器	Eppendorf 公司

主要实验试剂及耗材

Hotstar polymerase	Takara 公司
0.2ml PCR 管	Axygen 公司
0.5 ml Eppendorf 管	Axygen 公司
0.5-10 µl tip	Axygen 公司
1.5 ml Eppendorf 管	Axygen 公司
1.5ml Eppendorf 管	Eppendorf 公司
10 ml 一次性移液管	Falcon 公司
100-1000 µl tip	Axygen 公司
10-100 µl tip	Axygen 公司
15 ml 一次性离心管	Falcon 公司
15 ml 一次性离心管	Falcon 公司
50 ml 一次性离心管	Falcon 公司
AgNO ₃	上海试剂一厂
DMEM	GIBCO 公司
DNA marker	TakaRa 公司
dNTPs	Takara 公司
EDTA	Sigma 公司
Exonuclease I	New England Biolabs 公司

N,N-亚甲基丙烯酰胺	Bio-Rad 公司
NaOH	苏州桃源新亭化工厂
SAP	Roche 公司
SDS	New England Biolabs 公司
SPIN 柱子	QIAGEN 公司
T4 DNA 连接酶	TakaRa 公司
Taq DNA polymerase	Takara 公司
TEMED	Sigma 公司
冰醋酸	中国医药集团上海化学试剂公司
丙三醇	上海祥德精细化工厂
丙烯酰胺	Bio-Rad 公司
多聚-D-赖氨酸	Sigma 公司
多聚甲醛	Sigma 公司
肝素钠注射液	南京新天生物化学制药有限公司
过硫酸胺	Promega 公司
甲醛	湖南师大化学试剂厂
普通琼脂糖、低熔点琼脂糖	Biorad 公司
三氯甲烷	湖南师大化学试剂厂
无水乙醇	湖南师大化学试剂厂

主要试剂配制

名 称	各 组 成 成 分		用途
10×Tris- 硼 酸 缓 冲 液（TBE）	Tris 碱	108g	聚丙烯酰胺凝胶和琼脂 糖胶电泳缓冲液
	硼酸	55g	
	0.5M EDTA（pH8.0）	40ml	
	加水至	1000ml	

第一部分 遗传性非综合征耳聋(DFNA52)候选基因的生物信息学分析与突变检测

前 言

耳聋概况

耳聋(deafness)是各种不同程度听力损失(hearing loss)的总称。外耳和中耳传音结构、内耳感音器官及听觉神经通路上任何部位及附近的病变都可能引起耳聋。耳聋是一种严重影响人类生活质量的常见先天性疾病,它可由单一基因突变或不同基因的复合突变引起,也可由环境因素(如医疗因素,环境暴露,创伤,药物等)或基因和环境两者共同作用而致。

听力损失是最常见的感觉障碍。在世界范围内,每 1000 名新生儿中就有 1 名先天性耳聋患儿,50%患儿的耳聋与遗传因素有关^[1,2]。据中国残疾人联合会网站统计,中国 6000 万残疾人口中 2100 万为听力残疾者,居残疾人总数的首位。据人口调查统计,每年出生的新生儿中,重度听力障碍者约占 1‰~3‰。另外,随着全球人口老龄化加速,耳聋的绝对发病率正在显著增高。随年龄增长听阈差距增大,不同程度听力障碍者在人群中的比例,45~64 岁为 14%,65~75 岁为 30%,75 岁以上者高于 50%^[1,3]。尽管近年来遗传性耳聋研究领域有许多重要进展,但耳聋相关的基因的定位、分离、克隆及基因的突变还有待于更深入的研究。

遗传耳聋的分类

导致听力损害的原因众多,但大致可以分为两大类:环境相关因素和遗传相关因素。前者包括耳毒性药物(如氨基糖甙类等)的使用,头部产伤、外伤和肿瘤,各类细菌和病毒感染(如出生前巨细胞病毒、弓形体和风疹病毒等感染,出生后中耳炎和脑膜炎等),长期处于噪声环境以及早产、新生儿缺氧、低体重儿、严重的新生儿黄疸等。除以上原因外其中一半以上的听力损害是遗传因素所致。

遗传性耳聋是指由来自亲代的致聋基因或新发生的突变致聋基因所导致的耳部发育异常或代谢障碍,以致出现听力障碍。遗传性耳聋根据是否伴有耳外组织的异常或病变可分为综合征性耳聋(syndromic hearing loss, SHL)和非综合征性耳聋(nonsyndromic hearing loss, NSHL),其中 SHL 占 30%左右,而 NSHL 约占 70%,非综合征耳聋是最常见的感音神经性聋。在 NSHL 中,根据遗传方式不同将遗传性耳聋分为常染色体显性(DFNA, 15% ~ 20%)、常染色体隐性(DFNB, 80%)、性连锁(DFN X-linked, 1%)和线粒体遗传性耳聋(线粒体遗传所占的比例在不同的人群中从 1~20%之间不等)四类,见图 1-1。

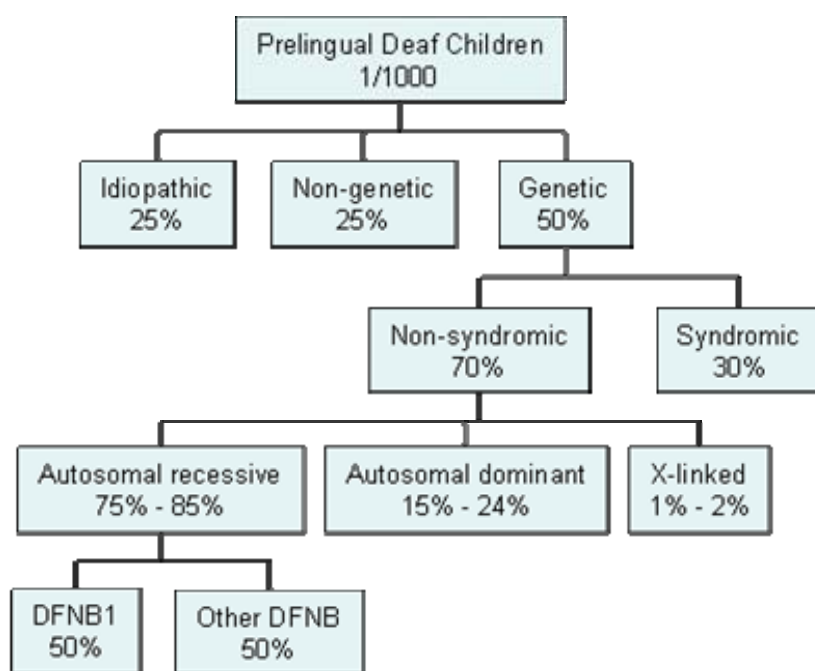


图 1-1 在儿童中学语前聋的病因分类

另外，耳聋的分类有很多种，比如根据其病变部位、性质、程度和病因进行不同的分类。临床最常用的是按病因及病变的部位分类，可分为传导性聋(conductive hearing loss)、感音神经性(sensorineural hearing loss)和混合性聋(mixed hearing loss)三种。传导性聋指因外耳、中耳或咽鼓管病变，以及迷路骨性被囊的耳硬化症导致声波不能传入内耳所致的听力损失。感音神经性聋为内耳、耳蜗神经、脑干听觉通路及听觉中枢病变所致听力损失的统称。混合性聋是指传导性和感音神经性聋同时存在。

耳聋的程度根据纯音测听所测得的言语频率听阈的平均值进行分类，可分为轻度聋(26-40dB)，中度聋(41-55dB)，中重度聋(56-70dB)，重度聋(71-90dB)和极度聋(90dB 以上)。

根据发病时间可分为先天性聋和后天性聋，语前聋和语后聋。先天性聋指出生时或出生后不久就已存在的听力障碍，是胚胎期或围生期致聋因素所致，其病因可分为遗传性聋和非遗传性聋两大类。语前聋指重度先天性聋或在婴幼儿期即失去听力者，由于不能通过语音交流接受言语信号，更无自身言语反馈，导致言语发育障碍，最终成为聋哑，多在出生时或一岁时发病，一般为非进行性。在语前聋中，常染色体隐性遗传是主要的遗传方式，其危害也最为严重。语后聋指在言语形成之后失去听力者（一般在 6-7 岁以后发病）。

耳聋基因的克隆与分类

遗传性耳聋在 16 世纪就有文献记载,但由于内耳部位深,体积小,研究手段受到限制,因此认识和理解控制听觉系统的基因研究进展很慢。近 10 多年来,随着现代科学技术的发展及分子生物学、遗传学的飞速发展,人们对遗传性聋的认识不断加深,并取得了显著进步,与耳聋相关的基因的定位、分离、克隆及基因的突变得到了深入的研究。

遗传性听力损害具有高度的遗传异质性,根据目前的研究表明与 NSHL 相关的位点有 120 多个 (<http://webhost.ua.ac.be/hhh/>),并根据遗传方式和定位的先后进行命名,以 DFNA、DFNB、DFN 和 DFNY 分别代表常染色体显性、常染色体隐性、X 连锁和 Y 连锁 NSHL 位点。有人估计,非综合征型耳聋相关的位点可能超过 100 个,而综合征型耳聋的位点可能多达 400 个^[5]。

截止 2007 年 9 月,已定位了 77 个非综合征型耳聋位点,其中常染色体显性遗传 50 个,常染色体隐性遗传 67 个,X-连锁遗传 5 个,Y-连锁遗传 1 个。并且克隆了 22 个常染色体显性遗传 NSHL 基因,25 个常染色体隐性遗传 NSHL 基因,2 个 X-连锁遗传 NSHL 基因^[4](Hereditary Hearing Loss Homepage, <http://webhost.ua.ac.be/hhh/>) (如表 1 所示),其中 7 个基因突变既可以导致 DFNA 也可以导致 DFNB,另外还有 11 个 NSHL 基因突变也可以导致 SHL。对于一个特定的 NSHL 位点,可能不止一个耳聋致病基因,如 DFNA2(*GJB3*, *KCNQ4*)和 DFNA3(*GJB2*, *GJB6*)。这提示染色体上特定区域有成簇耳聋基因的存在。

众多的位点被定位也说明了 NSHL 具有相当高的遗传异质性。通常,常染色体显性遗传的耳聋表现为学语后聋,听力损失或轻或重,可单独累及某个频率段或各个频率段,病程发展多呈进行性;而呈常染色体隐性遗传的耳聋表现为学语前聋,听力损失一般为中、重度以上,全频听力损失多见。

这些耳聋基因编码的产物分属不同的蛋白家族:包括细胞外基质蛋白,肌球蛋白和其他骨骼肌蛋白,间隙连接蛋白和离子通道,以及转录因子和一些功能未知的基因表达产物。它们涉及听力器官的胚胎期发育、功能单位的形成、正常功能的维持等多个方面。从 NSHL 基因产物的功能方面, NSHL 基因分为五类。

1. 细胞外基质成分

柯蒂器的毛细胞上覆盖的盖膜在感觉传导机制中扮演了重要的角色。它产生剪切力使毛细胞上纤毛束弯曲从而打开离子通道,引发产生听力的一系列反应。盖膜由胶原纤维和非胶原纤维的基质组成,而后的主要成分是 α -和 β -盖膜蛋白(tectorin)。将盖膜及其他内耳外胶质成分和它们下面的非感觉细胞相联系的是耳锚蛋白(otoancorin)。

表 1-1 常染色体显性遗传及常染色体隐性遗传耳聋基因

序号	位点	基因	序号	位点	基因
1		CRYM ^[10]	1		PRES ^[34]
2	DFNA1	DIAPH1 ^[11]	2	DFNB1	GJB2 ^[14]
3	DFNA2	GJB3 ^[12]	3	DFNB1	GJB6 ^[35]
4	DFNA2	KCNQ4 ^[13]	4	DFNB2	MYO7A ^[36]
5	DFNA3	GJB2 ^[14]	5	DFNB3	MYO15 ^[37]
6	DFNA3	GJB6 ^[15]	6	DFNB4	SLC26A4 ^[38]
7	DFNA4	MYH14 ^[16]	7	DFNB6	TMIE ^[39]
8	DFNA5	DFNA5 ^[17]	8	DFNB7/DFNB11	TMC1 ^[31]
9	DFNA6/DFNA14	WFS1 ^[18, 19]	9	DFNB8/DFNB10	TMPS3 ^[40]
10	DFNA8/DFNA12	TECTA ^[20]	10	DFNB9	OTOF ^[41]
11	DFNA9	COCH ^[21]	11	DFNB12	CDH23 ^[42]
12	DFNA10	EYA4 ^[22]	12	DFNB16	STRC ^[43]
13	DFNA11	MYO7A ^[23]	13	DFNB18	USH1C ^[44, 45]
14	DFNA13	COL11A2 ^[24]	14	DFNB21	TECTA ^[46]
15	DFNA15	POU4F3 ^[25]	15	DFNB22	OTOA ^[47]
16	DFNA17	MYH9 ^[26]	16	DFNB23	PCDH15 ^[48]
17	DFNA20/26	ACTG1 ^[27]	17	DFNB24	RDX ^[49]
18	DFNA22	MYO6 ^[28, 29]	18	DFNB28	TRIOBP ^[50, 51]
19	DFNA28	TFCP2L3 ^[30]	19	DFNB29	CLDN14 ^[52]
20	DFNA36	TMC1 ^[31]	20	DFNB30	MYO3A ^[53]
21	DFNA44	CCDC50 ^[32]	21	DFNB31	WHRN ^[54]
22	DFNA48	MYO1A ^[33]	22	DFNB36	ESPN ^[55]
			23	DFNB37	MYO6 ^[56]
			24	DFNB59	PJVK ^[57]
			25	DFNB67	LHFPL5 ^[58]

编码胞外基质的基因有：编码 α -盖膜蛋白(α - tectorin) 的 TECTA(与 DFNA8/12 相关)；编码胶原 XI 多肽亚单位 2 的 COL11A2(突变导致 DFNA13、Stickler 综合征 2 型和 OSMED 综合征)；编码耳锚蛋白的 OTOA(与 DFNB22 相关)；COCH(Coagulation factor C Homologue)(与 DFNA9 相关)及 DFNA5(突变导致 dfna5)。

2. 细胞骨架蛋白

耳蜗和前庭器中维持高度的组织性和特殊的结构，包括富含肌动蛋白的感觉毛细胞的静纤毛，这对于听觉和平衡处理来说极其重要。很多细胞骨架蛋白以及相关蛋白分子，在毛细胞发育的不同阶段、毛细胞的不同部位，参与纤毛及其他细胞骨架相关成分的发育成熟、正常形态的维持和生理功能的行使。与 NSHL

有关的蛋白分子包括：常规肌球蛋白(MYH9, MYH14)和非常规肌球蛋白(MYO1A, MYO3A, MYO6, MYO7A 和 MYO15),肌动蛋白(ACTG1), 肌动蛋白成束化蛋白(ESPN), 含 PDZ 结构域蛋白(USH1C, WHRN), 钙粘蛋白(CDH23), 原钙粘蛋白(PCDH15), 可能的肌动纤维多聚化蛋白(HDIA1), 一种新的纤毛相关蛋白(STRC), 突触囊泡转运蛋白(OTOF)。

3. 离子内环境稳态中涉及的分子

维持耳蜗内离子, 尤其是钾离子循环中维持离子内环境稳态, 对于听觉过程中涉及的信号转导非常重要。K⁺通过两条途径转运至内淋巴液: 侧路途径包括柯蒂器的支持细胞, 螺旋韧带和血管纹的纤维细胞; 中间途径是朝齿间细胞方向穿过支持细胞。在 NSHL 基因中, 对钾离子循环发挥重要作用的是钾离子通道(KCNQ4)以及缝隙连接蛋白基因(GJB2, GJB3, GJB6 和 GJA1)。调节其他离子稳定和平衡的还包括 SLC26A4、CLDN14、TMPRSS3、WFS1 等。

4. 转录因子

已经有四个编码转录因子作为 NSHL 基因被发现: POU4F3, POU3F4, EYA4 和 TFCP2L3。POU4F3 和 POU3F4 是属于转录因子中包含 POU 域的家族。它们有的在较迟发育阶段发挥作用, 控制存活和特殊神经元亚型的最终分化。POU3F4 突变与 X 连锁混合性听力下降(DFN3)有关。POU4F3 与 DFNA 15 相关。

5. 未知功能的基因

目前还有一些 NSHL 基因的功能尚不清楚。这些基因有: DFNA5、WFS1(编码 wolframin 蛋白)、TMIE(跨膜内耳表达基因, TransMembrane Inner Ear expressed gene)和 TMC1(跨膜耳蜗表达基因, TransMembrane Cochlea-expressed gene 1)。例如 DFNA5, 尽管对 DFNA5 作了仔细的分析, 还是没有找到任何一个与其同源的基因, 也没有找到任何已知的功能结构域。因而, 这个基因在没有更新的数据来提供其功能方面的信息之前, 只能以其位点命名为 DFNA5。这几种未知功能的基因已经有小鼠模型, 这些小鼠模型对基因功能研究提供了线索。

怀化耳聋家系的定位

2002 年, 本研究室收集到一个湖南怀化常染色体显性遗传的双侧进行性感音神经性耳聋家系, 共 39 人, 其中病人 15 个, 包括 6 名男性和 9 名女性(见图 1-2)。先证者 V-9, 女, 23 岁, 16 岁起双侧耳鸣并进行性听力减退, 不伴其它器官系统的异常。家系成员中耳聋患者的起病年龄在 16-30 岁, 均表现为双侧耳鸣,

听力减退为双耳对称性、以高频听力下降为主的感觉神经性耳聋, 逐渐发展到累及所有频率的重度听力减退。听力曲线表现为 1000Hz 以上频率气道与骨导同时下降, 下降程度不等, 与患者年龄和起病时间的长短有关。

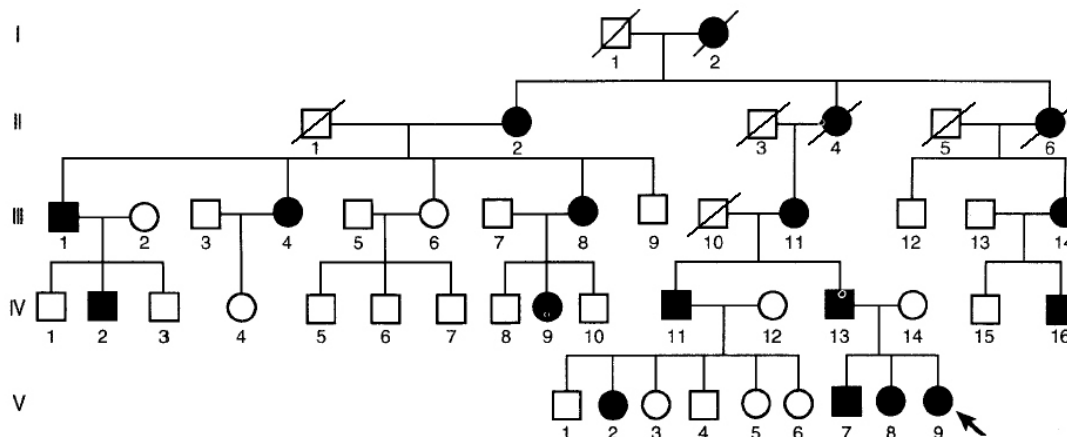


图 1-2 怀化籍耳聋家系家系图

对 CX31、CX26、POU4F3 基因的编码区及剪接位点的序列和 TECTA、DFNA5、MYO7A、DIAPH1 基因已发现突变的位点进行突变检测, 结果没有检测到已报道的突变。

对该耳聋家系进行全基因组扫描, 5 号染色体长臂的几个相邻的遗传标记 d5s471、d5s2115、d5s436 的 LOD 值分别为 1.17 ($\theta=0.1$)、3.08 ($\theta=0.00$)、3.62 ($\theta=0.00$)。为了精细定位该疾病基因, 及确定其两侧的位点, 在 d5s2115 和 d5s436 周围的区间内选取了更为密集的微卫星标记进行分析。连锁分析的结果将该耳聋家系的致病基因定位于 5 号染色体长臂, d5s2017 在 $\theta=0.00$ 取得最大的 LOD 值 6.89。

根据 Genethon 图谱上的标记顺序用 Cyrillic 构建了这个家系的单体型, 并用人工进行检查和调整, 通过分析受累个体中观察到的交换事件, 将该疾病基因定位在 5 号染色体长臂 D5S2056 和 D5S638 之间约 8.8cM 的区间。根据个体 II-2、III-12、III-14、IV-11 和 IV-13 的单体型推出 II-4 发生了重组, 将疾病共分离区的着丝粒端定位于 D5S2056 和 D5S2115 之间。她的后代 III-11、IV-11 和 IV-13 都遗传了这条重组染色体。另一个重要的重组事件发生在 IV-8。他是一个 42 岁的男性, 无任何听力下降的自觉症状, 纯音测听显示正常, 其在 D5S643 为纯合子, 很难确定是在 D5S2017 和 D5S643 之间发生了交换还是在 D5S643 和 D5S638 之间发生了交换。所以我们将共分离区的端粒端定位于 D5S643 和 D5S638 之间。这些结果表明该疾病基因定位于 D5S2056 和 D5S638 之间约 8.8cM 的区间, 在 GDB (Genome database) 中查询, 该区间相当于 5q23.3-q31.3 的细胞遗传学位点。

两个耳聋基因 DFNA1 (DIAPH1) 和 DFNA15 (POU4F3) 已被定位于 5q31,

但这两个基因经突变检测未发现突变。提示这是一个新的常染色体显性遗传进行性感音神经性耳聋位点，国际基因命名委员会(HUGO nomenclature committee)将其命名为 DFNA52。

与常染色体显性 NSHL 相关的基因有：间隙连接蛋白 (GJB2、GJB3、GJB6)；离子通道 (KCNQ4)；细胞外基质 (COCH、COL11A2、TECTA)；转录因子 (POU4F3)；肌球蛋白 (MYO7A、MYH9)；参与胞质分裂和肌动蛋白聚合 (DIAPH1)；DFNA5 (OMIM: 600994)。这些不同基因家族的存在说明有许多基因参与了听觉形成过程。怀化耳聋家系定位的区间比较大，因耳聋的遗传异质性很强，我们很难收集别的耳聋家系来缩小定位区域。定位区间内总计有 173 个 RefSeq Genes, 199 个 Know Genes，其中有许多基因是已克隆的耳聋基因的基因家族成员之一，很难根据一个基因的已知功能或推测的功能来选择候选基因。要克隆 DFNA52 的致病基因是困难的，候选基因的突变检测是又费时又费钱的，显然，需要运用高通量的生物信息学的分析方法来鉴定最可能的候选基因来做进一步的研究。

耳聋候选基因的生物信息学分析策略及耳聋基因数据库

当前，人类基因组研究的重心正在由“结构”向“功能”转移，一个以基因组功能研究为主要内容的所谓“后基因组时代”(post-genomics)，也即功能基因组(functional genomics)时代，已经到来。如何获取基因的功能信息，即与人类重大疾病和重要生理功能相关的基因信息，就摆在了我们面前。人类基因组计划已确定了一系列分布于全基因组 24 条染色体上的分子标记。根据这些分子标记的序列和定位信息，结合杂交或 PCR 的方法可以快速检测染色体上与遗传性疾病相关的热点位置，进而从中克隆疾病相关基因。定位候选克隆克服了经典的定位克隆纯粹依靠连锁分析进行染色体定位的繁冗而缓慢的弊端，大大加快了遗传性疾病易感基因的克隆工作的进程。总之，定位候选克隆策略(positional candidate cloning)是传统定位克隆(positional cloning)的改进和发展，随着人类基因组计划的完成，定位候选克隆在克隆疾病相关基因中的有效性而日益受到重视，已成为一种有效的克隆疾病相关基因的方法。

目前我们定位耳聋基因时常用的技术路线是：a.获取高信息量的非综合征遗传性耳聋家系；b.确定耳聋家系的遗传方式(显性、隐性、X-连锁或线粒体母系遗传)；c.根据确定的遗传方式选择候选的耳聋基因的 DNA 多态标记(在一种遗传方式中可以有多种表型，即发病时间、听力损失程度、累及频段、病程发展、伴随症状的不同形式，应根据所拥有的耳聋家系的表型进行候选 DNA 多态标记的筛选)；d.应该连锁分析软件确定或排除候选的耳聋基因位点；e.如与候选的耳聋

基因位点均无连锁关系,则应用基因组每隔 20cM 或 10cM,或 5cM 的图距单位选择一个 DNA 标记,对家系成员作全基因组 DNA 标记座位的基因型分析,所有家系成员的基因型和表型刻录输给计算机作连锁分析或区间定位分析,从而判断这个家系的表型性状的基因在基因组上的位置。通过单体型分析受累个体中观察到的交换事件,将该疾病基因定位于某一特异的染色体亚区,然后采用一些方法识别候选基因(candidate genes),然后逐个检查是否具有疾病基因的确切证据,以克隆其疾病基因。

用来识别疾病基因的方法最后得到的是候选基因(candidate genes),这些基因必须逐个进行检验,以确定与疾病的相关性。但是有二种疾病基因识别的方法是直接从假设的候选基因开始的,即先假设某一基因为疾病位点,然后再检查该基因与疾病的相关性。这种候选基因的入围可能依据其产物的特征与疾病发病过程一致。显然,在未知某种疾病的基因定位也可以提出其候选基因,这称为不依赖定位的候选基因法。如果候选基因与疾病基因染色体的定位一致,则对该候选基因的信心就更更大些。这种定位性候选基因法近年来大获成功。随着疾病基因和人类基因愈来愈多地在染色体上定位,该方法已渐渐成为疾病基因定位的主要方法。

鉴定人类致病基因的关键在于获得对疾病机制的理解,其本质需要发展新的诊断学和治疗学的方法。连锁分析和突变检测已较早地并成功运用于染色体定位和鉴定致病基因。符合孟德尔遗传的单基因病的疾病基因的克隆主要依赖于有好的家系和样本。通过连锁分析和遗传定位是人类遗传学研究的主要依据,虽然定位信息减少了致病候选基因的数目,但这种减少经常不足以迅速鉴定疾病基因,有时在寻找到致病性突变前需要分析大量的基因。分离这些位点的致病基因有时相当困难的:定位区间过大,包含数百个可能的候选基因,使得实验方法又费时又费钱。显然,要采用的高通量分析方法就需要运用计算机的方法来鉴定最可能的候选基因来做进一步的研究。

人类耳聋候选疾病基因的选择可从以下几个方面考虑。

(1).依据适当的表达方式和功能:

根据有些疾病的发病机理可直接提出具有合适表达方式或功能的疾病候选基因。例如听觉缺陷可能涉及发育胚胎的耳蜗形成之前或期间表达的基因。如果某一基因在此发育阶段于耳蜗表达则为一个候选基因。从内耳特异组织分离出可能参与听力形成的基因,在耳聋的家系中进行突变检测,从而克隆耳聋基因。从人类或老鼠的耳蜗中分离到的特异表达的基因都可以作为耳聋的候选基因。运用测序的方法分析人类胚胎耳蜗 cDNA 文库,生成的大量的 ESTs(表达序列标记, expressed sequence tags)提供人类耳蜗转录图谱,是定位候选克隆的候选基因

的资源。

(2).疾病模型动物表型同源基因:

如果某一动物疾病表型与某一人类疾病很相似,则人类疾病有可能是由于相关的同源基因突变所致。如果动物的疾病基因已知,就可以用以帮助人类疾病基因的识别和鉴定,也可以用于检验该基因是否与人类疾病相关。例如,小鼠内耳与人类内耳在结构和功能上相似,小鼠是研究人类耳聋基因的首选模型动物。小鼠全基因组的突变率很低,常规的基因敲除(knock-out)和转基因(transgenesis)突变规模小,且研究的是已知基因的功能。近年,采用化学诱导突变的表型筛选(phenotype screens)已成为日益关注的目标,乙烷亚硝基脲(N-ethyl-N-nitrosourea, ENU)诱变的大规模表型筛选能获得大量新突变,确定了一些新基因,同时也发现了与听觉及平衡功能有关的突变和基因。估计,小鼠存在近 200 个与听力损害有关的突变基因。小鼠突变对于展现哺乳动物基因功能和模拟人类疾病具有关键性的作用,对分析和探讨疾病的发病机制具有特别重要意义。与小鼠耳聋相关的基因有一些基因在人体的同源基因已证实是耳聋基因,其余基因仍是克隆和鉴别新的耳聋基因的良好候选基因。将一个突变基因在小鼠图谱中定位后,可以观察定位在附近的表型异常是否有一种表型与定位基因引起的缺陷一致。如果发现合理的联系,就可以对有问题的基因进行分子鉴定以识别突变。对有听觉缺陷的小鼠进行近交建立听觉缺陷小鼠模型,再对这些小鼠进行缺陷基因的鉴别和克隆,克隆出来的小鼠耳聋基因的人类同源基因就是人类耳聋基因的良好候选基因。耳聋基因 MY07A 就是通过此方法克隆的。

(3).与某一导致相似人类疾病的基因同源或功能相关:

有时不同的基因位点突变可能产生同样的临床结果,因此不同基因引起的表型有时可以区别,有时则不能。如果已识别某一基因为这种疾病基因,那么序列或功能上紧密相关的其他基因可能是相似疾病的候选基因。例如所识别基因可能是多基因家族的成员之一,其他成员也许是相似疾病的候选者。若某一基因的功能与某已定位的具有相似表型疾病基因密切相关也可以作为候选基因。基因功能上的相关体现在它们产物功能上的相关,如编码受体和其相应配体,或编码同一代谢途径中或发育过程中的相互作用的蛋白。

如今,随着新的分子生物学技术的发展,人类基因组计划的完成,特异性耳蜗特异性互补 DNA(complementary DNA, cDNA)文库数量的增加以及一些动物模型的建立,以及可获得的生物信息数据库的激增,为克隆新的遗传性耳聋基因提供了最可能的候选基因。

1. Morton 耳蜗 EST 数据库^[56] (Morton Cochlear EST Database, http://www.brigham-womens.org/bwh_hearing/)

附属 Bringham 妇女医院和哈佛大学的 Morton 实验室致力于各种哺乳动物

内耳的分子生物学的研究,研究重点是鉴定耳聋基因,了解在内耳中的各种蛋白质的功能和研究正常或听觉缺陷的耳中的基因表达情况。耳蜗 EST 数据库包括两个部分: Human Cochlear ESTs 和 Inner Ear cDNA Microarrays。

人类耳蜗 ESTs(Human Cochlear ESTs): Brigham 妇女医院病理科收集的人类耳蜗用于生成 RNA 和 cDNA 文库。这个文库捐赠给 IMAGE 协会,建立表达序列标志 ESTs(dbEST Library ID.371)。^①人类胚胎耳蜗 cDNA 文库 EST 数据(Morton Human Fetal Cochlea cDNA Library EST Data): 共列出了 15000 多个 ESTs。这个数据库提供了 ESTs 的 GenBank 登录号和鉴定出的相应的基因(若是已知的基因)。^②人类耳蜗 EST 消减文库(Subtracted Human Cochlear Library EST Data): 有少部分的 cDNA 文库来自于人类胚胎 mRNA(dbEST Library ID 1822)。根据选择性扩增差异表达基因(Clontech PCR-Select cDNA Subtraction)的方法将人类胚胎的耳蜗与脑, 睾丸, 肾, 子宫和纤维原细胞的 cDNA 标准化, 建立此文库, 共列出 3880 条 ESTs。

内耳 cDNA 微阵列(Inner Ear cDNA Microarrays): Morton 实验室建立了两个不同的内耳相关的 cDNA 微阵列。^①Human Cochlea: 人类耳蜗 cDNA 微阵列来源于 Morton 胚胎耳蜗 cDNA 文库的克隆。芯片包含 5000 多个 cDNA 克隆, 与阳性和阴性对照一样, 附加的耳聋基因也加入此芯片中。^②Mouse Organ of Corti: 小鼠 Corti 器微阵列来源于 NIH 的 Kachar 实验室的小鼠 Corti 器 cDNA 文库。此 cDNA 芯片包含 9000 多个 cDNA 克隆, 与阳性和阴性对照一样, 附加的耳聋基因也加入此芯片中。

2. 耳发育相关基因表达数据库^[55] (Table of gene expression in the developing ear, <http://www.ihr.mrc.ac.uk/hereditary/genetable/>)

此数据表由英国 MRC 听力研究所和欧委会建立及支助的。此数据库共包含 766 条数据, 最近更新日期为 2002 年 4 月。数据库收集了大量的在不同的动物种类的内耳发育过程中在不同时期有表达的基因。数据库表格的第一部分是所有染色体位置已知的小鼠(资料来自于 MGD 数据库)和人类(资料来自于 OMIM 数据库)基因的列表。这些基因已根据它们在小鼠染色体上的位置进行了升序排序, 即从 1 号染色体的着丝粒附近开始起始位置。例如, 当唯一的人类基因的位置已知, 将此基因插入数据表中与小鼠对应的保守同线性(conserved synteny)的位置上(见 <http://www.ncbi.nlm.nih.gov/Homology>)。数据表的第二部分列出染色体位置未知的基因和蛋白质, 其按照字母顺序排序。

检测基因或蛋白质表达的方法有利用特殊 RNA 探针的原位杂交(in-situ hybridisation)、或用抗体检测蛋白质表达的免疫定位(immunolocalisation)。其它

方法还包括 Northern blotting, 逆转录酶 PCR(reverse transcriptase polymerase chain reaction, RTPCR), 从耳组织中提取蛋白质及 mRNA 来产生 cDNA 和报道基因。

3. Ear Libraries of NEIBank^[57] (<http://neibank.nei.nih.gov/cgi-bin/libList.cgi?-tissue=ear>)

由美国国立卫生研究院的眼科研究所(National Institutes of Health, National Eye institute)建立的小鼠 Corti 器(OC)的表达序列标志(EST)数据库, 有 20000 个独立的克隆经过分离, 分析和聚类成 8690 个独立的基因族。大量的 Corti 器相关的新基因被鉴定出来。在序列比对时经常发现已知基因的选择性剪接本与 Corti 器功能相关。通过生物信息学分析将小鼠 Corti 器的 ESTs 找到对应的同线性的与人类常染色体显性和隐性耳聋相关的人类染色体位置, 这为鉴定人类疾病提供了有用的新的定位候选基因。通过此资源既可查看在听觉过程上在 Corti 器中基因的表达谱情况, 又是识别新的在听觉过程中起重要作用基因的工具。

数据库包括四部分: ①NbLib0011: dbEST human cochlea unnormalized; ②NbLib0053: NIDCD mouse organ of corti unnormalized (gi); ③NbLib0123: NIDCD mouse taste cell subtracted (ia); ④NbLib0010: NIDCD human fetal cochlea subtracted (n)。

4. 遗传性听力损害相关的小鼠品系和突变(Hereditary Hearing Impairment in Mice, <http://www.jax.org/hmr/>)

由国立耳聋和其它交际障碍研究所(National Institute on Deafness and other Communication Disorders, NIDCD)支助, Jackson 实验室通过大规模地表面型筛选鉴定出与遗传性听力损害相关的小鼠品系和突变。根据国际上最新的研究文献, 更新一些数据, 如人类听力障碍的小鼠模型、与听觉及平衡功能有关的小鼠基因和突变、耳聋位点在染色体上的图位、与采用小鼠模型的听觉研究的网络资源的链接。其研究重点主要是将小鼠作为人类遗传性听力障碍的动物模型。数据包括的内容有:

①人类听力障碍的小鼠模型(Mouse Models of Human Hearing Disorders): 列出人类非综合征及综合征耳聋基因和对应的小鼠基因和突变。

②近亲交配小鼠品系中老年性聋潜在的遗传因素(Genetic factors that underlie age-related hearing loss in inbred mouse strains): 数据表列出在近亲交配的小鼠品系中会导致老年性聋(age-related hearing loss, AHL)的遗传因素, 其遗传位置图谱, 基因定位的所采用的方法, 已知的易感和抗性位基因的品系, 和相关参考文献。

③耳聋的遗传修饰因子和二基因遗传 (Genetic modifiers and digenic inheritance of hearing loss): 数据表列出已报道的在人和鼠的遗传性疾病中可改变耳聋表达的遗传修饰因子位点和二基因遗传的案例。

④已报道导致内耳畸形或异常的小鼠突变(Mouse mutations reported to cause inner ear malformations or dysfunction): 数据表包括 200 多个单个基因位点。各个条目提供了以下信息: 基因符号, 基因名称, 染色体图位置, 突变符号, 突变名称, 突变类型, 耳相关的病状, 听觉测定, 前庭功能, 和相关参考文献。通过基因符号将条目按字母顺序排序。

5. 小鼠耳聋突变(Deaf Mouse Mutants, <http://www.sanger.ac.uk/PostGenomics/-mousemutants/deaf/>)

Sanger 研究所列出的小鼠耳聋突变数据, 数据表包括已知涉及耳聋和/或平衡功能障碍的主基因或位点。内容包括 9 个部分: 中耳或耳廓缺陷的小鼠突变; 内耳畸形的小鼠突变; 伴神经上皮缺陷(neuroepithelial defects)的小鼠突变; 伴内淋巴内环境稳定缺陷的小鼠突变; 伴盖膜或耳石膜(otolithic membrane)缺陷的小鼠突变; 伴外周神经缺陷的小鼠突变; 伴中枢听觉系统缺陷的小鼠突变; 伴听阈增高但病理未知的小鼠突变; 可作为耳聋的候选基因但无听觉系统相关信息的小鼠突变。

现在还发展出了一些与其它方法相结合的新思路, 如与消减杂交相结合, 筛选位于候选区域的差异表达基因, 大大提高了获得有价值基因的可能性。同时人类基因组计划中转录图谱工作的开展, 有很多 EST 已被精确定位于染色体的不同区带上, 可以直接进行功能研究。随着研究的深入, 越来越多的 EST 定位工作的完成, 基因转录图谱的不断完善, 可以直接进入功能鉴定和基因全长的克隆工作, 这可能将是定位克隆中最快的方法。在分析分离手段不断提高, 人类基因组计划紧锣密鼓地实施的今天, 各种新思路新方法层出不穷, 定位克隆方法的周期也将不断缩小, 为人类最终克隆各种疾病基因提供了可能性。

耳聋相关网络资源:

1. Morton Cochlear EST Database, <http://www.brighamandwomens.org/bwh-hearing/>)
2. Mouse OC library at UniGene:
<http://www.ncbi.nlm.nih.gov/UniGene/lib.cgi?ORG=Mm&LID=692>
3. Washington University Inner Ear Protein database:
<http://oto.wustl.edu/thc/innerear2d.htm>
4. Zebrafish Inner Ear EST database: <http://www.genoscope.cns.fr/zie/>
5. The Jackson Laboratory Hearing Web page:
<http://www.jax.org/research/hhim/documents/models.html>.
6. MRC Institute of Hearing Research:
<http://www.ihr.mrc.ac.uk/hereditary/genetable/index.shtml>
7. Inner Ear Gene Expression:
<http://www.mgh.harvard.edu/depts/coreylab/genomics.html>.
8. Hereditary Hearing Loss Web page: <http://www.uia.ac.be/dnalab/hhh>
9. National Institute on Deafness and other Communication Disorders
<http://www.nidcd.nih.gov/>
10. Association for Research in Otolaryngology <http://www.aro.org/>
11. Hereditary Hearing Loss Homepage <http://webhost.ua.ac.be/hhh/>
12. Ear resources on NEIBank cDNA libraries from cochlea
<http://neibank.nei.nih.gov/Ear/NbEar.shtml>
13. Deaf Mouse Mutants <http://www.sanger.ac.uk/PostGenomics/mousemutants/deaf/>

第一章 生物信息学方法分析 DFNA52 的候选基因

应用生物信息学方法分析 DFNA52 在 D5S2056 和 D5S638 之间候选基因。综合耳蜗基因表达谱, 基因功能, 耳聋临床表现特征, 小鼠耳聋致病基因以及识别疾病候选基因的生物信息学软件(如 G2D, DGP, PROSPECTR, GeneSeeker) 的预测结果等资料, 建立了识别耳聋候选基因的策略, 优先选择了 52 个候选基因(包括小鼠耳聋基因 *Neurog1*, *SMAD5*)。

1.1.材料与方法

1.1.1 研究对象:

一个遗传了 5 代的常染色体显性遗传的 NSHL 家系来自湖南怀化, 共 39 人, 其中病人 15 个, 包括 6 名男性和 9 名女性(见图 1-1)。这个家系所有的研究对象都进行病史询问、体格检查(包括耳的外形、外耳道、鼓膜的检查)、纯音测听(PTA)(包括 125HZ、250HZ、500HZ、1000HZ、2000HZ、4000HZ 和 8000HZ 的气导和骨导听力)和畸变产物耳声发射(中心频率分别为 250HZ、500HZ、1000HZ、2000HZ、4000HZ 和 8000HZ)检测。先证者 V-9, 女, 23 岁, 16 岁起双侧耳鸣并进行性听力减退, 不伴其它器官系统的异常。家系成员中耳聋患者的起病年龄在 16-30 岁, 均表现为双侧耳鸣, 听力减退为双耳对称性、以高频听力下降为主的感觉神经性耳聋, 逐渐发展到累及所有频率的重度听力减退。听力曲线表现为 1000Hz 以上频率气道和骨导同时下降, 下降程度不等, 与患者年龄和起病时间的长短有关。畸变产物耳声发射的听阈图类似于纯音听力图。曾患中耳炎、曾用过氨基糖甙类药物、有长期的噪音接触史的耳聋病人不纳入本研究。

1.1.2. 研究方法:

候选基因突变分析基本实验步骤, 如图1-3所示。

1.2 生物信息学分析定位区间候选基因

在 UCSC Genome Browser (<http://genome.ucsc.edu/cgi-bin/hgGateway>) 查询 D5S2056 到 D5S638 区间的基因。结果显示在这个区间总计有 173 个 RefSeq Genes, 199 个 Know Genes, 如图 1-4 所示。其细胞遗传学距离为 5q31.1-5q32, 物理距离为 chr5:134,325,465-146,314,598, 大小 11,989,134bp。

根据 NCBI Map View 36.1 (<http://ncbi.nih.gov>)分析 DFNA52 定位区间 D5S2056(139.33cM)至 D5S638 (148.63cM)的候选基因。区间内有 152 个已命名的基因和 26 个无特征性的基因。

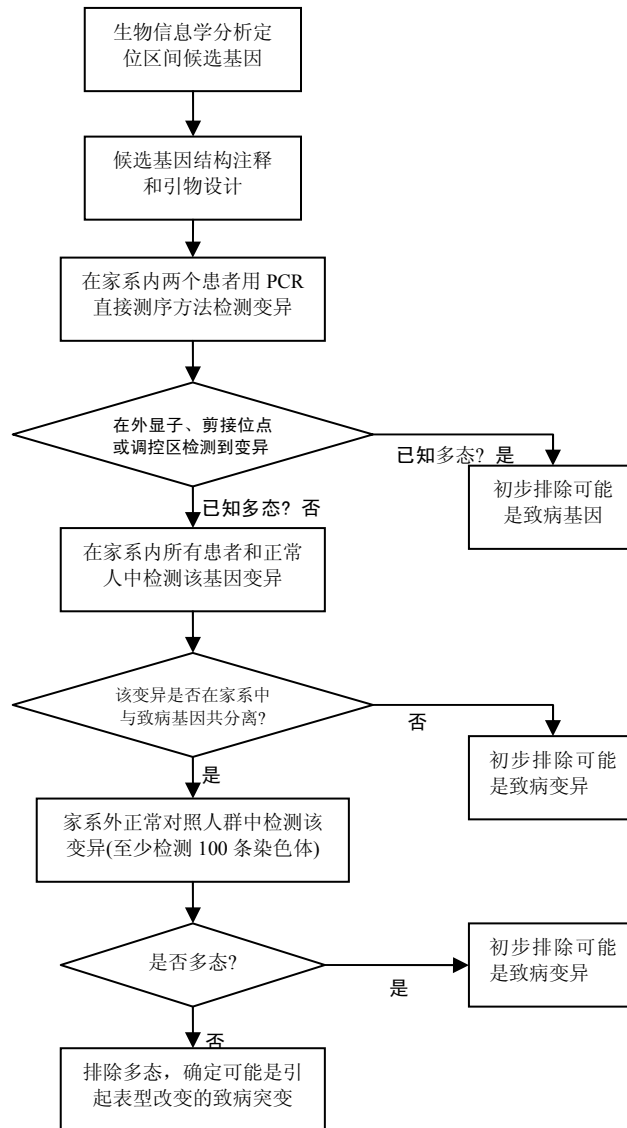


图 1-3 候选基因突变分析流程图

因这些导致常染色体显性遗传的非综合征型耳聋的基因是不同的，许多基因参与了听觉形成过程，包括间隙连接家族、细胞骨架蛋白、转录因子、离子通道、肌球蛋白和一些功能未知的基因，很难根据一个基因的已知功能或推测的功能来选择候选基因。本研究根据耳蜗基因表达谱、基因功能、疾病的表型、小鼠模型的遗传学研究(表型，表达)和发表的文献的研究情况以及利用生物信息学候选疾病基因鉴定方法来选择候选基因。



根据 UniGene(<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=unigene>)中的基因表达谱情况, 三个耳基因表达数据库 gene expression in the developing ear

(<http://www.ihr.mrc.ac.uk/hereditary/genetable/>)^[55]、Human Cochlear cDNA library and EST database(<http://hearing.bwh.harvard.edu/estinfo.htm>)^[56]和 ear libraries on NEIBank^[57] (<http://neibank.nei.nih.gov/cgi-bin/libList.cgi?tissue=ear>), 选取了在数据库中位于 D5S2056 和 D5S638 区间的基因作为候选基因。

1.2.2. 根据 mouse 耳蜗 cDNA 文库及 EST 的序列找候选基因

通过分析 NCBI 的 Mouse Organ of Corti cDNA pBluescript、NCBI 中的耳聋和前庭器组织 cDNA 减除杂交序列, 通过 BLAT 对 mouse genome 做序列比对, mouse 对应的人类 DFNA52 定位区间的基因, 选择其中的基因作为 DFNA52 的候选基因。

1.2.3. 人类/小鼠同源性比较和小鼠耳聋模型

不同种属的同源性分析和比较基因组学也能为鉴定相关候选基因提供重要的工具。人类染色体 5q31 (DFNA52) 同源对应小鼠 13,18 染色体的区域(见表 1-2)。小鼠和其它物种同源性也能有助于预测和研究基因产物的功能和性质。有近 200 多种小鼠基因突变与内耳异常有关(Hereditary Hearing Impairment in Mice, <http://www.jax.org/hmr/index.html>), 有些基因的人类同源基因已证实是耳聋基因, 其余基因仍是克隆和鉴别新的耳聋基因的人类耳聋基因的良好候选基因。通过分析 Mouse Genome Informatics(<http://www.informatics.jax.org/>)中基因的功能和表型情况, 及 Deaf Mouse Mutants(<http://www.sanger.ac.uk/PostGenomics/mousemutants/deaf/>)情况, DFNA52 区间内的 Neurog1(neurogenin 1)是小鼠耳聋相关基因, Neurog1 是内耳感觉神经元发育所必需的, neurog1 的缺失突变体(null mutants)表现为异常的内耳表型, 突变体的耳发育的感觉上皮的毛细胞较小, 其形态正常但紊乱和数目减少^[58]。在 Pitx1 (paired-like homeodomain transcription factor)突变性研究中表现为中耳听骨形态改变异常^[59]。

表 1-2 小鼠 13,18 染色体对应人染色体 5q31 的区域内的耳聋基因

Mouse chr	cM	gene
11	13	Otx1, orthodenticle homolog 1 [targeted mutation; Jackson waltzer, jv]
11	17	Foxi1, forkhead box I1 [targeted mutation]
11	25	Prop1, paired like homeodomain factor 1 [Ames dwarf, df]
11	33.9	Myo15, myosin XV [shaker 2, sh2] NS
11	34.45	Pmp22, peripheral myelin protein, 22 kDa [trembler, Tr] S
11	46	co cocked
11	56	Cola1, procollagen, type I, alpha 1 [Mov13 transgene disruption; targeted mutation] S
11	73.5	Ts, tail-short

11	77	Sans, scaffold protein containing ankyrin repeats and SAM domain [spontaneous mutation js, Jackson shaker] S
11	79	Actg, actin, gamma, cytoplasmic [DFNA20/26] *
13	21	mu, muted
13	25	Tcfap2a, transcription factor AP-2, alpha [Doard, ENU-induced mutation]
13	35	Neurod3, neurogenic differentiation 3 [targeted mutation]
13	36	Ntrk2, neurotrophic tyrosine kinase, receptor, type 2 [targeted mutation]
13	37	Agtbp1, ATP/GTP binding protein 1 [Purkinje cell degeneration, pcd]
13	38	Ptch, patched homolog [mesenchymal dysplasia, mes]
13	39	Slc12a7, solute carrier family 12, member 7; K-Cl co-transporter KCC4 [targeted mutation]
13	40	Mass1, monogenic, audiogenic seizure susceptibility 1 [Frings, BUB/BnJ strains, targeted mutation] S
18	3	Tw, twirler
18	6	Aqp4, aquaporin 4 [targeted mutation] ALDH7A1
18	16	Diap1, diaphanous homolog 1 [DFNA1] *
18	24	Pou4f3, POU domain, class 4, transcription factor 3 [targeted mutation; dreidel, ddl] NS
18	27	Slc12a2, solute carrier family 12, member 2 (Na-K-Cl co-transporter) [targeted and spontaneous mutations]
18	27	sy, shaker-with-syndactylism deletion [contains Slc12a2]
18	55	Mbp, myelin basic protein [shiverer, shi]

1.2.4. Gene Expression Omnibus 数据库关于 Cochlear hair cell line differentiation time course 的两个 GDS51 和 GDS45。分析其中的基因。

GDS51 在 DFNA52 定位区间对应的 mouse 的 13,18 号染色体的基因有: Neurog1、Il9、Pabpc2、Npy6r、Pitx1、Hbegf、Hars、Wnt8a、Diap1、Nr3c1、Ppp2r2b、Spock1

表 1-3 耳蜗毛细胞细胞株分化时程 (Mu11K-B)

登录号:	GDS51
标题:	Cochlear hair cell line differentiation time course (Mu11K-B)
DataSet type:	gene expression array-based (cDNA / in situ oligonucleotide)
注释:	10 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:6530415D11 product:unknown EST, full insert sequence
概要:	Gene expression profile of differentiation of inner ear hair cell precursors in vitro. A conditionally immortal cell line derived from the mouse cochlea, UB/OC-1, examined from 0 to 14 days.

GDS45 在 DFNA52 定位区间对应的 mouse 的 13,18 号染色体的基因有: Pura、Ube2d2、Hbegf、Kif2a、Pitx1、Tgfbi、Fgf1、Ctnna1、Hspa9a、Egr1、Diap1、Gfra3

表 1-4 耳蜗毛细胞细胞株分化时程 (Mu11K- A)

标题:	GDS45
DataSet type:	Cochlear hair cell line differentiation time course (Mu11K-A)
注释:	gene expression array-based (cDNA / in situ oligonucleotide)
概要:	2 days neonate thymus thymic cells cDNA, RIKEN full-length enriched library, clone:E430020F23 product:unknown EST, full insert sequence
标题:	Gene expression profile of differentiation of inner ear hair cell precursors in vitro. A conditionally immortal cell line derived from the mouse cochlea, UB/OC-1, examined from 0 to 14 days.

1.2.5.根据已克隆的耳聋基因及其表达情况和所致耳聋临床表现

DFNA52 特征性临床表现为语后聋, 进行性发展, 高频聋, 根据遗传性 NSHL 基因及其表达情况和所致耳聋临床表现, 见表 1-5, 有这种临床表现的已克隆的 DFNA 的非综合征耳聋基因主要有肌球蛋白(myosin)超家族, KCNQ4(突变导致 DFNA2)^[60], 它们在内耳表达部位多为耳蜗内外毛细胞, 螺旋内感觉神经元等。DFNA52 定位区间内无肌球蛋白家族基因, 但有 KCTD16 (potassium channel tetramerisation domain containing 16)。

表 1-5 常染色体显性遗传非综合征耳聋位点及耳聋基因的临床表现

位点	染色体位置	基因	基因产物类型	耳聋的临床特点
DFNA1	5q31	DIAPH1	Cytoskeletal protein	语后, 进行性, 低频聋
DFNA2	1p34	GJB3 / KCNQ4	Gap junction component /channel component	学语后聋,, 进行性, 高频聋, 10-30 岁发病
DFNA3	13q12	GJB2/ GJB6	Gap junction component	语前, 非进行性, 高频聋
DFNA4	19q13	MYH14	细胞骨架蛋白	语后, 进行性, 中或全频聋
DFNA5	7p15	DFNA5	(Novel protein)	语后, 进行性, 高频聋
DFNA6	4p16.3	WFS1	维持离子稳定	语后, 进行性, 低频聋
DFNA7	1q21-q23	unknown		语后, 进行性, 高频聋
DFNA8	11q22-24	TECTA	Extracellular	语前, 非进行性, 中或全频聋
DFNA9	14q12-q13	COCH	Extracellular	语后, 进行性, 高频聋, 伴前庭症状
DFNA10	6q22-q23	EYA4	A member of the vertebrate Eye family of transcriptinonal activators	语后, 进行性, 中或全频聋, 20-60 岁, 多为 30-40 岁发病
DFNA11	11q12.3-q21	MYO7A	Motor molecule	语后, 进行性, 中或全频聋, 10-20 岁发病, 多变性的前庭功能障碍。
DFNA12	11q22-q24	TECTA	细胞外基质	语前, 非进行性, 中或全频聋
DFNA13	6p21	COL11A2	Collagen protein	语前, 非进行性, 中或全频聋
DFNA14	4p16	WFS1	维持离子稳定	语后, 进行性, 低或中频聋
DFNA15	5q31	POU4F3	Transcription factor	语后, 进行性, 中或高频聋
DFNA16	2q24	unknown		语后, 波动性, 高频聋
DFNA17	22q	MYH9	细胞骨架蛋白	学语后, 进行性, 高频聋, 蜗管退化

DFNA18	3q22	unknown		语后, 进行性, 高频聋
DFNA19	10 (pericentr.)	unknown		语前, 非进行性, 高频聋
DFNA20	17q25	ACTG1	细胞骨架蛋白	学语后聋, 高频聋, 各个频段, 进行性发展
DFNA21	6p21	unknown		
DFNA22	6q13	MYO6	细胞骨架蛋白	语后, 进行性, 高频聋, 伴前庭症状
DFNA23	14q21-q22	unknown		语前, 非进行性, 低频轻至中度, 高频中至深度聋
DFNA24	4q	unknown		语前, 非进行性, 低或中频聋
DFNA25	12q21-24	unknown		语后, 进行性, 高频聋, 发病年龄偏晚
DFNA26	17q25	ACTG1	细胞骨架蛋白	语后, 进行性, 高频聋
DFNA27	4q12	unknown		语后, 进行性, 高频聋
DFNA28	8q22	TFCP2L3	转录因子	语后, 进行性, 中或高频聋, 7 岁开始聋, 15 岁左右达重至深度聋
DFNA29	reserved			
DFNA30	15q25-26	unknown		语后, 进行性, 高频聋
DFNA31	6p21.3	unknown		
DFNA32	11p15	unknown		语后, 进行性, 高频聋
DFNA33	reserved			
DFNA34	1q44			语后, 缓慢进行性, 高频聋
DFNA35	reserved			
DFNA36	9q13-q21	TMC1	功能不明	语后, 迅速进行性, 高频聋
DFNA37	1p21			语后, 进行性, 高频聋, 逐渐波及中低频, 早龄开始发病
DFNA38	见 DFNA6/14			语后, 进行性, 低频聋
DFNA39	4q21.3	DSPP		学语后, 进行性, 高频聋
DFNA40	16p12			
DFNA41	12q24-qter	unknown		语后, 进行性, 高频聋
DFNA42	4q28			
DFNA43	2p12			
DFNA44	3q28-29	CCDC50		20 岁后开始全频下降
DFNA45	reserved			
DFNA46	reserved			
DFNA47	9p21-22	unknown		10 岁后开始进行性耳聋, 中至重度
DFNA48	12q13-q14	MYO1A	细胞骨架蛋白	
DFNA49	1q21-q23	unknown		
DFNA50	7q32	unknown		
DFNA51	9q21			reserved
DFNA52	5q31	unknown		reserved
DFNA53	14q11-q12	unknown		语后聋, 高频, 进行性, 累及全频
DFNA54		unknown		语后聋, 低频

* reserved 表明没文献发表, HUGO Nomenclature Committee 保留的命名

1.2.6.根据已发表文献的研究情况

Abe S 等人^[61]通过分析人类内耳和前庭组织的基因芯片(包含 23,040 基因)表达谱情况鉴定 cDNA Microarray 数据,发现在耳蜗和前庭器组织中有 52 个基因的信号强度比 29 个其它组织混合物强 10 倍,这 52 个基因中 PCDHGC3 是在 DFNA52 定位区间内的,将其做为候选基因。

Alsaber R 等人^[62]运用生物信息学方法分析各个未克隆的非综合型耳聋位点的候选基因,列出的这些候选基因对于经过连锁分析定位的耳聋家系提供了突变分析的出发点。DFNA52 区间内的候选基因有: FGF1、GFRA3、IK、PCDH1、DIAPH1、POU4F3、TTID、NRG2、PCDHAC1、PCDHAC2、NDFIP1。

1.2.7 识别候选基因的生物信息学软件

随着基因组学和功能基因组学相关信息的飞速发展,识别疾病基因的方法迅速形成。如今数据库成为选择候选疾病基因的必要方法。疾病的特征定位信息和功能信息的结合是选择候选疾病基因常见的策略。在近几年,许多候选疾病基因的生物信息学方法已形成。这些区分优先次序的方法选择最有可能的候选基因进行分析,将加速疾病基因的发现过程。这些识别人类疾病候选基因的不同方法覆盖了不同的方向。它们运用功能的和文献资料,基因特殊性特征,基于解剖学的基因/蛋白质表达数据或表型比较分析。

(1) Genes2Diseases (G2D) (http://www.ogic.ca/projects/g2d_2/)^[63,64]

G2D 系统是基于生物医学文献检索和结合病理学条件的基因语义学术语,通过 GO 注释和疾病相关基因的同源性分析鉴定出候选基因。

在 G2D 界面上输入 phenotype 和 location 资料,其预测的候选基因为: TGFBI、PCDHB15、NPY6R(假基因)、TRPC7、ZMAT2、KIF20A、SLC4A9、SLC23A1、HNRPA0

(2) DGP^[65](<http://cgg.ebi.ac.uk/services/dgp/>)

DGP 是一个人类基因其涉及遗传性疾病的概率的数据库。使用这些序列特征的疾病基因预测系统可在线获得(<http://cgg.ebi.ac.uk/services/dgp/>)。在 search by gene or position 的网页中选择 5 号染色体, Limits of the region 选择从 134325465 至 146314598,其 DGP 生成结果情况如下表 1-6。DGP 预测在 DFNA52 区间内的基因可能涉及人类遗传性疾病的概率情况,其概率高依次为: NR3C1、PCDH12、POU4F3、PITX1、ANKHD1、HBEGF、TGFBI、PPP2R2B 等。

表 1-6 从 5 号染色体的碱基对 134325465 至 146314598 搜索基因

Gene Symbol	Probability	known phenotype
NR3C1	0.780718	Yes
PCDH12	0.77415	No
POU4F3	0.759831	Yes
PITX1	0.741876	No
EIF4EBP3(ANKHD1),FLJ20288	0.734616	No
DTR (HBEGF)	0.728437	Yes
TGFBI	0.72384	Yes
PPP2R2B	0.702065	Yes
LARS	0.697761	No
GRAF(ARHGAP26)	0.694918	Yes
IK	0.686479	No
SPRY4	0.684969	No
DIAPH1	0.683384	Yes
TTID	0.678684	Yes
FGF1	0.677605	No
TCERG1	0.661627	No
MATR3	0.660132	No
MADH5	0.652244	No
CDC25C	0.64936	No
ARAP3	0.63254	No
H2AFY	0.626571	No
HSPA9B	0.615334	No
PCDHA13,PCDHA3,PCDHAC1,PCDHAC2	0.606412	No
KIAA0141	0.60117	No
PCDH1	0.595187	No
WNT8A	0.588213	No
PCDHGA12,PCDHGC3,PCDHGC4,PCDHGC5	0.584626	No
MGC29463	0.575485	No
CTNNA1,CTNNA2,VR22	0.575397	No
KIF20A	0.570886	No

(3) PROSPECTR^[66]

在 PROSPECTR 网页服务(<http://www.genetics.med.ed.ac.uk/prospectr/>), 使用者可在任一条目或基因组区域中将涉及疾病的可能性的基因进行排序。这种方法用 GO 术语、内在功能域及基因表达数据扩展。

用 PROSPECTR 搜索 markers D5S2056 和 D5S638 之间的基因情况如表 1-7。

表 1-7 在 D5S2056 至 D5S638 之间所有基因的得分(列出前 21 个基因)

tRank	Symbol & description	Score
1	POU4F3	0.72
2	SMAD5	0.68
3	KLHL3	0.68

4	JMJD1B	0.68
5	ARHGAP26	0.68
6	PCDHB2	0.65
7	PCDHB14	0.65
8	PCDHB5	0.65
9	PCDHB10	0.65
10	PCDHB1	0.65
11	GFRA3	0.65
12	PCDHB12	0.65
13	PCDHB15	0.65
14	PCDHB7	0.65
15	PCDHB16	0.65
16	WNT8A	0.65
17	PCDHB8	0.65
18	PCDHGA4	0.65
19	SLC23A1	0.61
20	PCDH1	0.61
21	TTID	0.61

(4) GeneSeeker^[67,68]

GeneSeeker 使用人类和小鼠表达和表型各种数据库 (<http://www.cmbi.ru.nl/GeneSeeker/>)。这些信息结合了定位区间内两个种属的基因数据。系统使用不同的在线数据库,而不是本地数据,因此能及时进行数据挖掘。GeneSeeker 方法不同于其它利用跨种属的候选区分优先次序的方法。

在 GeneSeeker 网页上 Genetic location 选择 5q31.1-5q32, Expression location 选择 ears/ear OR auditory OR cochlea, 其搜索结果显示, 位于 5q31.1-5q32 中在 ear/auditory/cochlea 有表达的人类基因有: DIAPH1, FGF1, NR3C1, POU4F3, TGFBI, ZMAT2。

1.2. 结果

根据上述耳蜗基因表达谱数据库, 基因功能, 耳聋临床表现特征, 小鼠耳聋致病基因和其它文献的研究结果, 以及识别疾病候选基因的生物信息学软件的预测结果, 我们选择了 DNFA52 区间的候选基因有 52 个, 见表 1-8(附录 1), 因此, 最终选择了 52 个候选基因进行突变检测。

表 1-8. DFNA 52定位区间内的候选基因表

基因名	基因描述	UniGene	gene expression in the developing ear	Human Cochlear cDNA library and EST database	NEIBank Libraries	cDNA,EST Blast	Reference	Function
ANKHD1	ankyrin repeat(锚蛋白重复) and KH domain containing 1				+			
ARHGAP26	Rho GTPase activating protein(GTP 酶活化蛋白) 26							
BRD8	bromodomain containing 8	+			+			
CDC23	CDC23 (cell division cycle 23 细胞分裂周期, yeast)	+		+	+			
CTNNA1	catenin 连环蛋白(cadherin-associated protein)	+	+	+	+	+		
EGR1	early growth response 1(早期生长反应因子 1)	+		+	+	+		
ETF1	eukaryotic 真核生物 translation termination factor 1							
FGF1	fibroblast growth factor 1(成纤维细胞生长因子)		+	+		+	+	
GFRA3	GDNF family receptor alpha 3	+		+	+	+	+	
GNPDA1	glucosamine-6-phosphate deaminase 1(葡萄糖胺-6-磷酸脱氨基酶)	+						
H2AFY	H2A histone family, member Y	+			+			
HBEGF	heparin-binding EGF-like growth factor		+			+		
HDAC3	histone deacetylase 3(组蛋白脱乙酰基酶 3)	+		+	+			
HNRPA0	heterogeneous nuclear ribonucleoprotein A0 不均一核内核糖核酸蛋白	+		+	+			
HSPA9B	heat shock 70kDa protein 9B (热休克蛋白)	+		+	+	+		
IK	IK cytokine, down-regulator of HLA II(细胞因子)					+	+	
JMJD1B	jumonji domain containing 1B	+			+			
KCTD16	potassium channel tetramerisation domain(钾通道四聚体化)containing 16							+
KIF20A	Kinesin(驱动蛋白) family member 20A					+		
KLHL3	kelch-like 3	+		+	+			
LARS	leucyl-tRNA synthetase(亮氨酸-tRNA 合成酶)	+			+			
LOC492311	similar to bovine IgA regulatory protein	+			+			
LRRTM2	leucine rich repeat transmembrane neuronal 2							+

MATR3	matrin 3 基质蛋白	+		+	+	+		
NDFIP1	Nedd4 family interacting protein 1	+			+		+	
NEUROG1	neurogenin 1 神经原		+			+	+	
NR3C1	nuclear receptor(细胞核受体) subfamily 3, group C	+		+	+			
NRG2	neuregulin 2 神经调节蛋白						+	
PACAP	proapoptotic caspase adaptor protein							
PAIP2	poly(A) binding protein interacting protein 2	+		+	+	+		
PCBD2	pterin-4 alpha-carbinolamine dehydratase/dimerization cofactor of hepatocyte nuclear factor 1 alpha (TCF1) 2				+			
PCDH1	protocadherin 1 原钙粘蛋白						+	
PCDHAC1	protocadherin alpha subfamily C, 1						+	
PCDHAC2	protocadherin alpha subfamily C, 2						+	
PCDHB3	protocadherin beta 3							+
PCDHB5	protocadherin beta 5	+			+			
PCDHB11	protocadherin beta 11	+						
PCDHGC3	protocadherin gamma subfamily C, 3	+			+		+	
PITX1	paired-like homeodomain transcription factor 1					+		
PPP2R2B	protein phosphatase 2 (formerly 2A), regulatory subunit	+		+	+			
SLC23A1	solute carrier family 23							+
SMAD5	SMAD, mothers against DPP homolog 5	+			+			
SPOCK1	sparc/osteonectin, cwcv and kazal-like domains proteoglycan (testican) 1	+	+	+	+	+		
TAF7	TAF7 RNA polymerase II, TATA box binding protein (TBP)-associated factor							
TCERG1	transcription elongation regulator 1	+		+	+			
TGFB1	transforming growth factor, beta-induced	+		+	+	+		
TRPC7	transient receptor potential cation channel, subfamily C							+
MYOT	Myotilin(TTID)						+	
UBE2D2	ubiquitin-conjugating enzyme E2D 2	+			+	+		
WNT8A	wingless-type MMTV integration site family, member 8A					+		
YIPF5	Yip1 domain family, member 5	+			+			
ZMAT2	zinc finger, matrin type 2	+			+			

1.3. 分析与讨论

我们之前运用基因组扫描的方法定位了一个新的常染色体显性遗传神经性耳聋位点 DFNA52，定位于 5q31.1-q32，D5S2056 至 D5S638 之间的区间内^[69]。定位于 DFNA52 区间的两个耳聋基因，POU4F3 (OMIM:602460)与 DIAPH1 (OMIM:602121)，未检测到突变。此外，DFNA52 的遗传区间与 DFNA54^[70]和 DFNB60 的区间有重叠，很有可能在 5q31 另有一个或多个非综合征型耳聋基因，见图 1-5。

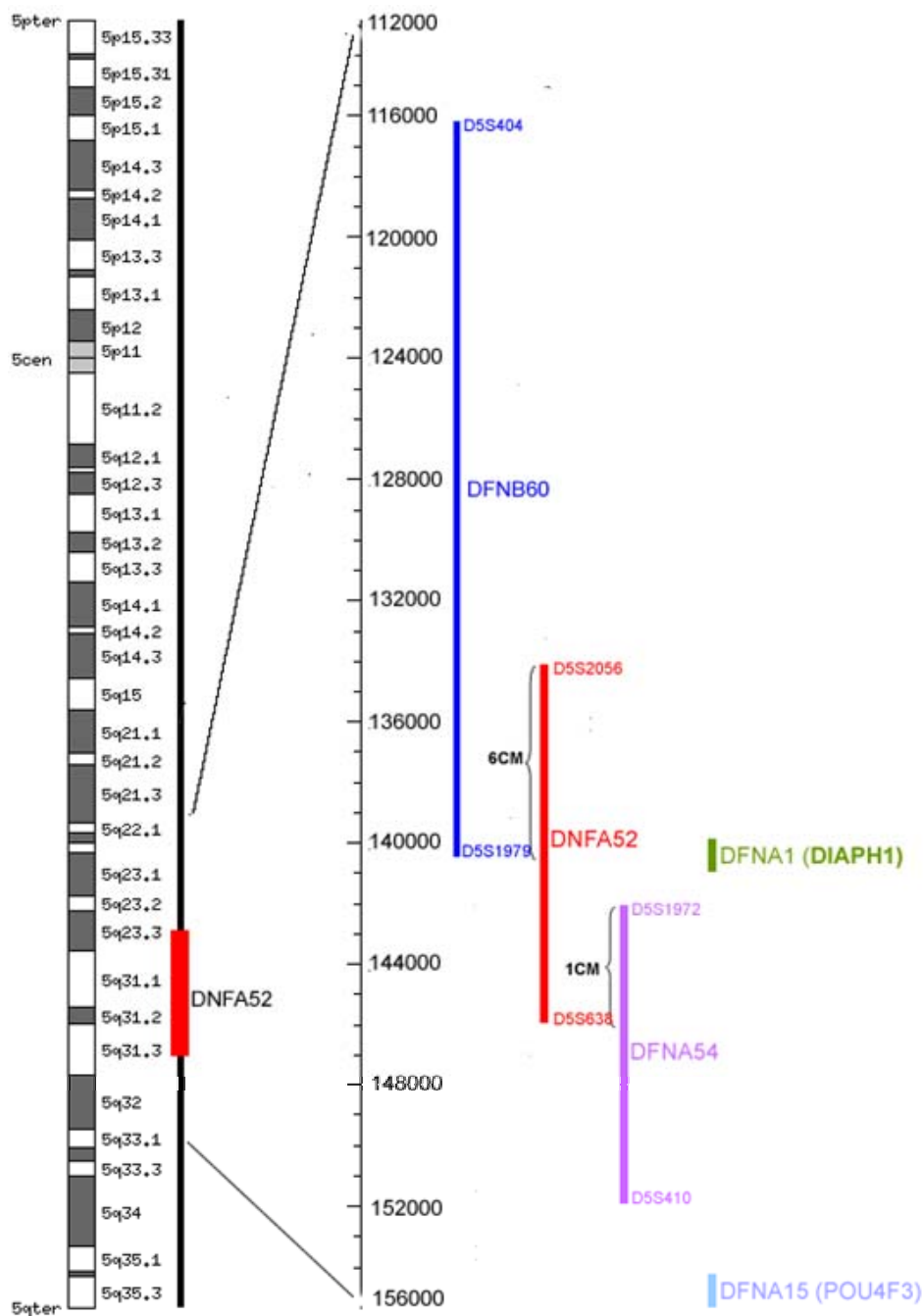


图 1-5 5q31 上非综合征型耳聋相关位点遗传距离示意图

为鉴定在耳蜗的基因表达情况和鉴定耳聋的候选基因,通过构建人类和小鼠的耳蜗 cDNA 文库,一些耳聋基因表达数据库已产生,促进了对耳蜗发育,内环境稳定性和听力下降的发病机理的理解^[72,73,57](Anne B. Skvorak 1999; Luijendijk MW 2003; Pompeia C 2004)。我们分析 UniGene 中的基因表达谱和耳基因表达数据库,包括耳发育过程中基因表达情况表^[55](table of gene expression in the developing ear)(Holme RH 2003), 人类耳蜗 cDNA 文库和 EST 数据库^[56] (Human Cochlear cDNA library and EST database)(Resendes BL 2002), 及 NEIBank 的耳文库^[57] (ear libraries on NEIBank)(Pompeia C 2004)。

连锁分析和突变筛选已较早的并成功地运用于染色体定位和鉴定致病基因。遵守孟德尔遗传的单基因病的致病突变的鉴定主要依赖于有好的病人和家系。虽然定位信息减少了致病候选基因的数目,这种减少经常不足以迅速鉴定疾病基因。分离这些位点的致病基因有时是困难的:定位区间过大,包含数百个可能的候选基因,使得实验室方法费时和费钱。有时在寻找到致病性突变前需要分析大量的基因。显然,高通量分析方法需要使用计算机方法来鉴定最可能的候选基因来进一步研究。随着基因组学和功能基因组学相关信息的飞速发展,许多候选疾病基因的生物信息学方法已形成。这些区分优先次序的方法选择最有可能的候选基因进行分析,将加速疾病基因的发现过程^[74]。

根据 NCBI 数据库的注释,DFNA52 区间遗传距离有 8.8cM,物理距离有 12Mb,包括已知基因和一些预测的或无特征的基因共 176 个。因这些导致常染色体显性遗传的非综合症型耳聋的基因是不同的,许多基因参与了听觉形成过程,包括间隙连接家族、细胞骨架蛋白、转录因子、离子通道、肌球蛋白和一些功能未知的基因,很难根据一个基因的已知功能或推测的功能来选择侯选基因。本研究根据基因表达、基因功能、疾病的表型、小鼠模型的遗传学研究(表型,表达)和发表的文献的研究情况以及利用识别候选疾病基因的生物信息学软件来选择候选基因,最终我们从 DFNA52 区间内优先选择了 52 个基因来进行突变筛选。

第二章 DFNA52 的候选基因的突变检测

本课题的研究内容是用定位候选克隆的策略克隆 DFNA52 的致病基因, 我们已通过全基因组扫描将一怀化耳聋家系的致病基因定位于 5q31-q32, 两点连锁分析在 D5S2017 得到最大 LOD 值 6.89, 精细定位提示疾病基因定位于 D5S2056 和 D5S638 之间 8.8cM 的区域, 本研究应用生物信息学方法分析定位区间的候选基因, 优先选择了 52 个基因进行突变检测分析, 以期能找到 DFNA52 的致病基因。

2.1. 材料与方法

2.1.1. 研究对象:

遗传了 5 代的常染色体显性遗传的湖南怀化 NSHL 家系共 47 人, 其中病人 18 个, 包括 6 名男性和 12 名女性(见图 1-2)。先证者 V-9, 女, 23 岁, 16 岁起双侧耳鸣并进行性听力减退, 不伴其它器官系统的异常。家系成员中耳聋患者的起病年龄在 16-30 岁, 均表现为双侧耳鸣, 听力减退为双耳对称性、以高频听力下降为主的感觉神经性耳聋, 逐渐发展到累及所有频率的重度听力减退。

2.1.2. 候选基因的选择:

从 DFNA52 区间内选择 52 个候选基因进行突变检测, 见表 1-8。

2.1.3. 引物设计

在基因的外显子以及外显子与内含子交界位置设计引物(部分检测基因的外显子个数、引物名称和各引物序列, 见附录 2)。设计引物使用加拿大生物研究所提供的 primer3 在线软件^[75](<http://www-genome.wi.mit.edu/cgi-bin/primer/primer3.cgi/>)或 Primer5。为了使 PCR 扩增有效性、特异性更好, 我们对以下参数进行了设定: 引物中的 GC 含量设定在 40%-60%之间; 退火温度设定在 57℃-63℃之间; 为了使引物更好的结合到模板链上, 引物的 3'端一般设定为以 G 或 C 结尾; 为了便于测序, 扩增片段一般设定在 200bp-500bp 之间。

2.1.4. PCR 扩增

10 μ l 反应体系见表 2-1; 突变检测中热启动 PCR 反应程序见图 2-1, 95℃ 15min 热启动; 94℃ 30s, 60℃ 30s, 72℃ 1min, 10 个循环; 94℃ 30s, 56℃ 30s,

72℃1min，25 个循环；72℃10min；4℃保存。选取两个患者进行 PCR 扩增。

表 2-1 突变检测 PCR 反应体系

试剂 (Reagents)	体积 (Volume)
10×Buffer	1 μl
Q-Buffer	2 μl
MgCl ₂	0.5 μl
dNTPs (10 mM)	0.2 μl
Hot Star Taq 酶 (5 U/μl)	0.05 μl
Primers (100ng/μl)	0.3 μl
DNA 模板 (50 ng/μl)	1 μl
ddH ₂ O Add to	10.0 μl

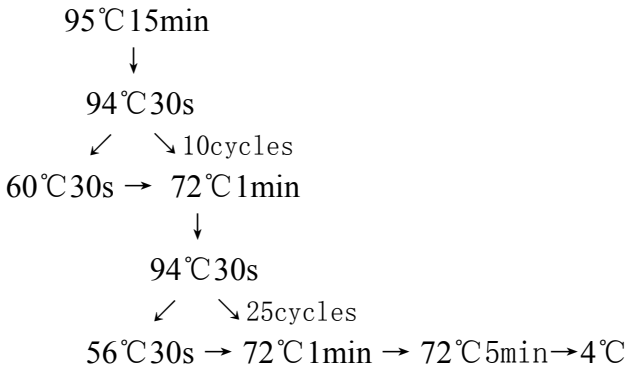


图 2-1 突变检测中热启动 PCR 反应程序

2. 1. 5. 聚丙烯酰胺凝胶电泳

取 2 μl PCR 反应产物，在 6%聚丙烯酰胺凝胶上(6%聚丙烯酰胺凝胶：30%丙烯酰胺 6ml，10×TBE 1.5ml，TEMED 25 μl，10%APS 300 μl，加一蒸去离子水至 30ml)，300V 电泳 40 分钟，硝酸银染色，检测 PCR 产物结果。若产物特异性好，为目的片段大小，则可测序。

2. 1. 6. 测序

电泳检测完后,对扩增特异性好的 PCR 产物用虾碱性磷酸酶和核酸外切酶处理（表 2-2），37℃酶切 40min, 80℃15min 灭活虾碱性磷酸酶和核酸外切酶，4℃保存。处理完后的产物用 3100 基因分析仪测序。

表 2-2 虾碱性磷酸酶和核酸外切酶处理的试剂组成

试剂 (Reagents)	体积 (Volume)
10×SAP Buffer	0.5 μl
SAP (1U/μl)	2μl

Exonuclease I (20U/ μ l)	0.25 μ l
PCR 产物	7.25 μ l

2.1.7. 测序结果分析

测序结果用 DNASTAR 软件中的 EditSeq、SeqMan 程序分析。全部测序结果与 GenBank 中登记的序列比对。每检测到一个变异, 用 SNPblast 程序 (www.ncbi.nlm.nih.gov/blast/)来排除多态, 然后用家系内的所有 DNA 样本来证实这个变异在家系内是否与疾病共分离。

2.2. 结果

2.2.1 突变检测结果

在对上述 52 个候选基因的编码区、剪切位点、部分内含子区和部分调控区的突变检测中, 15 个基因未发现任何序列变异, 在其余 37 个候选基因共发现 108 个 SNPs。查询单核苷酸多态数据库 dbSNP, 已报道的 SNPs 有 93 个, 未收录的有 15 种(测序结果见表 2-3)。经再对家系内其他患者、家系内正常人及家系外正常对照的突变分析, 结果表明这 15 个 SNPs 与疾病表型无共分离。

表 2-3 候选基因突变检测的突变情况

Genes(mRNA AC)	Loc.in gene	Base change	Codon change	refSNP ID
PCDHB5	Ex1	c.1996C>T	Gly>Gly	Not reported
		c.2375T>C	Ser>Pro	Reported rs400562
SPOCK	Prom	c.-153a>t c.-51_52ins+T		In bQSNP Quantification of SNP allele frequencies databases shown respectively as H13426、QH13427
	IVS5	Ivs5+90_91dup taaa		Reported rs34109254
HDAC3	5'UTR	c.-96g>c		Not reported
	IVS1	IVS1+27 a>g		Reported rs32954
	Ex3	c.165A>G	Gln>Gln	Reported rs2530223
	IVS11	IVS11-49 t>c		Reported rs2547547
HSPA9B	IVS2	IVS2-79_81 del tct		Not reported
	Ex9	c.948 A>G	Lys>Lys	Reported rs1042665
	Ex16	c.1933 C>T	Leu>Leu	Reported rs10117
PAIP2	3' UTR	c.*701_702insATT		Reported rs10672787
TGFBI	IVS1	IVS1-90 c>t		Reported rs10044956
	Ex6	c.651 g>c	Leu>Leu	Reported rs1442
	IVS4	IVS4+57 G>A		Reported rs3749786
	Ex7	c.45 A>C	Asn>Asn	Not reported
	Ex13	c.1803 G>A	Leu>Leu	Reported rs35151677
	IVS13	IVS13-71 a>g		Reported rs13168506
	IVS13	IVS13-55 a>t		Reported rs13188659
	IVS14	IVS14+44 t>c		Reported rs6880837
	IVS15	IVS15-73 g>a		Reported rs3749782
TCERG1	IVS1	IVS1+90 c>t		Reported rs10044956
	IVS2	IVS2-84 a>t		Reported rs2304055
	IVS8	IVS8+17 t>c		Reported rs12657671
	IVS21	IVS21-52 c>t		Reported rs4418147
	Ex22	c.3121A>G	Ser>Ser	Reported rs4705103
	3'UTR	c.*250C>T		Reported rs3822506
PPP2R2B	5'UTR	c.-118A>T (NM_181678)		Reported rs17524553
	5'UTR	c.-177C>T (NM_181674)		Reported rs11741933

KLHL3	Ex5	c.471 A>G	Ala>Ala	Reported	rs2905608
	IVS5	IVS5+21 t>a		Reported	rs2301710
	IVS8	IVS8-21 c>t		Reported	rs2074348
	Ex12	c.1383 G > A	Glu>Glu	Reported	rs2301708
	IVS12	IVS12+89 T>C		Reported	rs7713180
	IVS14	IVS14+29 C>T		Reported	rs2074345
	IVS14	IVS14+42 a>g		Reported	rs2074344
	IVS13	c.*3062a>g		Reported	rs3813314
	3'UTR	c.*4205 a>g		Reported	rs6863414
FGF	3'UTR	c.*1256G>T		Reported	rs34001
	3'UTR	c.*1353a>g		Reported	rs34000
	3'UTR	c.*1505G>C		Reported	rs33999
neurog1	5'UTR	c.-60T>G		Reported	rs8192558
KCTD16	IVS4	IVS4+15 c>t		Not reported	
pcdhgc3	Ex1	c.2115 T>C	Phe>phe	Reported	rs1423148
wnt8a	5'UTR	c.-51T>C		Reported	rs4574537
	IVS3	IVS3+51 t>c		Reported	rs4835762
	Ex6	c.741G>A	Glu>Glu	Reported	rs6596422
	3'UTR	c.*6C>T		Reported	rs12520886
	IVS7	IVS7+45g>C		Reported	rs3860103
Hbegf	Ex3	c.273c>t		Reported	rs415020
GNPDA1	IVS1	IVS1+103 c>g		Reported	rs2232193
	Ex2	c.69 c>t	Arg>arg	Reported	rs164080
PCBD2	3'UTR	c.*327g>a		Not reported	
JMJD1B	Ex2	c.264 c>t	Ile>Ile	Reported	rs4835678
	Ex6	c.766 G>A	Ala>Thr	Reported	rs6865472
	IVS11	IVS11+53 a>t		Reported	rs4835680
	Ex14	c.3489T>C	Gly>Gly	Reported	rs7726234
	IVS16	IVS16+54 a>c		Reported	rs740077
PCDH1	Ex3	c.1540 g>a	Ala>Thr	Reported	rs3822357
	Ex3	c.2250 t>c	Ala>Ala	Reported	rs3797054
PCDHAC1	Ex1	c.807 g>a	Gln>gln	Reported	rs155361
	Ex1	c.1269 c>g	Pro>pro	Reported	rs15582
	Ex1	c.1492 C>G	Leu>Val	Reported	rs246074
PCDHAC2	Ex1	c.117 t>a	Pro>pro	Reported	rs155364
	Ex1	c.361 t>c	Leu>leu	Reported	rs155808
	Ex1	c.702 C>T	Ile>Ile	Reported	rs155807
ARHGAP26	3'UTR	c.-125g>a		Reported	rs2276993
	Ex3	255 A>G	Arg>arg	Reported	rs185200
	IVS4	IVS4+73 c>a		Reported	rs27779
	IVS6	IVS6+43 a>g		Reported	rs37216
	IVS12	IVS12-20 t>c		Reported	rs192194
	Ex14	c.1245 T>G	Gly>gly	Reported	rs2270068
	IVS16	IVS16+97 g>a		Reported	rs6580266
	IVS18	IVS18+117 c>t		Reported	rs10045199
	IVS19	IVS19+16 a>t		Reported	rs246607
	Ex22	c.2355 c>t	Asn>Asn	Reported	rs258819
Lars	IVS19	IVS19+63t>c		Reported	rs2962504
PCDHB3	Ex1	c.2250G>A	Gln>Gln	Reported	rs17844406
Pitx1	Ex2	c.896G>C	Gly>Ala	Reported	rs479632
TTID	Ex2	c.220 A>C	Lys>Gln	Reported	rs6890689
SMAD5	IVS1	IVS1+39 t>c		Reported	rs7718415
	IVS4	IVS4+39 a>g		Reported	rs10479111
	IVS7	IVS7+37 a>g		Reported	rs4585442
	IVS7	IVS7-31 t>a		Reported	rs4146185
PACAP	5UTR	c.-436 G>A		Not reported	
	IVS2	IVS2+17 Del C	c.-388delC	Not reported	
PCDHB11	Ex1	c.1425 T>C	Ala>Ala	Not reported	
	Ex1	c.1584 C>G	Asp>Glu	Reported	rs799834
	Ex1	c.1873 C>A	Arg>Ser	Not reported	
	Ex1	c.2163 G>T	Ser>Ser	Reported	rs638054
SLC23A1	IVS1	IVS1+6_44del39bp		Not reported	
	Ex11	c.1261 T>G	Ser>Ala	Reported	rs6596474
ETF1	IVS9	IVS9+99 T>C		Reported	rs4835685

	IVS10	IVS10+84 t>c		Reported rs2242599
YIPF5	IVS1	IVS1-18 a>g		Not reported
ZMAT2	IVS1	IVS1-106 t>c		Reported rs801168
Trpc7	Ex1	c.501 C>T	Pro>Pro	Reported rs2546661
	IVS2	IVS2+42 c>t		Reported rs2277052
	Ex7	c.1918 C>T	Tyr>Tyr	Not reported
NRG2	IVS8	IVS8+13 c>g		Reported rs2287683
	Ex11	c.2203 C>T	Pro>Ser	Not reported
	Ex11	c.2246 G>T	Arg>Gln	Not reported
BRD8	IVS6	IVS6+87 c>t		Reported rs13155519
	IVS23	IVS23+37 g>a		Reported rs217270
	Ex26	c.3593 A>G		Reported rs412051
IK	IVS10	IVS10+25 g>a		Reported rs778596
	IVS19	IVS19+162 t>G		Reported rs4801
H2AFY	IVS2	IVS2+31 c>g		Reported rs17168241

2.2.2 测序结果图

15 个未报道的 SNPs 测序结果 (部分略, 见图 2-2 至 2-9), 对家系内其他患者、家系内正常人及家系外正常对照的检测, 这此变异在家系内与疾病表型不共分离。

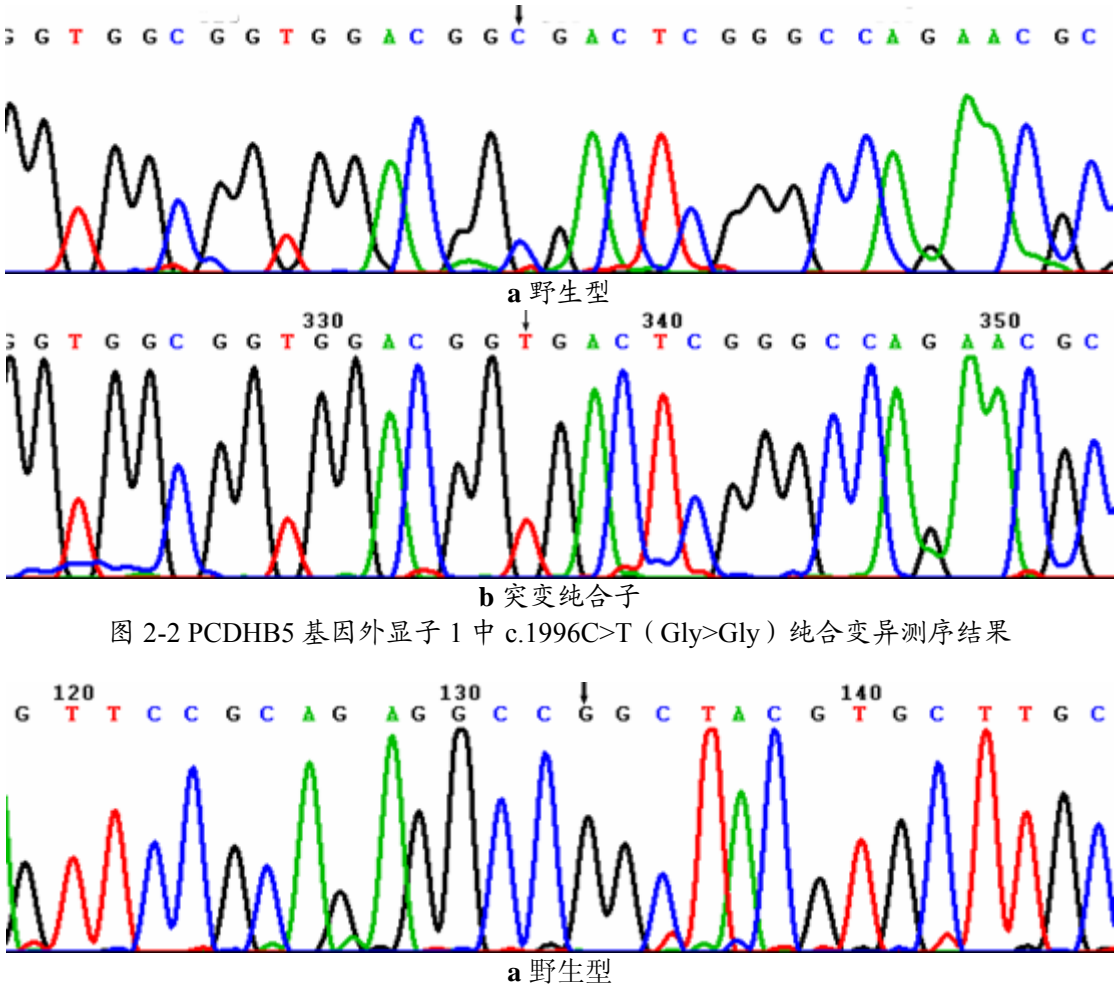


图 2-2 PCDHB5 基因外显子 1 中 c.199C>T (Gly>Gly) 纯合变异测序结果

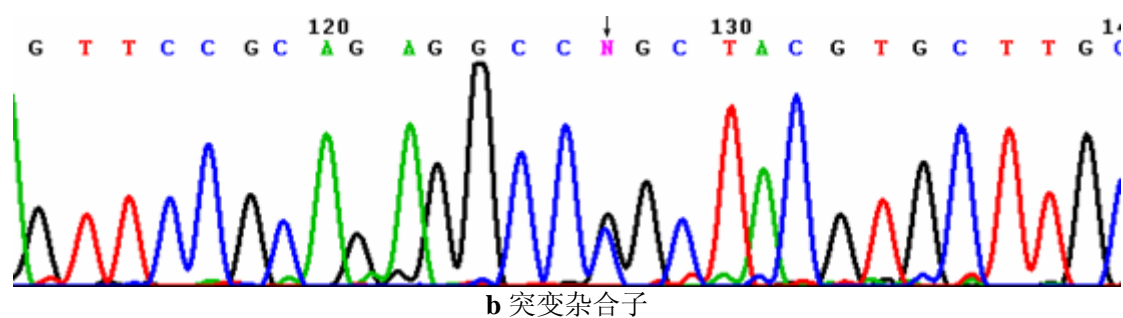


图 2-3 HDAC3 基因的 5'UTR 区 c.-96g>c 杂合变异测序结果

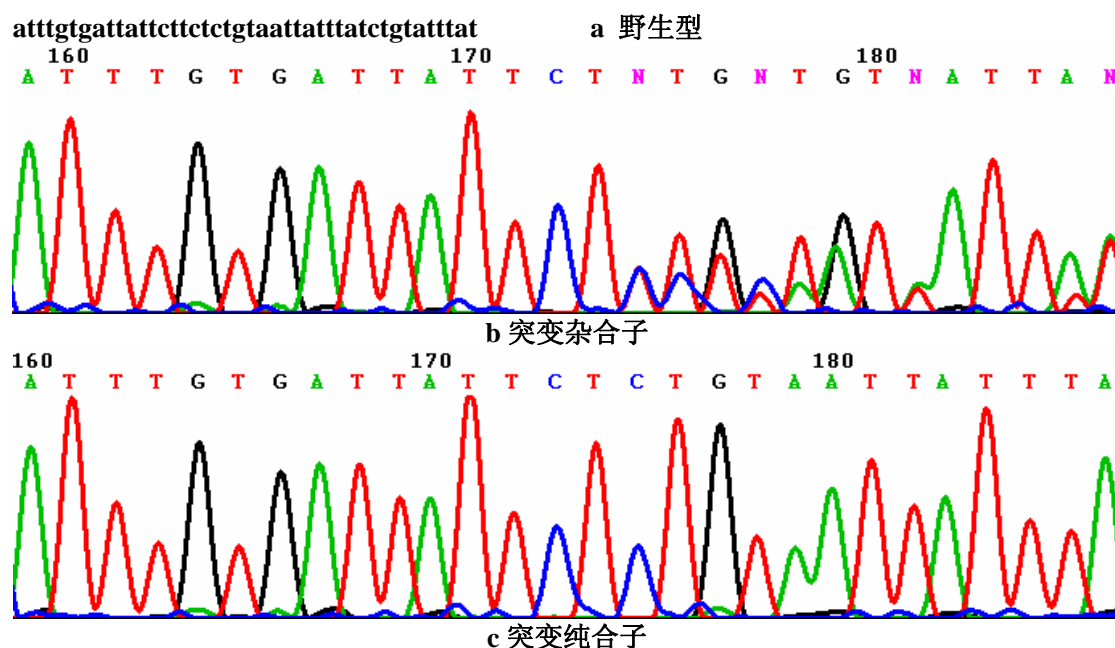


图 2-4 HSPA9 基因 IVS2 内 IVS2-79_81 del tct 杂合与纯合变异测序结果

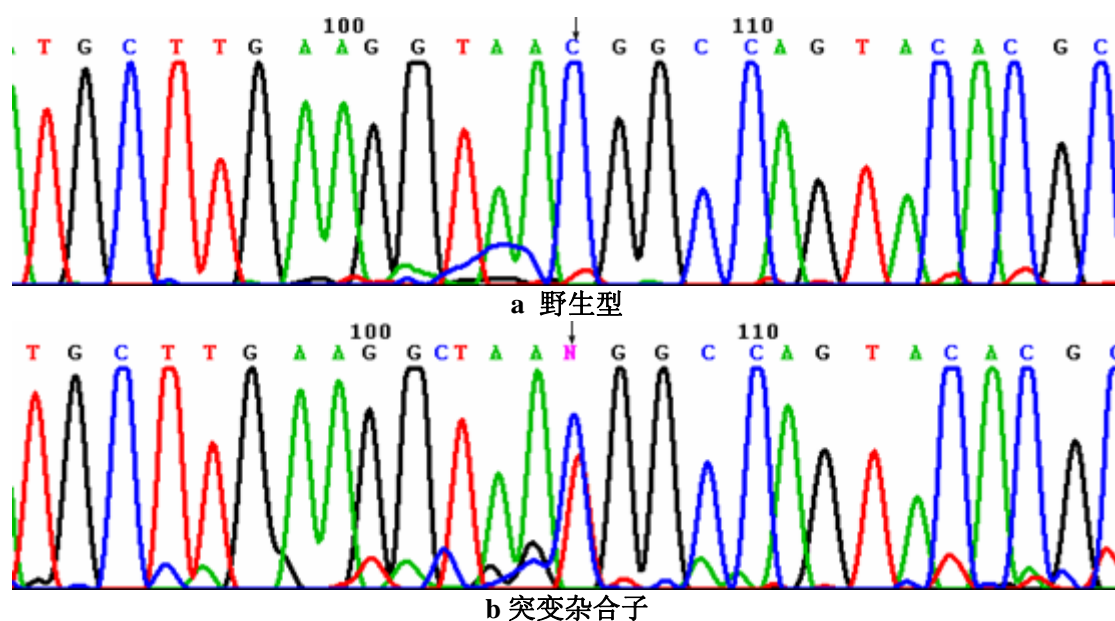


图 2-5 TGFBI 基因外显子内 c.45 C>T (Asn>Asn) 杂合变异测序结果

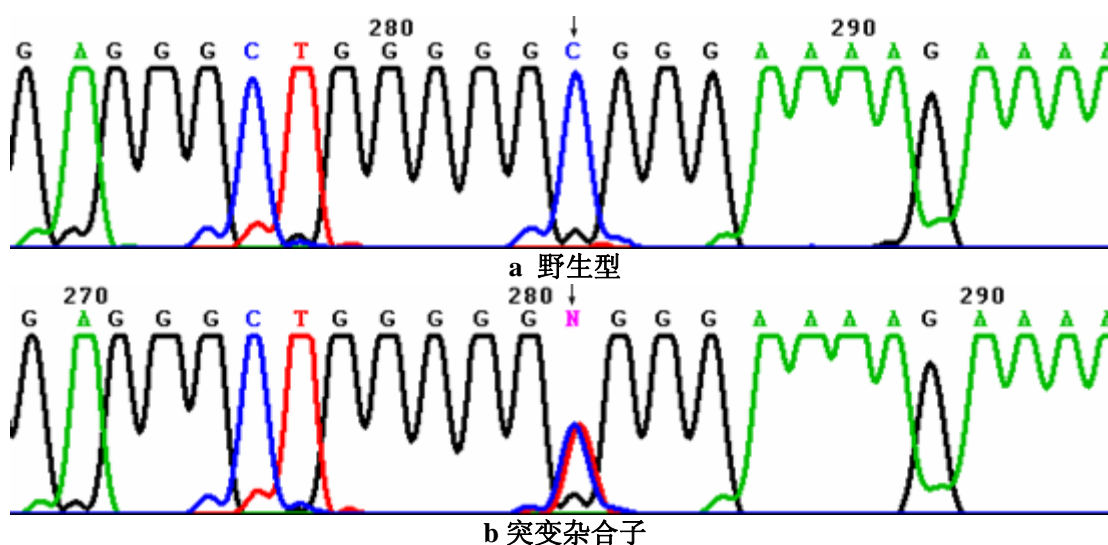


图 2-6 KCTD16 基因内含子 4 内 IVS4+15 c>t 杂合变异测序结果

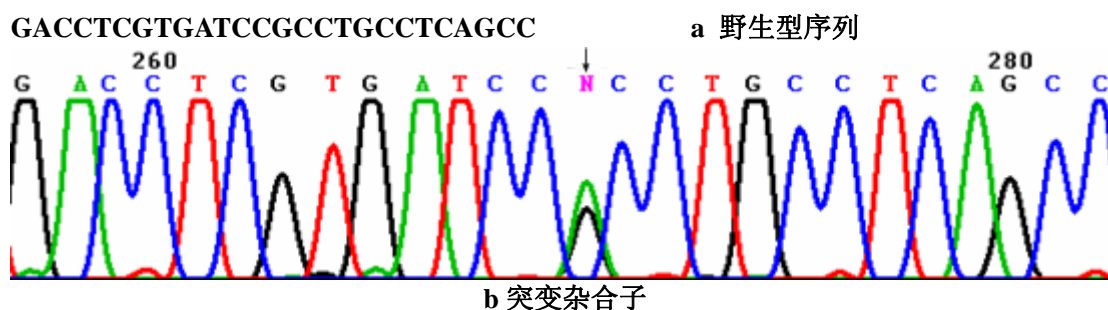


图 2-7 PCBD2 基因 3'UTR 内 c.*327g>a 杂合变异测序结果

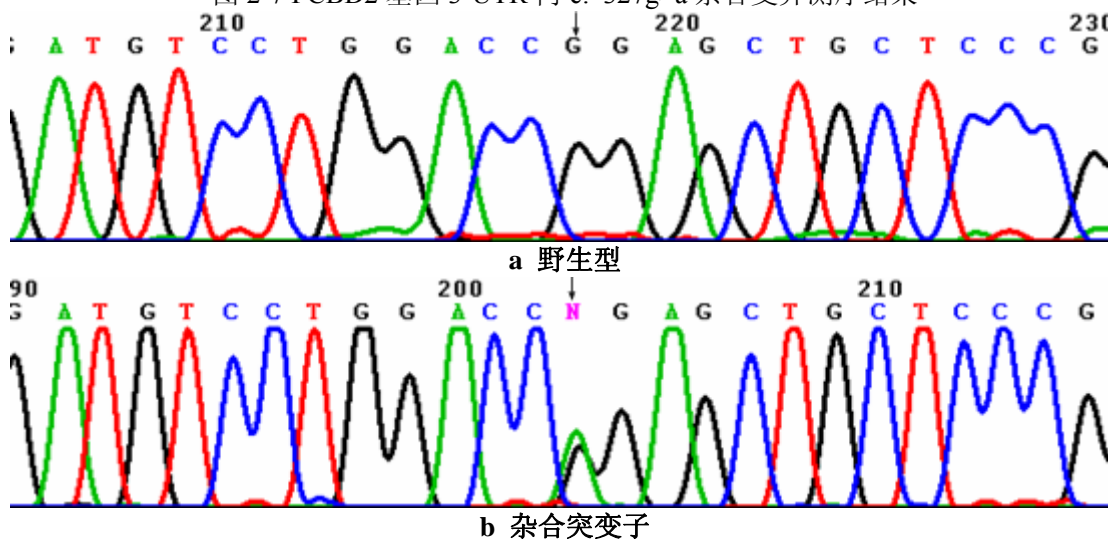


图 2-8 PACAP 基因 5UTR 内 c.-436 G>A 杂合变异测序结果

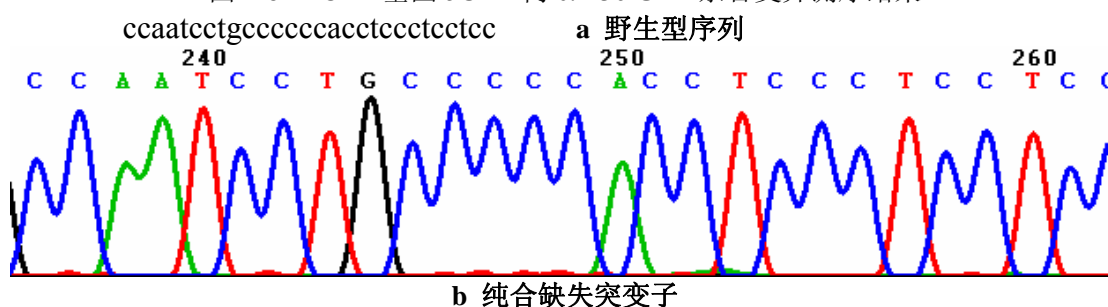


图 2-9 PACAP 基因 IVS2 内 IVS2+17 Del C(c.-388delC)纯合变异测序结果

2.3. 分析与讨论

我们从 DFNA52 区间内优先选择了 52 个基因来进行突变筛选。分析这些基因的编码区和剪切位点未发现与家系共分离的致病性突变。这表明这些基因可能均不是 DFNA52 的致病基因。

Neurog1(neurogenin 1)是在 DFNA52 位点区间内的最好的候选基因之一。原神经源性基因 neurogenin 1 (ngn1)是内耳感觉神经元发育所必要的,在 ngn1 无效突变体中是完全不表达的。Ngn1 无效突变体耳发育成较小的感觉上皮,毛细胞在形态上正常,但顺序紊乱且数目减少^[58](MA Q 2000)。我们测序了 Neurog1 基因的整个编码区,UTR 区及在上游的保守序列^[76](BLADER P 2004),以寻找致病性变异。突变结果分析显示整个基因仅一个序列改变,在 5'UTR 区有个 SNP(rs8192558),但是此变异在家系内与疾病不分离。

同时,另一个感兴趣的候选基因是溶质运载家族成员 SLC23A1。溶质运载家族成员之一的 SLC26A4 突变导致 DFNB4^[33](Li et al. 1998),Pendred 综合征^[77](Everett et al. 1997)和前庭水管扩大综合征^[78](Usami et al. 1999)。在包括东亚和南亚的不同的人群中遗传性耳聋 SLC26A4 的突变约占 10%^[79](Park et al. 2003)。然而,通过对家系内的两个受累患者的 DNA 样本测序分析,SLC23A1 所有的 15 个编码蛋白质的外显子均未发现有任何变异。

在我们研究中,选择 PCDHB5,PCDHAC1,PCDHAC2,PCDHB3,PCDHB11 和 CTNNA1 作为 DFNA52 位点的功能候选基因。钙粘蛋白基因(CDH23)已证实是 Usher 综合征 1D 型^[80]及 DFNB12 的致病基因^[37](Bolz et al. 2001; Bork et al. 2001)。已知编码原钙粘蛋白 15 的 PCDH15 基因导致 Usher 综合征 1F 型^[81](USH1F)和 DFNB23^[43]。而且,在小鼠耳聋突变体 waltzer (Di Palma F 2001)和 Ames-waltzer(Alagramam KN,2001)中报道是两个钙粘蛋白相关的基因。然而,对这 5 个基因的编码区进行突变检测其结果是阴性的。因此可排除它们是致病基因。但是,存在 3 个紧密连锁的串联在人类染色体 5q31 上的原钙粘蛋白基因族, α , β 和 γ 三个基因族,包括 57 个基因和 6 个假基因。尽管其它定位于区间内的原钙粘蛋白也适合作为 DFNA52 位点的候选基因,但筛选所有的原钙粘蛋白基因是比较困难的,因此其它原钙粘蛋白基因在本研究中没有全做突变检测。

SMAD5 基因是转化生长因子- β (transforming growth factor- β , TGF- β)信号下游细胞质内的信号转导分子^[84],它主要介导骨形态形成蛋白(bone morphogenetic proteins, BMPs)的信号,可将信号从细胞膜直接转导至细胞核,从而对 TGF- β 胞内信号转导起着关键的作用。国内杨仕明研究小组发现 SMAD5 基因敲除小鼠随着月龄增加而听力下降加重,那些敲除了 SMAD5 基因的杂合子小鼠表现出听力下降,而同样条件下的野生型小鼠则听力正常^[85,86]。在敲除

SMAD5 基因的小鼠内耳切片的研究中,发现了敲除鼠的内耳毛细胞存在缺失现象^[87,88],说明内耳组织结构的病变是小鼠听力下降的原因。而其后的研究发现 SMAD5 基因在小鼠内耳组织中确有表达,说明 SMAD5 基因在内耳听觉功能中可能具有一定的作用^[86]。并且有研究显示 Smad4 条件基因敲除后小鼠由于基因的缺陷而出现了重度感音神经性耳聋,纯合子小鼠出现全聋^[89]。我们在对 SMAD5 基因突变检测中,发现了四个已知 SNP,这些 SNPs 未引起氨基酸的改变。王秋菊等人对 SMAD5 基因在中国 143 名耳聋患者中的突变筛查研究中,也表明 SMAD5 在耳聋人群中可能不是造成耳聋的危险因素^[90]。

总之,我们在 DFNA52 位点的怀化耳聋家系中筛选了 52 个基因,尽管由于这些基因的表达谱和功能情况适合作为 DFNA52 的候选基因,但未发现有在家系内与疾病共分离的致病性突变。通常,这些发现提示这些基因不太可能是涉及 DFNA52。但是不能完全排除这些基因是 DFNA52 的致病基因,因为我们未筛查这些基因的调控区间(启动子, 5' 和 3' UTRs),在内含子区间内的功能变异的可能性尚未排除。此外,可能在此区间内存在新的致病基因有待于发现。新的候选基因还有待于经突变检测来证实。随着基因图谱的更新,基因表达情况的深入研究,也会发现新的在耳蜗表达的基因,其将成为新的候选基因。将来的工作包括继续筛查其它候选基因和/或在此家系研究中补充新的后代,和发现新的致病基因定位与 DFNA52 位点有重叠的耳聋家系,以缩小定位的遗传距离,这将有助于分离 DFNA52 的致病基因。

第一部分 结论

1. 运用生物信息学的方法, 综合耳蜗基因表达谱, 基因功能, 耳聋临床表现特征, 小鼠耳聋致病基因以及识别疾病候选基因的生物信息学软件的预测结果等资料, 建立了识别耳聋候选基因的策略, 在 DFNA52 定位区间内优先选择了 52 个候选基因(包括小鼠耳聋基因 *Neurog1*, *SMAD5*)。

2. 对 52 个候选基因的外显子以及外显子与内含子交界区域进行突变检测, 未发现致病性突变, 只发现 108 个 SNPs, 其中 15 个未报道。说明这些候选基因可能与 DFNA52 发病无关。

3. 小鼠耳聋相关基因 *Neurog1*, *SMAD5* 在怀化耳聋家系中未发现致病性突变, 提示这两个基因可能不是 DFNA52 的致病基因。

参考文献

- [1] Morton NE. Genetic epidemiology of hearing impairment. *Ann NY Acad Sci* 1991;630:16–31.
- [2] Marazita ML, Ploughman LM, Rawlings B, Remington E, Arnos KS. Genetic epidemiological studies of early-onset deafness in the US school-age population. *Am J Med Genet* 1993;46:486–91
- [3] Liu XZ, Xu LR, Hu Y, Nance WE, Sismanis A, Zhang SL, Xu Y. Epidemiological studies on hearing impairment with reference to genetic factors in Sichuan, China. *Ann Otol Rhinol Laryngol* 2001;110:356–63.
- [4] Van Camp G SR, Hereditary Hearing Loss Homepage <http://dnalab-www.uia.ac.be/dnalab/hhh/>. Mar, 2005.
- [5] Karen PS. New interventions in hearing impairment. *BMJ* 4:622-625 (2000).
- [6] Abe S, Katagiri T, Saito-Hisaminato A, et al., Identification of CRYM as a candidate responsible for nonsyndromic deafness, through cDNA microarray analysis of human cochlear and vestibular tissues. *Am J Hum Genet*, 2003, 72(1): 73-82.
- [7] Lynch ED, Lee MK, Morrow JE, et al., Nonsyndromic deafness DFNA1 associated with mutation of a human homolog of the *Drosophila* gene *diaphanous*. *Science*, 1997, 278(5341): 1315-1318.
- [8] Xia JH, Liu CY, Tang BS, et al., Mutations in the gene encoding gap junction protein beta-3 associated with autosomal dominant hearing impairment. *Nat Genet*, 1998, 20(4): 370-373.
- [9] Kubisch C, Schroeder BC, Friedrich T, et al., KCNQ4, a novel potassium channel expressed in sensory outer hair cells, is mutated in dominant deafness. *Cell*, 1999, 96(3): 437-446.
- [10] Kelsell DP, Dunlop J, Stevens HP, et al., Connexin 26 mutations in hereditary non-syndromic sensorineural deafness. *Nature*, 1997, 387(6628): 80-83.
- [11] Grifa A, Wagner CA, D'Ambrosio L, et al., Mutations in GJB6 cause nonsyndromic autosomal dominant deafness at DFNA3 locus. *Nat Genet*, 1999, 23(1): 16-18.
- [12] Donaudy F, Snoeckx R, Pfister M, et al., Nonmuscle myosin heavy-chain gene MYH14 is expressed in cochlea and mutated in patients affected by autosomal dominant hearing impairment (DFNA4). *Am J Hum Genet*, 2004, 74(4): 770-776.

- [13] Van Laer L, Huizing EH, Verstreken M, et al., Nonsyndromic hearing impairment is associated with a mutation in DFNA5. *Nat Genet*, 1998, 20(2): 194-197.
- [14] Bernalova IN, Van Camp G, Bom SJ, et al., Mutations in the Wolfram syndrome 1 gene (WFS1) are a common cause of low frequency sensorineural hearing loss. *Hum Mol Genet*, 2001, 10(22): 2501-2508.
- [15] Young TL, Ives E, Lynch E, et al., Non-syndromic progressive hearing loss DFNA38 is caused by heterozygous missense mutation in the Wolfram syndrome gene WFS1. *Hum Mol Genet*, 2001, 10(22): 2509-2514.
- [16] Verhoeven K, Van Laer L, Kirschhofer K, et al., Mutations in the human alpha-tectorin gene cause autosomal dominant non-syndromic hearing impairment. *Nat Genet*, 1998, 19(1): 60-62.
- [17] Robertson NG, Lu L, Heller S, et al., Mutations in a novel cochlear gene cause DFNA9, a human nonsyndromic deafness with vestibular dysfunction. *Nat Genet*, 1998, 20(3): 299-303.
- [18] Wayne S, Robertson NG, DeClau F, et al., Mutations in the transcriptional activator EYA4 cause late-onset deafness at the DFNA10 locus. *Hum Mol Genet*, 2001, 10(3): 195-200.
- [19] Liu XZ, Walsh J, Tamagawa Y, et al., Autosomal dominant non-syndromic deafness caused by a mutation in the myosin VIIA gene. *Nat Genet*, 1997, 17(3): 268-269.
- [20] McGuirt WT, Prasad SD, Griffith AJ, et al., Mutations in COL11A2 cause non-syndromic hearing loss (DFNA13). *Nat Genet*, 1999, 23(4): 413-419.
- [21] Vahava O, Morell R, Lynch ED, et al., Mutation in transcription factor POU4F3 associated with inherited progressive hearing loss in humans. *Science*, 1998, 279(5358): 1950-1954.
- [22] Lalwani AK, Goldstein JA, Kelley MJ, et al., Human nonsyndromic hereditary deafness DFNA17 is due to a mutation in nonmuscle myosin MYH9. *Am J Hum Genet*, 2000, 67(5): 1121-1128.
- [23] Zhu M, Yang T, Wei S, et al., Mutations in the gamma-Actin Gene (ACTG1) Are Associated with Dominant Progressive Deafness (DFNA20/26). *Am J Hum Genet*, 2003, 73(6): 1082-1091.
- [24] Melchionda S, Ahituv N, Bisceglia L, et al., MYO6, the human homologue of the gene responsible for deafness in Snell's waltzer mice, is mutated in autosomal dominant nonsyndromic hearing loss. *Am J Hum Genet*, 2001, 69(3): 635-640.

- [25] van Wijk E, Krieger E, Kemperman MH, et al., A mutation in the gamma actin 1 (ACTG1) gene causes autosomal dominant hearing loss (DFNA20/26). *J Med Genet*, 2003, 40(12): 879-884.
- [26] Peters LM, Anderson DW, Griffith AJ, et al., Mutation of a transcription factor, TFCP2L3, causes progressive autosomal dominant hearing loss, DFNA28. *Hum Mol Genet*, 2002, 11(23): 2877-2885.
- [27] Kurima K, Peters LM, Yang Y, et al., Dominant and recessive deafness caused by mutations of a novel gene, TMC1, required for cochlear hair-cell function. *Nat Genet*, 2002, 30(3): 277-284.
- [28] Donaudy F, Ferrara A, Esposito L, et al., Multiple mutations of MYO1A, a cochlear-expressed gene, in sensorineural hearing loss. *Am J Hum Genet*, 2003, 72(6): 1571-1577.
- [29] Liu XZ, Ouyang XM, Xia XJ, et al., Prestin, a cochlear motor protein, is defective in non-syndromic hearing loss. *Hum Mol Genet*, 2003, 12(10): 1155-1162.
- [30] del Castillo I, Villamar M, Moreno-Pelayo MA, et al., A deletion involving the connexin 30 gene in nonsyndromic hearing impairment. *N Engl J Med*, 2002, 346(4): 243-249.
- [31] Liu XZ, Walsh J, Mburu P, et al., Mutations in the myosin VIIA gene cause non-syndromic recessive deafness. *Nat Genet*, 1997, 16(2): 188-190.
- [32] Wang A, Liang Y, Fridell RA, et al., Association of unconventional myosin MYO15 mutations with human nonsyndromic deafness DFNB3. *Science*, 1998, 280(5368): 1447-1451.
- [33] Li XC, Everett LA, Lalwani AK, et al., A mutation in PDS causes non-syndromic recessive deafness. *Nat Genet*, 1998, 18(3): 215-217.
- [34] Naz S, Giguere CM, Kohrman DC, et al., Mutations in a novel gene, TMIE, are associated with hearing loss linked to the DFNB6 locus. *Am J Hum Genet*, 2002, 71(3): 632-636.
- [35] Scott HS, Kudoh J, Wattenhofer M, et al., Insertion of beta-satellite repeats identifies a transmembrane protease causing both congenital and childhood onset autosomal recessive deafness. *Nat Genet*, 2001, 27(1): 59-63.
- [36] Yasunaga S, Grati M, Cohen-Salmon M, et al., A mutation in OTOF, encoding otoferlin, a FER-1-like protein, causes DFNB9, a nonsyndromic form of deafness. *Nat Genet*, 1999, 21(4): 363-369.

- [37] Bork JM, Peters LM, Riazuddin S, et al., Usher syndrome 1D and nonsyndromic autosomal recessive deafness DFNB12 are caused by allelic mutations of the novel cadherin-like gene CDH23. *Am J Hum Genet*, 2001, 68(1): 26-37.
- [38] Verpy E, Masmoudi S, Zwaenepoel I, et al., Mutations in a new gene encoding a protein of the hair bundle cause non-syndromic deafness at the DFNB16 locus. *Nat Genet*, 2001, 29(3): 345-349.
- [39] Ouyang XM, Xia XJ, Verpy E, et al., Mutations in the alternatively spliced exons of USH1C cause non-syndromic recessive deafness. *Hum Genet*, 2002, 111(1): 26-30.
- [40] Ahmed ZM, Smith TN, Riazuddin S, et al., Nonsyndromic recessive deafness DFNB18 and Usher syndrome type IC are allelic mutations of USH1C. *Hum Genet*, 2002, 110(6): 527-531.
- [41] Mustapha M, Weil D, Chardenoux S, et al., An alpha-tectorin gene defect causes a newly identified autosomal recessive form of sensorineural pre-lingual non-syndromic deafness, DFNB21. *Hum Mol Genet*, 1999, 8(3): 409-412.
- [42] Zwaenepoel I, Mustapha M, Leibovici M, et al., Otoancorin, an inner ear protein restricted to the interface between the apical surface of sensory epithelia and their overlying acellular gels, is defective in autosomal recessive deafness DFNB22. *Proc Natl Acad Sci USA*, 2002, 99(9): 6240-6245.
- [43] Ahmed ZM, Riazuddin S, Ahmad J, et al., PCDH15 is expressed in the neurosensory epithelium of the eye and ear and mutant alleles are responsible for both USH1F and DFNB23. *Hum Mol Genet*, 2003, 12(24): 3215-3223.
- [44] Khan SY, Ahmed ZM, Shabbir MI, Kitajiri SI, Kalsoom S, Tasneem S, Shayiq S, Ramesh A, Srisailpathy S, Khan SN, Smith RJ, Riazuddin S, Friedman TB (2007) Mutations of the RDX gene cause nonsyndromic hearing loss at the DFNB24 locus. *Hum Mutat*. 28(5):417-423
- [45] Riazuddin S., Khan S. N., Ahmed Z. M., Ghosh M., Caution K., Nazli S., Kabra M., Zafar A. U., Chen K., Naz S., Antonellis A., Pavan W. J., Green E. D., Wilcox E. R., Friedman P. L., Morell R. J., and Friedman T. B. (2006). Mutations in TRIOBP, which encodes a putative cytoskeletal-organizing protein, are associated with nonsyndromic recessive deafness. *Am J Hum Genet* 78: 137-43.
- [46] Shahin H., Walsh T., Sobe T., Abu Sa'ed J., Abu Rayan A., Lynch E. D., Lee M. K., Avraham K. B., King M. C., and Kanaan M. (2006). Mutations in a novel isoform of TRIOBP that encodes a filamentous-actin binding protein are

- responsible for DFNB28 recessive nonsyndromic hearing loss. *Am J Hum Genet* 78: 144-52.
- [47] Wilcox ER, Burton QL, Naz S, et al., Mutations in the gene encoding tight junction claudin-14 cause autosomal recessive deafness DFNB29. *Cell*, 2001, 104(1): 165-172.
- [48] Walsh T, Walsh V, Vreugde S, et al., From flies' eyes to our ears: mutations in a human class III myosin cause progressive nonsyndromic hearing loss DFNB30. *Proc Natl Acad Sci USA*, 2002, 99(11): 7518-7523.
- [49] Mburu P, Mustapha M, Varela A, et al., Defects in whirlin, a PDZ domain molecule involved in stereocilia elongation, cause deafness in the whirler mouse and families with DFNB31. *Nat Genet*, 2003, 34(4): 421-428.
- [50] Naz S, Griffith AJ, Riazuddin S, et al., Mutations of ESPN cause autosomal recessive deafness and vestibular dysfunction. *J Med Genet*, 2004, 41(8): 591-595.
- [51] Ahmed ZM, Morell RJ, Riazuddin S, et al., Mutations of MYO6 are associated with recessive deafness, DFNB37. *Am J Hum Genet*, 2003, 72(5): 1315-1322.
- [52] Shabbir, M.I., Ahmed, Z.M., Khan, S.Y., Riazuddin, S., Waryah, A.M., Khan, S.N., Camps, R.D., Ghosh, M., Kabra, M., Belyantseva, I.A. and Friedman, T.B. (2006). Mutations of human TMHS cause recessively inherited nonsyndromic hearing loss. *J Med Genet* 43: 634-40.
- [53] J. Sambrook E.F. Fritsch T. Maniatis. *Molecular cloning a laboratory manual* 2nd ed . Cold spring Harbor Laboratory press. 1989
- [54] Lathrop GM, Lalouel JM. Easy calculations of lod scores and genetic risks on small computers. *Am J Hum genet* 1984;36:460-465 .
- [55] HOLME RH, B. T., STEEL KP, 2003 Table of gene expression in the developing ear. [<http://www.ihr.mrc.ac.uk/Hereditary/genetable>].
- [56] Resendes BL, R. N., Szustakowski JD, Resendes RJ, Weng Z, Morton CC. , 2002 Gene Discovery in the Auditory System: Characterization of Additional Cochlear-Expressed Sequences. *J Assoc Res Otolaryngol*. 3: 45-53.
- [57] Pompeia C, H. B., Belyantseva ia, Noben-Trauth K, Beisel K, Gao J, Buchoff P, Wistow G, Kachar B, 2004 Gene expression profile of the mouse organ of Corti at the onset of hearing. *Genomics* 83: 1000-1011.
- [58] Ma q, A. D., Fritsch B. , 2000 Neurogenin 1 null mutant ears develop fewer, morphologically normal hair cells in smaller sensory epithelia devoid of innervation. *J Assoc Res Otolaryngol*. 1: 129-143.

- [59] Lanctot C, Moreau A, Chamberland M, Tremblay ML, Drouin J. Hindlimb patterning and mandible development require the Ptx1 gene. *Development*, 1999, 126(9):1805-10.
- [60] Coucke, P. J. V. H., P.Kelley, P. M.Kunst, H.Schatteman, I.Van Velzen, et al. Mutations in the KCNQ4 gene are responsible for autosomal dominant deafness in four DFNA2 families. *Hum. Molec. Genet*, 1999, 8: 1321-1328.
- [61] Abe S, Katagiri T, Saito-Hisaminato A, et al. Identification of CRYM as a candidate responsible for nonsyndromic deafness, through cDNA microarray analysis of human cochlear and vestibular tissues. *Am J Hum Genet*, 2003, 72:73-82
- [62] Alsaber R, Tabone CJ, Kandpal RP. Predicting candidate genes for human deafness disorders: a bioinformatics approach. *BMC Genomics*. 2006 Jul 19; 7:180.
- [63] Perez-Iratxeta, C., Wjst, M., Bork, P. and Andrade, M.A. G2D: a tool for mining genes associated with disease. *BMC Genet*. 2005, 6, 45.
- [64] Perez-Iratxeta, C., Bork, P. and Andrade, M.A. Association of genes to genetically inherited diseases using data mining, *Nat. Genet*. 2002, Vol. 31, pp. 316–319.
- [65] Lopez-Bigas, N. and Ouzounis, C.A. Genome-wide identification of genes likely to be involved in human genetic disease, *Nucleic Acids Res*. 2004, Vol. 32, pp. 3108–3114.
- [66] Adie, E.A., Adams, R.R. and Evans, K.L. Speeding disease gene discovery by sequence based candidate prioritization, *BMC Bioinformatics*, 2005 Vol. 6, p. 55.
- [67] van Driel, M.A., Cuelenaere, K., Kemmeren, P.P. et al. (2003), A new web-based data mining tool for the identification of candidate genes for human genetic disorders, *Eur. J. Hum. Genet*. Vol. 11, pp. 57–63.
- [68] van Driel, M.A., Cuelenaere, K., Kemmeren, P.P. et al. (2005), GeneSeeker: Extraction and integration of human disease-related information from web-based genetic databases, *Nucleic Acids Res*. Vol. 33, pp. W758–W761.
- [69] Xia, J. D., H.; Feng, Y.; Zhang, H.; Pan, Q.; Dai, H.; Long, Z. et al, A novel locus for autosomal dominant nonsyndromic hearing loss identified at 5q31.1-32 in a Chinese pedigree. *J. Hum. Genet*, 2002, 47: 635-640.
- [70] Gurtler N, Kim Y, Mhatre A, Schlegel C, Mathis A, Lalwani AK. et al. DFNA54, a third locus for low-frequency hearing loss. *J Mol Med*, 2004 82(11):775-80.

- [71] Delmaghani S, del Castillo FJ, Michel V, Leibovici M, Aghaie A, Ron U, Van Laer L, Ben-Tal N, Van Camp G, Weil D, Langa F, Lathrop M, Avan P, Petit C. Mutations in the gene encoding pejvakin, a newly identified protein of the afferent auditory pathway, cause DFNB59 auditory neuropathy. *Nat Genet.* 2006 Jul;38(7):770-8.
- [72] Anne B. Skvorak, Z. W., Andrew J. Yee, Nahid G. Robertson and Cynthia C. Morton, Human cochlear expressed sequence tags provide insight into cochlear gene expression and identify candidate genes for deafness. *Human Molecular Genetics*, 1999, Vol. 8, No. 3 8: 439-452.
- [73] Luijendijk Mw, V. D. P. T., Van Duijnhoven G, Den Hollander Ai, Ten Caat J, Van Limpt V, Brunner Hg, Kremer H, Cremers Fp. Cloning, characterization, and mRNA expression analysis of novel human fetal cochlear cDNAs. *Genomics* , 2003, 82: 480-490.
- [74] van Driel MA, Brunner HG. Bioinformatics methods for identifying candidate disease genes. *Hum Genomics.* 2006 ,2(6):429-32.
- [75] Steve Rozen and Helen J. Skaletsky . Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ, 2000,pp 365-386
- [76] Blader P, L. C., Rastegar S, Scardigli R, Nicod Jc, Simplicio N, Plessy C, Fischer N, Schuurmans C, Guillemot F, Strahle U. Conserved and acquired features of neurogenin1 regulation. *Development* .2004,131: 5627-5637.
- [77] Everett LA, Glaser B, Beck JC, Idol JR, Buchs A, Heyman M, Adawi F, Hazani E, Nassir E, Baxevanis AD, Sheffield VC, Green ED. Pendred syndrome is caused by mutations in a putative sulphate transporter gene (PDS). *Nat Genet.* 1997 Dec;17(4):411-22.
- [78] Usami S, Abe S, Weston MD, Shinkawa H, Van Camp G, Kimberling WJ. Non-syndromic hearing loss associated with enlarged vestibular aqueduct is caused by PDS mutations. *Hum Genet.* 1999 Feb;104(2):188-92.
- [79] Park HJ, Shaukat S, Liu XZ, Hahn SH, Naz S, Ghosh M, Kim HN, Moon SK, Abe S, et.al. Origins and frequencies of SLC26A4 (PDS) mutations in east and south Asians: global implications for the epidemiology of deafness. *J Med Genet.* 2003 Apr;40(4):242-8.
- [80] Bork JM, Peters LM, Riazuddin S, Bernstein SL, Ahmed ZM, Ness SL,

- Polomeno R, Ramesh A, Schloss M, Srisailpathy CR, Wayne S, Bellman S, Desmukh D, Ahmed Z, et al. Usher syndrome 1D and nonsyndromic autosomal recessive deafness DFNB12 are caused by allelic mutations of the novel cadherin-like gene CDH23. *Am J Hum Genet.* 2001 Jan;68(1):26-37.
- [81] Ahmed ZM, Riazuddin S, Bernstein SL, Ahmed Z, Khan S, Griffith AJ, Morell RJ, Friedman TB, Riazuddin S, Wilcox ER. Mutations of the protocadherin gene PCDH15 cause Usher syndrome type 1F. *Am J Hum Genet.* 2001 Jul;69(1):25-34.
- [82] Di Palma F, Holme RH, Bryda EC, Belyantseva IA, Pellegrino R, Kachar B, Steel KP, Noben-Trauth K. Mutations in Cdh23, encoding a new type of cadherin, cause stereocilia disorganization in waltzer, the mouse model for Usher syndrome type 1D. *Nat Genet.* 2001 Jan;27(1):103-7.
- [83] Alagramam KN, Murcia CL, Kwon HY, Pawlowski KS, Wright CG, Woychik RP. The mouse Ames waltzer hearing-loss mutant is caused by mutation of Pcdh15, a novel protocadherin gene. *Nat Genet.* 2001 Jan;27(1):99-102.
- [84] Ten-Dijke P, Hill CS. New insights into TGF-beta-Smad signalling. *Trends Biochem Sci.* 2004;29(5):265-273
- [85] 杨仕明, 刘清明, 郭维, 等. Smad5 基因敲除小鼠听生理和耳蜗形态实验观察. *解放军医学杂志*, 2005, 30(7): 579—581.
- [86] 郭维、杨仕明、刘清明、孙彦询, 等. Smad5 基因缺陷导致小鼠听力重度损失. *中华耳科学杂志*, 2006, 4(3):190-193
- [87] 郭维, 杨仕明, 胡吟燕, 等. Smad5 基因在小鼠耳蜗中的表达及定位. *解放军医学杂志*, 2005, 30(7): 585—586.
- [88] 孙建和, 杨仕明, 孙彦询, 杨晓, 等. Smad5 基因敲除小鼠耳蜗扫描电镜观察, 2006, 4(3):194-196
- [89] 侯昭晖、杨仕明、郭维、胡吟燕, 等. Smad4 条件基因敲除小鼠听功能初步研究, 2006, 4(3):176-180
- [90] 王秋菊, 李庆忠, 纵亮, 郭维等. SMAD5 基因在中国耳聋患者中的突变筛查研究. *中华耳科学杂志*, 2006, 4(3):200-204
- [91] Modamio-Hoybjor S, Mencia A, Goodyear R, Del Castillo I, Richardson G, Moreno F, Moreno-Pelayo MA. A Mutation in CCDC50, a Gene Encoding an Effector of Epidermal Growth Factor-Mediated Cell Signaling, Causes Progressive Hearing Loss. *Am J Hum Genet.* 2007, 80:1076-1089.

第二部分 LOR 基因 730insG 突变导致的兜甲蛋白角化病

前 言

鱼鳞病(Ichthyosis)是一种以皮肤干燥、伴有鱼鳞状鳞屑为特征的遗传性角化障碍性疾病。临床上分为两种主要形式：①基本形式：病变主要局限在皮肤，根据遗传方式、形态学和组织学上不同，又可分为四型：显性遗传寻常型鱼鳞病、性连锁隐性遗传鱼鳞病、显性遗传先天性鱼鳞病样红皮病、隐性遗传先天性鱼鳞病样红皮病。此外还有四型少见的鱼鳞病即显性遗传可变性红斑角化病、隐性遗传迂回线状鱼鳞病、胎儿鱼鳞病(丑胎)、火棉胶婴儿。②鱼鳞病样综合征：病变累及到多系统、多器官，鱼鳞病只是多系统病变的一部分，如 Rud 综合征、Refsum 综合征、Sjogren-Larsson 综合征、Netherton 综合征、Vohwinkel 综合征(又称残毁性掌跖角皮症)等。在鱼鳞病中最常见的、多发的是寻常型鱼鳞病，在所有的鱼鳞病患者中约占 95%，其发病率约在每 250 个人中就有一个患者^[1]。在中国人群中，发病率约为 2.29%^[2]。在鱼鳞病样综合征中一种“兜甲蛋白角化病(Loricrin keratoderma)”已引起一些学者的关注，兜甲蛋白角化病是一个新的疾病术语，指一类由 LOR 基因突变所引起的遗传性皮肤病。本前言中重点介绍寻常性鱼鳞病和兜甲蛋白角化病。

1. 寻常性鱼鳞病

寻常性鱼鳞病(Ichthyosis vulgaris, IV, OMIM 146700)是一种存在表型多样性的常染色体显性遗传的皮肤疾病。寻常型鱼鳞病是鱼鳞病中最常见的、多发性的，在所有的鱼鳞病患者中约占 95%，其发病率约在每 250 个人中就有一个患者^[1]。在中国人群中，发病率约为 2.29%^[2]。在幼儿期就可表现为症状，其外显率约 91%或完全外显^[3,1]。寻常性鱼鳞病主要临床特点是：出生时皮肤外观正常，出生数月后发病，此后随年龄的增长病情逐渐加重，皮肤干燥，皮肤表面出现细薄的片状鳞屑，皮损以四肢伸侧为重尤以小腿明显，同时伴有掌跖角化增厚，手背、上臂、股外侧皮肤增厚，有时还伴有指甲变脆、毛发稀疏等。一般患者在 1~4 岁时症状已经较明显。对称性鱼鳞状皮肤可表现不同的情况，从几乎不可见的皮肤干燥粗糙至厚的角化片状体。掌和趾常出现过度角化，导致表现干燥。在颊和颈的一侧，在上臂的背面，臀部和腿，可出现毛发角化(滤泡角化过度)。皮损在冬季寒冷干燥季节加重，在潮湿温暖的夏季减轻。部分患者成年后皮损可逐渐减轻。寻常性鱼鳞病病理特征表现为中度角化过度，表皮变薄，角质层轻、中度增厚，

颗粒层减少或缺乏，其颗粒层松弛或包含形态异常的角质透明蛋白粒(keratohyalin granules)^[4-6]，生发层显得扁平，毛囊孔和汗腺可以有角质栓塞，皮脂腺减少，真皮血管周围有浸润。

关于寻常型鱼鳞病的发病机制还不清楚，定位并克隆寻常性鱼鳞病将有助于阐明其分子基础。由于获得合适的寻常型鱼鳞病家系比较困难，所以关于寻常型鱼鳞病致病基因的连锁分析及定位研究较少。2002年，Compton 等人对一个美国的以颗粒层缺失为表型的寻常性鱼鳞病家系定位研究，其研究显示与1q21上的表皮分化基因簇(EDC)连锁^[7]。2003年，zhong 通过连锁分析对两个中国家系中鉴定出一个寻常性鱼鳞病位点位于染色体1q22上，在D1S1653($\theta=0.00$)上其两点 LOD 值为 2.47^[8]。Zhong 研究结果表明 1q22 上的寻常性鱼鳞病位点与 EDC 无重叠，但从 D1S1653 至 EDC 的基因组 DNA 仅有 4Mb。2006 年，Smith 等研究中度或重度表型的 15 个有血缘的鱼鳞病患者，鉴定 FLG 基因有纯合型或混合杂合型突变 R501X 和 2282del4，首次克隆了一个 IV 的致病基因—filaggrin(FLG)^[3]，提供了丝聚合蛋白基因可以导致鱼鳞病的直接证据。

2. 兜甲蛋白角化病

兜甲蛋白角化病(Loricrin keratoderma)是一个新的疾病术语，指一类由 LOR 基因突变所引起的遗传性皮肤病。兜甲蛋白角化病的表型具有多样性。临床上，这些患者被诊断为鱼鳞病型 Vohwinkel 综合征^[10-16] (残毁性掌跖角皮症)，进行性对称性红皮肤角化病(erythrokeratoderma)^[17]，或出生时为火棉胶婴儿的先天性鱼鳞病样红皮病。

伴随鱼鳞病的残毁性掌跖角化病是一种罕见的常染色体显性遗传病。主要表现为自幼掌趾部弥漫性角化过度，表面呈蜂窝状小凹陷；指或趾的中间指、趾骨被收缩的纤维束围绕，形成“假性阿洪病”，一般第 5 趾最常受累，进行性的可以全部累及，严重地发生趾（指）自行离断；伴非掌跖部位的全身性非红斑性鱼鳞病。在多个随鱼鳞病的残毁性掌跖角化病的家系中，发现 LOR 羟基末端连续 6 个 G 碱基后单个 G 碱基插入（730insG）导致其后的甘氨酸环及高谷氨酰胺-赖氨酸区翻译编码框移位，改变了其后的 84 个氨基酸密码子的编码，并使终止密码子延迟出现，导致突变型蛋白较野生型延长 22 个氨基酸。

进行性对称性红斑角皮症是常染色体显遗传性皮肤病。典型表型常为固定的、边界清楚的角化性红色斑块，常在出生后不久发病，皮损逐渐加重，在青春期皮损分布最为广泛，以后可逐渐消退。主要表现为双侧掌跖部位的弥漫性红斑，表面逐渐出现角化过度性斑块，皮损逐渐扩大，可累及手背，足背，脛前，肘，膝以及大腿伸侧，少数可累及臀部及颈、面部。皮损边界清楚，伴角化过度性鳞屑，对称分布。可有指（趾）甲增厚失去光泽。病理主要表现为表皮角化过度，

伴有角化不全,棘层明显增厚,真皮非特异性炎症细胞浸润。Ishida-Yamamoto 在进行性对称性红斑角皮症患者中,发现 LOR 羟基末端连续 4 个 C 碱基后单个 C 碱基插入 (709insC), 导致其后的翻译编码框移位, 改变了其后 91 个氨基酸密码子的编码, 并使终止密码子延迟出现, 导致突变型蛋白较野生型延长 20 个氨基酸^[17]。

火棉胶婴儿系属于几种遗传型的混合型。分娩后, 新生儿皮肤光亮紧张, 如同上过釉, 整个皮肤似一层干燥的火棉胶薄膜, 此膜发光, 呈棕黄色。火棉胶样膜止于腔口周围, 粘膜不受累。头发通常穿膜而出。指甲偶尔为膜所盖。在面部, 常有双侧性睑外翻, 口唇半外翻, 鼻塌, 鼻孔为角质碎片堵塞, 需加以清除。耳朵因火棉胶样损害而变形, 形成皱褶。皮肤僵硬, 触摸时好似一种塑料纸。上、下肢和手指呈半屈状, 足趾屈曲分开。手和足背水肿。于出生后 24 小时内或立刻可见到火棉胶样膜有皸裂, 皸裂累及膜全层, 其边缘翘起, 不透明。膜下为表皮深层, 潮湿, 高低不平, 呈红斑样。薄片状脱屑由皸裂部位开始, 于 15~30 天内脱屑累及全身, 最后脱屑的部位是头颅和肢端。膜下皮肤呈红斑样, 伴糠秕状鳞屑, 以后发展演变为各型鱼鳞病。

所有已报道的突变均是由于单个碱基插入, 使其后的氨基酸发生改变, 终止密码子延迟出现, 引起蛋白质功能的改变。在被诊断为 Vohwinkel 综合征的患者上鉴定出 Loricrin 基因第一个突变。然而, 其临床特征不同于最初的 Vohwinkel 在 1929 年报道的 Vohwinkel 综合征, 目前已知是连接蛋白 26(connexin 26)导致的 Vohwinkel 综合征。Loricrin 突变的患者和最初的 Vohwinkel 综合征患者均表现为掌趾部弥漫性角化过度, 表面呈蜂窝状小凹, 趾(指)发生纤维性收缩窄带, 进行性假趾(指)断症改变(pseudoainhum), 但在一些临床方面有所不同。最初 Vohwinkel 综合征患者的四肢关节上表现为特殊的线状或海星状的角化过度, 并伴有高频性听力障碍。但是 Loricrin 突变的患者没有海星状角化病或听力障碍, 非掌趾部位而表现为不同的皮肤病损, 可以描述为全身性的非红皮病样的鱼鳞病, 或诊断为鱼鳞病变异型 Vohwinkel 综合征(ichthyotic variant of Vohwinkel's syndrome, OMIM 604117)。在其它一些病例中, 表现为全身性的界限清楚的红斑性过度角化斑块, 在临床上被诊断为红皮肤角化病(erythrokeratoderma; OMIM 602036)。还有其它病例出生时为火棉胶婴儿, 后来发展成鱼鳞病样红皮病, 认为是先天性非大疱性鱼鳞病样红皮病的一种形式。为避免混淆真正的因连接蛋白导致的 Vohwinkel 综合征, Ishida-Yamamoto 等人将其命名为兜甲蛋白角化病(Loricrin keratoderma, LK)。在所有文献的兜甲蛋白角化病患者中, 没有明显的基因型与表型相对应的表现。在患者中皮肤损害程度存在差异, 甚至在同一家系中。尽管在 Loricrin 突变的患者中存在一些临床上的差异, 但几乎所有显微镜

下的超微结构和免疫组织学特征相同, 提供了重要的诊断价值。

3. 丝聚合蛋白与兜甲蛋白

丝聚合蛋白与兜甲蛋白均是表皮分化复合物基因的成员, 都参与了细胞套膜形成及角化过程。表皮分化复合物基因(epidermal differentiation complex, EDC; OMIM:601588)是一呈簇状分布于染色体 1q21 上约 2.5Mb 范围内的基因家族, 它们与表皮的终末分化密切相关。这些蛋白表达在角质形成细胞的各层。它们的基因突变, mRNA 表达量的改变, 或是蛋白结构功能的变化, 和多种角化异常的皮肤病及皮肤肿瘤有关。EDC 由 3 个基因家族组成: 第 1 个家族编码细胞角质化包膜的前体蛋白, 组成编码人类表皮的结构蛋白, 包括: 外皮蛋白(involucrin, IVL)、兜甲蛋白(Loricrin, LOR)、富含脯氨酸小蛋白(small proline-rich proteins, SPRR)等。SPRR 又可以分为 4 个亚群: 2 个 SPRR1, 7 个 SPRR2, 1 个 SPRR3 和 1 个 SPRR4。该家族的特征是基因大小相对比较小, 且中心区域包含短串联重复, 外皮蛋白为 39 个重复, LOR 为 20 个, SPRR 为 3~14 个; 第 2 个家族编码位于表皮颗粒层的中间丝相关蛋白, 在表皮角化过程中与角质形成细胞的角蛋白丝相结合, 包括: 丝裂蛋白原, 毛透明蛋白(trichohyalin, THH)。其特征是序列不仅包含多串联重复, 而且在氨基端有 EF 手结构域(即 Ca^{2+} 结合区域), 故该家族又称为“融合基因”。第 3 个家族为钙结合蛋白。它的特征是具有两个 EF 手结构域(即 Ca^{2+} 结合区域)。

丝聚合蛋白(Filaggrin): 在表皮颗粒层的起始部位, 角质形成细胞合成一种 450~600kD 的重磷酸化蛋白即前丝聚合蛋白(Profilaggrin), 这种蛋白参与透明角质颗粒的形成, 前丝聚合蛋白是一种不溶性的中性蛋白, 其氨基酸组成以丝氨酸, 甘氨酸, 谷氨酰胺, 组氨酸及精氨酸为主, 当进入到颗粒层和角质层之间的过渡带时, 前丝聚合蛋白经过去磷酸化和蛋白质水解形成丝聚蛋白(一个前丝聚合蛋白分子含 12 个丝聚蛋白)。丝聚蛋白是一种 36kD 的蛋白质, 其编码基因位于人类染色体 1q21 区, 丝聚蛋白作为一种基质, 具有“胶水”样作用, 能促进角质白中间丝(KIF)的致密包裹, 形成巨纤维。

兜甲蛋白(Loricrin): 人的 LOR cDNA 长 1.3 kb, 包含 36 bp 的 5' 非编码区, 945 bp 的开放阅读框和 237bp 的 3' 非编码区及 38 bp 的 polyA 尾。兜甲蛋白是一种 26kD 的不溶性蛋白质, 由 315 个氨基酸组成。最初聚集于透明角质颗粒层中, 这种蛋白富含甘氨酸, 丝氨酸, 半胱氨酸。序列分析表明, 甘氨酸和丝氨酸残基构成了“甘氨酸 n—丝氨酸 2—甘氨酸 n—丝氨酸 2”的结构形式, 这种结构再由半胱氨酸残基连接起来, 形成几个富含甘氨酸的长片段特殊序列, 其间由富含谷氨酰胺和谷氨酰胺/赖氨酸的短片段分隔连接。推测这种富含甘氨酸的序列具有高度柔韧性和延展性, 并能形成“甘氨酸环”与角蛋白多肽的“头”功能区

相互作用，为表皮提供柔韧性和延展性。兜甲蛋白通过表皮转谷氨酰酶参与细胞套膜的形成，通过异二肽链直接与细胞套膜交联。

2004 年，本研究收集到一个已遗传了四代的常染色体显性遗传的掌跖角化和鱼鳞病样家系，经我们初步诊断为寻常性鱼鳞病。采集了其中三代共 22 个家系成员的血样，其中患者 7 人。该家系是一诊断准确的 4 代家系，又具有比较明确的遗传方式，因此适于运用基因组扫描和连锁分析法进行基因的定位研究。本研究决定首先对该疾病表型相关的已定位区间进行连锁分析，若得不到与已定位区间连锁的证据，则进行全基因组扫描及连锁分析，对该疾病基因进行染色体定位，再在定位区间采用候选基因克隆的方法，选择候选基因进行突变检测，以期克隆该疾病基因，并为从分子水平进一步研究该病的发病机制，及更好地预防、诊断、治疗此类疾病奠定良好的基础。

3.1. 材料与方法

3.1.1 研究对象

一个来自河北省的掌跖角化和鱼鳞病样家系（图 3-1），该家系共有 4 代 27 个家系成员。经皮肤检查，并对先证者（III-4）进行了病理组织切片检查，家系中有患者 7 人，在家系中表现连续 3 代分布，其中男性患者 3 人，女性患者 4 人，无明显性别差异，符合常染色体显性遗传特征。

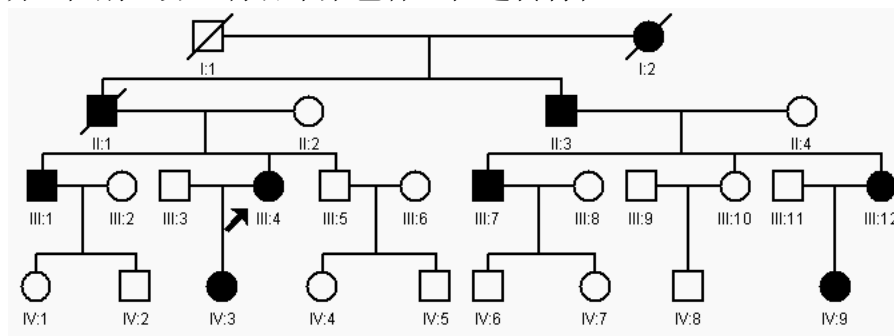


图 3-1 中国河北鱼鳞病样家系，箭头所示为先证者

3.2 研究方法

3.2.1 DNA 抽提

在患者知情同意的前提下，采集了 22 个家系成员外周血样各 5ml（除 III-9、III-11），常规苯酚、氯仿抽提基因组 DNA。

3.2.2 对已定位区间进行排除性定位分析

首先选取包括 Keratin 2a、KRT1、KRT10、LI1、LI2、LI3、LI5、IV 等已定位区间或已克隆的遗传性鱼鳞病相关基因位点附近微卫星 DNA 标记（ABI PRISM™ Linkage Mapping Set Version 2）：D1S2726、D1S252、D1S498、D1S484、D1S2878、D2S2382、D2S325、D3S1277、D3S1289、D12S368、D12S83、D14S283、D14S275、D17S938、D17S1852、D17S798、D17S1868、D19S226、D19S414 等 19 个，优化组合为 5 组，每组容纳 3~4 对引物，标记相同荧光染料的位点片段范围至少相差 15bp 以上，进行多重 PCR 扩增，之后对该疾病家系进行连锁分析。

3.2.2.1 对微卫星位点引物进行多重PCR

采用 Touchdown 程序结合多重 PCR 反应方法，将 3~4 种不同引物混合，放入同一管中进行扩增，总反应体积 5μl（见表 3-1）。同时取法国人 CEPH-02（人类多态研究中心提供的欧洲人标准样品）作为阳性对照。PCR 反应用 96 孔微量板，在 Perkin Elmer 9600 热循环仪上进行。反应条件采用两相循环（见图 1-4）：94℃ 12min，预变性，第一相循环时，94℃ 变性 30s，起始退火温度 63℃ 60s，以后每

个循环下降 0.5°C ， 72°C 延伸 $1\text{min}50\text{s}$ ，共15个循环。目的是使每对引物扩增时都达到最佳退火温度，都能扩增出产物；第二相循环， 94°C 变性 30s ， 56°C 退火 60s ， 72°C 延伸 $1\text{min}50\text{s}$ ，24个循环，最后 72°C 延伸 15min 。对于扩增不好的位点，再进行单重PCR扩增。

表 3-1 多重 PCR 反应体系

Reagent (试剂)	Volume (体积 μl)
10 \times GeneAmp PCR buffer II	0.5 μl
10mM dNTPs	0.1 μl
25mmol/L MgCl_2	0.6 μl
5u/ μl Ampli Taq Gold TM polymerase	0.04 μl
10mmol/L primer	0.04 μl /pair
30ng/ μl Genomic DNA	1 μl
sterile deionized H_2O	add ddH ₂ O to 5 μl

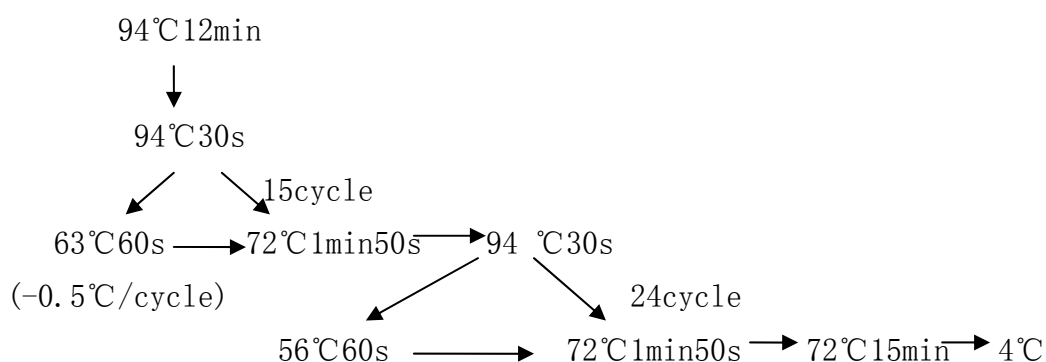


图 3-2 PCR 循环程序

3.2.3 上样和电泳

PCR 扩增产物的电泳和分析使用 ABI3100 DNA Genetic Analyzer 完成，该仪器采用毛细管电泳。在 96 孔板的每一孔中分别加入 $10\mu\text{l}$ 的 HI-DI 缓冲液， $1.0\mu\text{l}$ PCR 扩增产物， $0.5\mu\text{l}$ GS-400ROX 分子量内标。轻微振荡混匀，短暂离心。将 96 孔板在 PCR 仪上 94°C 加热变性 3min 后，放入冰中骤冷，以保持 DNA 的单链状态。将 96 孔板置于 ABI3100 DNA Genetic Analyzer 的进样器上，设定电泳程序，自动进样后开始电泳，恒定电压 3000V 、恒定温度 60°C ，电泳 2.5h 。

3.2.4. 数据处理

3.2.4.1 收集数据

用 ABI PrismTM 3100 analyzer Collection software version 1.1 进行数据收集。

3.2.4.2 转换数据

应用 GeneScan Analysis 3.7 软件进行分子量内标校正, 确定片段大小的精确性。并自动将 ABI3100 收集的凝胶电泳图象转换为数字信号文件, 然后进行分子量内标校正和扩增片段大小分析。

3.2.4.3 基因分型

用 Genotyper 2.1 软件进行样本基因分型, 具体步骤: (1) 准确鉴定每一泳道内的分子量内标, 然后依据内标绘出分子量的回归曲线; (2) 校峰, 以进一步确定每个峰的真实性; (3) 对各产物进行检测, 根据回归曲线确定每一产物的分子量大小, 并由此转化为片段长度大小。

3.2.4.4 孟德尔遗传模式检验

使用 Designer 软件及 Linkage 软件检测基因型数据的错误并对产物进行孟德尔遗传模式的校对与检查。Designer 软件是利用每个个体的基因型信息来检测父母与小孩的一致性, Linkage 软件是利用每个个体的基因型信息来检测同胞中是否有多余的等位基因。这 2 个软件检查的目的是寻找基因型数据是否有判读错误, 从而检查出家系中是否存在不符合孟德尔遗传的情况。之后, 对错误读取数据进行人工修正。

3.2.4.5 连锁分析

使用 Linkage version5.1 软件包进行连锁分析。在常染色体显性遗传模式下假设男性和女性重组率相同、设疾病基因的外显率为 0.95、疾病等位基因的频率为 0.0001、标记位点的各等位基因频率设为相等, 然后使用 Linkage version5.1 软件包中的 MLink 程序进行两点连锁分析, 计算两点 LOD 值。

3.2.5 精细定位

通过排除性定位扫描分析结果, 在 D1S252 处得到 LOD 值 2.17, D1S498 处得到 LOD 值 1.70, 在 D1S484 处得到 LOD 值 1.08。其它位点 LOD 值 < -2, 说明该家系致病基因与 1 号染色体连锁, 排除了该家系致病基因与其他已知位点的连锁关系, 进一步在 Marshfield 数据库中 D1S252、D1S498 和 D1S484 附近选取 10 个微卫星标记, 平均相距 2cM 左右, 重复 3.2.2.1 以下步骤, 分组先进行多重 PCR 扩增, 再进行基因组扫描, 然后计算 LOD 值进行精细定位研究。

3.2.6 基因组扫描和连锁分析

用 1 号染色体上微卫星荧光标记引物 (ABI PRISM™ Linkage Mapping Set Version 2) 进行基因组扫描。在 5ul 反应体系中进行多重 PCR, 包含 30ng gDNA, 10X PCR 缓冲液 0.5ul, 10mM dNTPs 0.1ul, 25mM MgCl₂ 0.6ul, 每条引物 80pmol 和 0.2U GOLD Taq 酶。

PCR 扩增产物的电泳和分析使用 ABI3100 DNA Genetic Analyzer 完成, 该仪器采用毛细管电泳。在 96 孔板的每一孔中分别加入 10μl 的 HI-DI 缓冲液, 1.0μl

PCR 扩增产物, 0.5 μ l GS-400ROX 分子量内标。轻微振荡混匀, 短暂离心。将 96 孔板在 PCR 仪上 94 $^{\circ}$ C 加热变性 3min 后, 放入冰中骤冷, 以保持 DNA 的单链状态。将 96 孔板置于 ABI3100 DNA Genetic Analyzer 的进样器上, 设定电泳程序, 自动进样后开始电泳, 恒定电压 3000V、恒定温度 60 $^{\circ}$ C, 电泳 2.5h。

应用 GeneScan Analysis 3.7 软件进行分子量内标校正, 确定片段大小的精确性。并自动将 ABI3100 收集的凝胶电泳图象转换为数字信号文件, 然后进行分子量内标校正和扩增片段大小分析。

用 Genotyper 2.1 软件进行样本基因分型, 具体步骤: (1) 准确鉴定每一泳道内的分子量内标, 然后依据内标绘出分子量的回归曲线; (2) 校峰, 以进一步确定每个峰的真实性; (3) 对各产物进行检测, 根据回归曲线确定每一产物的分子量大小, 并由此转化为片段长度大小。

利用 Linkage5.1 软件计算两点连锁值。在常染色体显性遗传模式下假设男性和女性重组率相同、设疾病基因的外显率为 0.95、疾病等位基因的频率为 0.0001、标记位点的各等位基因频率设为相等, 然后使用 Linkage version5.1 软件包中的 MLink 程序进行两点连锁分析, 计算两点 LOD 值。

为了确定该家系疾病基因的最小候选区间, 根据微卫星标记的基因分型结果, 利用 Cyrillic 软件构建单体型后, 再进行人工校正。

3.2.7. 构建单体型图

为了确定该家系疾病基因的最小候选区间, 根据微卫星标记的基因分型结果, 利用 Cyrillic 软件构建单体型后, 再按照家系中交换数最少的原则对单体型进行人工修正。

3.2.8 多点连锁分析

使用 genhunter 软件包进行多点连锁分析。在常染色体显性遗传模式下设定男性和女性重组率相同、设疾病基因的外显率为 0.95、疾病等位基因的频率为 0.0001、标记位点的各等位基因频率设为相等, 进行多点连锁分析, 计算多点 LOD 值。

3.2.9 候选基因突变分析

通过 UCSC(www.genome.ucsc.edu) 检索, 在本研究的定位区间内共有 KownGene 455 个, refseq Gene 200 个, 在此区间选择合适的功能候选基因进行突变检测, 将有望克隆到鱼鳞病致病基因。此区间包括表皮分化复合物基因(EDC), EDC 与表皮的终末分化密切相关, 包括钙结合蛋白家族、小分子脯氨酸丰富蛋白(small proline rich proteins)家族、LCE (late cornified envelope) 家族、套膜蛋白、兜甲蛋白、角蛋白细丝等基因, 它们的功能都与细胞套膜形成、表皮分化有关。选择表皮分化复合物基因的成员丝聚合蛋白(Filaggrin)与兜甲蛋白

(Loricrin)为候选基因，因已报道 FLG 突变导致寻常性鱼鳞病，LOR 突变导致鱼鳞病型 Vohwinkel 综合征。对候选基因 FLG，LOR 进行突变分析。

3. 2. 10 引物的设计

使用加拿大生物研究所提供的 primer3 在线软件(<http://www-genome.wi.mit.edu/cgi-bin/primer/primer3.cgi/>)设计 PCR 引物。根据 *FLG* 基因的序列 (NM_002016)，设计了 *FLG* 基因 2 号外显子、3 号外显子非重复序列的引物(表 1)；同时参照 Smith、Sandilands 等^[3,9]订购了 *FLG* 基因部分重复序列的引物。根据 LOR 基因的序列(NM_000427)设计了 LOR 基因外显子的引物。引物由上海博亚生物技术有限公司合成。为了使 PCR 扩增有效性、特异性更好，我们对以下参数进行了设定：引物中的 GC 含量设定在 40%-60%之间；退火温度设定在 57℃-63℃之间；为了使引物更好的结合到模板链上，引物的 3'端一般设定为以 G 或 C 结尾；为了便于测序，扩增片段一般设定在 200bp-500bp 之间。

表 3-2 FLG 基因突变检测引物一览表

外显子 编号	引物 名称	正向引物序列	反向引物序列
Exon2	flg-1	5'ttctcatgtccattctccattc3'	5'aagagctcaaaataacccttgc3'
Exon3	flg-2	5'ggctgataatgtgattctgtctg3'	5'tgtttctcttgggctcttgg3'
	flg-3	5'gaccctcaagtctggaaagaag3'	5'atcagtgggtgcataggcttca3'
	flg-4	5'ggaagaggctaagtgaagacttg3'	5'ctgcagagtgcccatgact3'
	flg-5	5'caggatgaagcctatgacacca3'	5'gtttgtctgcttgacttctgg3'
	flg-6	5'gcactggaggaagacaaggatc3'	5'gaactagaactgtgaggactgccac3'
	flg-7	5'tgaggaacagtcaggagacg3'	5'accttttgcctttcagtgc3'
	flg-8	5'tcacaagattctgcgtatcactc3'	5'gccacataaacctgggtcct3'
	flg-9	5'ggtgaatccgggttagaca3'	5'cccagaatctcctaaaataactcca3'

表 3-3 LOR 基因突变检测引物一览表

Genes	Exons	引物名称	引物序列	Tm	bp
Loricrin	exon1	Lor-1f	GCTGCCTCCGTAAGTATCCTC	57.1	
		Lor-1r	TGGAGAAACAGCCAGAGCC	57.9	399
		Lor-2f	GTTCTGGCTGCGGCTACT	59.1	
		Lor-2r	GACCTGCTGCGAGGAGAC	59.6	500
		Lor-3f	GCTCTGTCTGCGGCTACTCT	59.9	
		Lor-3r	AAACACCTCCAACCTCTTCG	59.2	431

3. 2. 11 PCR 扩增条件

10 μ l 反应体系（见表 3-4）；反应条件（见图 3-3）：首先选取先证者 M4453（III-4）的 DNA 样本进行 PCR 扩增。

表 3-4 热启动 PCR 试剂组成

试剂 (Reagents)	体积 (Volume)
10 \times Buffer	1 μ l
Q-Buffer	2 μ l
MgCl ₂	0.6 μ l
dNTPs (10 mM)	0.2 μ l
Hot Star Taq 酶 (5 U/ μ l)	0.05 μ l
Primers (100ng/ μ l)	0.3 μ l
DNA 模板 (50 ng/ μ l)	1 μ l
ddH ₂ O Add to	10.0 μ l

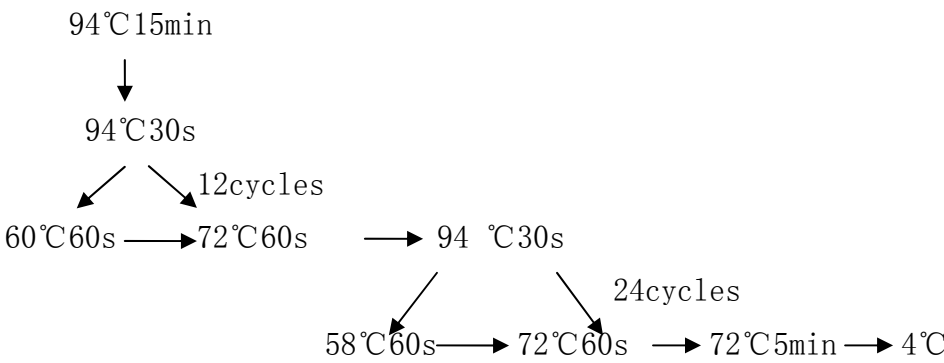


图 3-3 PCR 反应条件

3. 2. 12 聚丙烯酰胺凝胶电泳

取 2 μ l PCR 反应产物，加入 0.2 μ l 上样缓冲液，混匀，上样于 6%聚丙烯酰胺凝胶上(6%聚丙烯酰胺凝胶：30%丙烯酰胺 6ml，10 \times TBE 1.5ml，TEMED 25 μ l，10%APS 300 μ l，加去离子水至30ml)，300V 电泳 40 分钟，硝酸银染色，检测 PCR 产物。

3. 2. 13 测序

若 PCR 产物特异性好， 则可将 PCR 产物用虾碱性磷酸酶和核酸外切酶处理（冰上操作）消化残余的寡核苷酸及单链 DNA 片段，反应体系见表 3-5。37 $^{\circ}$ C 孵育 60min， 80 $^{\circ}$ C 15min 灭活虾碱性磷酸酶和核酸外切酶，4 $^{\circ}$ C 保存。酶切后的 PCR 产物无需进一步纯化，可直接作为模板进行测序。测序采用合成的专用测序引物，

以及 BigDNA[®] Terminator v3.1 Cycle Sequencing Kit 测序试剂盒, 测序反应体系参照该试剂盒说明书。测序反应完成后, 取反应产物 5 μ l, 于 ABI3100 DNA Genetic Analyzer 完成测序。

表 3-5 虾碱性磷酸酶和核酸外切酶处理的试剂组成

试剂 (Reagents)	体积 (Volume)
10 \times SAP Buffer	0.5 μ l
SAP (1U/ μ l)	2 μ l
Exonuclease I (20U/ μ l)	0.25 μ l
PCR 产物	7.25 μ l

3.2.14 序列分析

测序结果用 DNASTAR 软件中的 EditSeq、SeqMan 程序分析, 以 NCBI 数据库 (<http://www.ncbi.nlm.nih.gov>) 中 FLG、LOR 基因标准序列 (NM_002016 及 NM_000427) 为参照, 进行序列比对分析。

3.2. 结果

3.3.1 临床表现

先证者 (III-4) 临床特征表现为皮肤干燥, 四肢伸侧及背部皮肤具灰白色至淡棕色菱形或多角形鳞屑, 鳞屑中央固着, 边缘翘起, 如鱼鳞状, 面部皮肤正常; 掌跖可见线状皲裂和掌纹加深 (图 3-4A, B), 病理组织切片表现为表皮中等角化过度, 颗粒层稍变薄、真皮浅层血管扩张, 轻度淋巴细胞、组织细胞浸润 (图 3-4C)。

3.3.2 排除性定位分析

通过对与表型相关的已定位区间及已克隆基因附近微卫星 19 个 DNA 标记 (ABI PRISM[™] Linkage Mapping Set Version 2): D1S2726、D1S252、D1S498、D1S484、D1S2878、D2S2382、D2S325、D3S1277、D3S1289、D12S368、D12S83、D14S283、D14S275、D17S938、D17S1852、D17S798、D17S1868、D19S226、D19S414 进行连锁分析, 结果在 D1S252 处得到 LOD 值 2.17, D1S498 处得到 LOD 值 1.70, 在 D1S484 处得到 LOD 值 1.08, 将疾病基因初步定位在 1 号染色体上。其它染色体微卫星 DNA 标记的 LOD 值 < -2, 排除了该家系致病基因与其它已知位点的连锁 (见表 3-6)。



图 3-4 A: 先证者下肢伸侧及背部皮肤具灰白色至淡棕色菱形或多角形鳞屑



图 3-4 B: 先证者的双掌，蜂窝状掌跖角化，可见明显线状皲裂和掌纹加深；

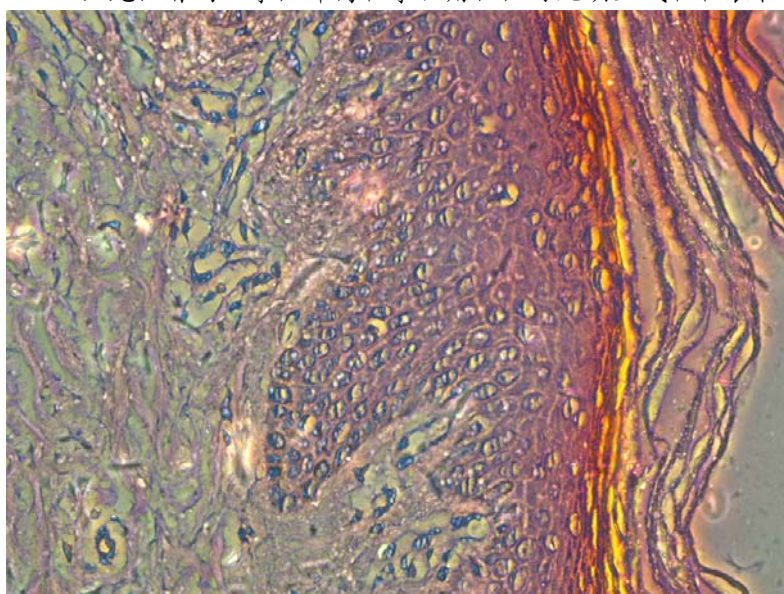


图 3-4 C: 患者皮损病理切片（1000X）

表 3-6 两点连锁分析 LOD 值（排除分析）

微卫星 DNA	不同重组率（ θ ）的 LOD 值						
标记位点	0.0	0.1	0.2	0.3	0.4	0.5	
D1S2726	-4.67	-0.47	-0.26	-0.16	-0.08	0.00	a
D1S252	2.17	1.97	1.56	1.06	0.52	0.00	a
D1S498	1.70	1.35	0.99	0.60	0.21	0.00	a
D1S484	1.08	0.98	0.77	0.51	0.24	0.00	a
D1S2878	-3.91	-0.18	0.08	0.12	0.08	0.00	a
D2S2382	-9.93	-0.29	0.07	0.15	0.11	0.00	a
D2S325	-11.12	-1.99	-0.99	-0.46	-0.15	0.00	a
D3S1277	-4.94	-1.37	-0.61	-0.24	-0.06	0.00	a
D3S1289	-10.18	-1.31	-0.57	-0.26	-0.11	0.00	a
D12S368	-3.03	-0.81	-0.46	-0.25	-0.1	0.00	a
D12S83	-3.75	-0.35	-0.16	-0.07	-0.02	0.00	a
D14S283	-3.75	-0.49	-0.18	-0.06	-0.01	0.00	a
D14S275	-5.37	-1.0	-0.46	-0.21	-0.07	0.00	a
D17S938	-10.45	-1.66	-0.65	-0.20	-0.01	0.00	a
D17S1852	-9.94	-0.81	-0.25	-0.04	0.03	0.00	a
D17S798	-3.26	0.58	0.57	0.42	0.22	0.00	a
D17S1868	-3.71	-0.12	0.12	0.17	0.13	0.00	a
D19S226	-9.52	-1.47	-0.61	-0.23	-0.05	0.00	a
D19S414	-3.59	-0.59	-0.20	-0.05	-0.01	0.00	a

3.3.3 精细定位

在 Marshfield 数据库中 D1S252、D1S498 和 D1S484 的附近选取 D1S187、D1S2746、D1S2881、D1S189、D1S534、D1S2696、D1S2346、D1S305、D1S3466、D1S1153 等 10 个微卫星标记，平均相距 2cM 左右，重复连锁分析步骤，然后计算 LOD 值进行精细定位研究。

进一步利用 Marshfield 数据库选择的遗传标记进行精细定位，结果在 D1S2696 得到最大两点 LOD 值为 3.46($\theta=0$)，在 D1S187 处 LOD=3.11($\theta=0$)，证明该家系致病基因肯定与此区域连锁。同时附近的遗传标记也得到 2 左右的 LOD 值，也给该家系致病基因与此区域连锁提供了支持证据。遗传标记和疾病位点之

间的两点 LOD 值见表 2。

表 3-7 疾病位点和 1 号染色体上 12 个标记之间的两点 LOD 值

Locus	LOD Score at $\theta=$					
	0.0	0.1	0.2	0.3	0.4	0.5
D1S2726	-4.67	0.47	-0.26	-0.16	-0.08	0.0
D1S187	3.11	2.55	1.93	1.26	0.56	0.0
D1S2746	1.93	1.51	1.07	0.6	0.19	0.0
D1S2881	1.97	1.59	1.18	0.75	0.35	0.0
D1S189	2.29	1.86	1.39	0.91	0.43	0.0
D1S252	2.17	1.97	1.56	1.06	0.52	0.0
D1S534	2.54	2.07	1.55	0.99	0.42	0.0
D1S2696	3.46	2.85	2.18	1.45	0.67	0.0
D1S2346	2.17	1.75	1.30	0.82	0.35	0.0
D1S305	-1.35	1.37	1.10	0.69	0.29	0.0
D1S3466	2.84	2.32	1.76	1.16	0.53	0.0
D1S1153	-1.03	1.69	1.40	0.95	0.45	0.0

3.3.4 单体型图的构建

根据家系中交换数最少的原则对单体型进行人工修正后,单体型分析被用来确定共分离区的边界(图 3-5)。结果显示,发现所有患者均携带同一单体型;III-1、III-4、IV-3 在 D1S2346 和 D1S305 之间发生了重组,提示疾病基因位点位于 1 号染色体长臂 D1S305 的着丝粒端及 1 号染色体短臂;IV-9 在 D1S2726 和 D1S187 之间发生了重组,进一步提示疾病基因位点位于 1 号染色体短臂 D1S2726 的着丝粒端。这些结果尤其是这些重组事件的发生则将该鱼鳞病家系致病基因定位于 D1S2726 和 D1S305 之间 15cM 区域。单体型分析结果见图 3,加框的基因型或连锁群示与疾病共分离的单体型。微卫星标记从上到下次序: 1pter-D1S2726-D1S187-D1S2746-D1S2881-D1S189-D1S252- D1S534-D1S2696-着丝粒-D1S2346-D1S305-D1S3466-D1S1153-1qter。

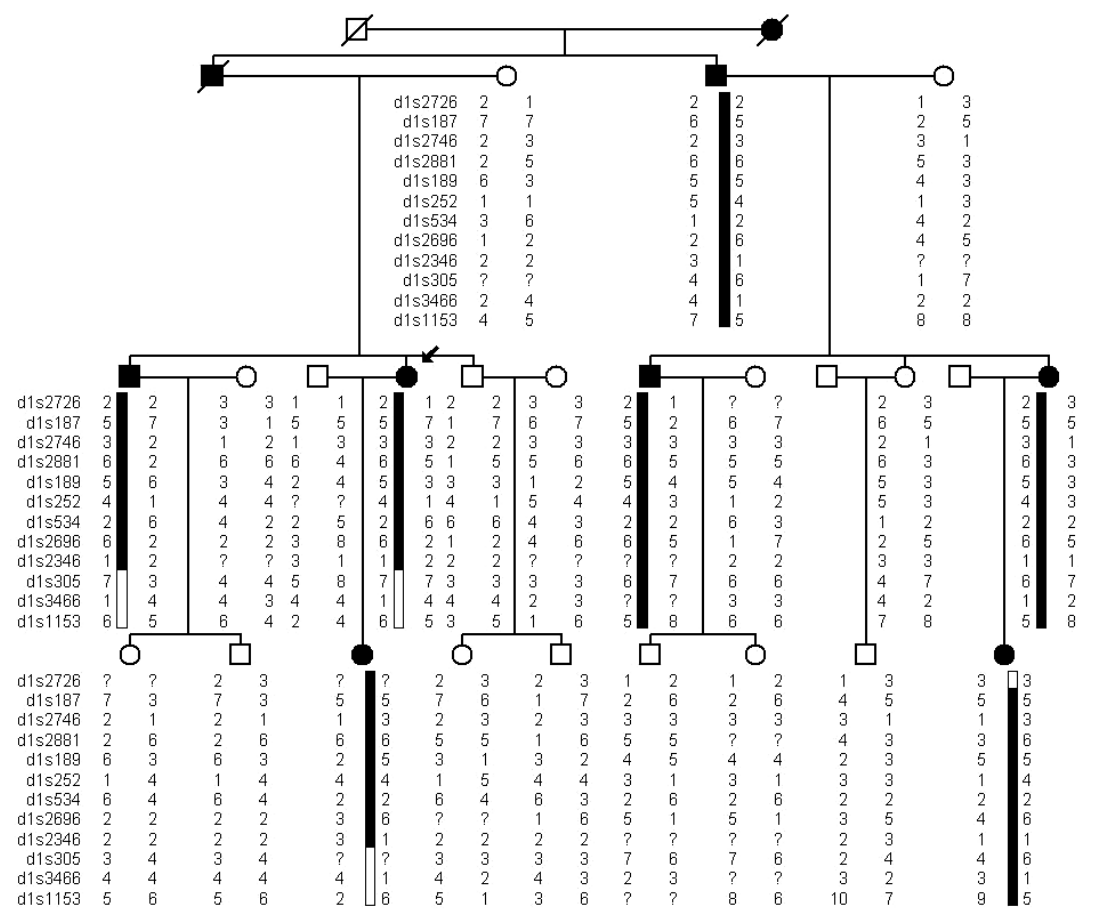


图 3-5 鱼鳞病样家系 1 号染色体上 12 个微卫星标记的单体型分析

3. 3. 5 突变检测结果

扩增产物经 6%聚丙烯酰胺凝胶电泳检测，所有引物均扩增出与预期大小基本符合的片段。*FLG* 测序结果分析，先证者（III-4）外显子 2 和外显子 3 的非重复序列及部分重复序列，均未发现与疾病相关的突变，也未见任何多态。而在 LOR 基因上检测到一个单个核苷酸(G)的插入，730insG(见图 3-6)。然而，在 LOR 基因的核苷酸 730(730insG)的 6 个正常 G 碱基(密码子 230-231)后一个 G 碱基的插入，导致在密码子 232 的移码突变，其改变了其后 84 个氨基酸和引起终止密码子延迟，突变体蛋白质延长了 22 个氨基酸。所有受累者通过直接测序来检测，发现都有这个杂合子突变，而在家系正常人及正常对照人群中未见些突变。通过从受累者的亚克隆 PCR 产物测序其正向克隆证实为两个不同的等位基因，我们确认这个结果(图 3-6)。

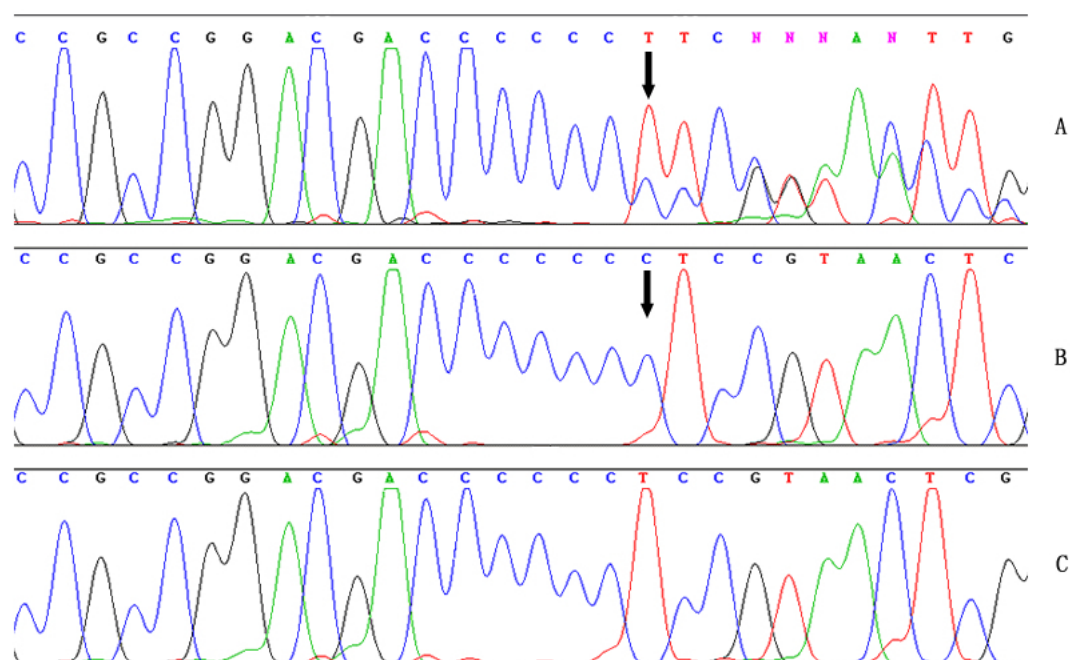


图 3-6. Loricrin 基因单个 G 碱基插入 (A) Sequence of the Loricrin gene around the G insertion in an affected individual. (B) Cloned affected and (C) cloned unaffected allele of the patient.

3.3. 讨 论

在本研究中，我们收集到的一个组织学特征高度相似的鱼鳞病样家系。通过连锁分析及单体型分析，疾病位点定位于染色体 1 p13.3-q21.3。在对 FLG 的非 filaggrin 重复序列编码区和启动子区的突变检测未发现致病性突变。但是我们在 LOR 基因上检测到一个已报道过的单个核苷酸(G)的插入(730insG)突变。LOR 羟基末端连续 6 个 G 碱基后单个 G 碱基插入 (730insG) 导致其后的甘氨酸环及高谷氨酰胺-赖氨酸区翻译编码框移位，改变了其后的 84 个氨基酸密码子的编码，并使终止密码子延迟出现，导致突变型蛋白较野生型延长 22 个氨基酸。最初我们将家系诊断为寻常性鱼鳞病，但家系的致病基因并不是已克隆的寻常性鱼鳞病致病基因 FLG，而是 LOR，经过查阅文献，LOR 基因单个碱基插入突变是导致 vohwinkel 综合症变异型(伴随鱼鳞病的 vohwinkel 综合症，残毁性掌跖角皮症)。并且，此河北家系中在 II3 和 III7 两个患者的小指上可见轻度的环状缩窄，其他患者都表现为弥漫性鱼鳞病皮肤病。我们怀疑最初的诊断有误，经过第二次家系调查，证实 II3 和 III7 两个患者的小指上可见轻度的环状缩窄，并且 II3 的第 5 趾多年前已自行离断。根据第二次检查到的临床特征和 LOR 突变结果，我们确定诊断此家系为兜甲蛋白角化病中的 vohwinkel 综合症变异型。

通过查阅已发表的有关 LOR 基因突变的文献，现在因 LOR 基因的一个核苷酸插入的突变导致的疾病现在共同命名为兜甲蛋白角化病(Loricrin keratoderma, LK)^[19]。LOR 基因突变导致的兜甲蛋白角化病表型具有多样性，在临床上被诊

断为 vohwinkel 综合症变异型(伴随鱼鳞病的 vohwinkel 综合症)、进行性对称性红皮肤角化病 (progressive symmetric erythrokeratoderma, PESK) 及出生时为火棉胶婴儿的先天性鱼鳞病样红皮病^[13]。其中共有 6 个家系中都是由于此 730insG 突变导致的(3 个英国家系,2 个日本家系,一个德国家系),见表^[10-15]。已报道 Loricrin 基因的杂合突变导致伴随鱼鳞病的 Vohwinkel 综合征^[10-16]和进行性对称性红皮肤角化病(erythrokeratoderma) ^[17]。Richard 认为日本的对称性红皮肤角化病家系属于 Vohwinkel 综合征的一种,因为出现了断肢性掌跖角皮病(假阿洪病,自发性断趾病),而断肢性掌跖角皮病是不出现了红皮肤角化病中的^[18]。LK 共同的临床表现包括蜂窝状的断肢性掌跖角化过度, 和形成指状缢痕带(假指/趾断症,假阿洪病)。其它临床表现包括在鱼鳞病型 Vohwinkel 综合征表现为中度鱼鳞病,和在红皮肤角化病中表现为广泛性红斑的角化病^[20]。Gedlicke 提出两个 LK 的必要特征: 蜂窝状的掌跖角化病和弥漫性鱼鳞病样皮肤病^[15]。组织学特征包括角化不全的角化过度症, 伴随颗粒层增厚和 Loricrin 突变体在细胞核中积聚。

表 3-8 Loricrin 基因单个碱基插入突变导致的兜甲蛋白角化病的临床表现

种族	基 因 突 变类型	蜂 窝 状 掌 趾角化病	假 指 / 趾 断 症, 和 / 或 自行断肢	其它皮肤异常	参考文献
1 英国	730insG	+	重度	弥漫性全身性鱼鳞病	[16]
2 苏格兰	730insG	+	重度-中度	出生时身上有白色物质, 膝和肘过 度角化, 轻度至中度鱼鳞病	[11]
3 日本	730insG	+	重度	鱼鳞病	[12]
4 日本	730insG	+	重度	火棉胶婴儿, 全身性鱼鳞病样红斑	[13]
5 英国	730insG	+	重度		[14]
6 德国	730insG	+	无	轻度的全身性先天性鱼鳞病和弥 漫性掌跖角化病	[15]
7 日本	709insC	+	重度-中度	全身性界限清楚的红斑性过度 角化斑块	[17]
8 北爱尔兰	662insG	+	中度	指关节角化过度, 全身性非红斑性鱼鳞 病	[16]
9 中国	730insG	+	中度	鱼鳞病	本次研究

在这个中国家系中, 患者都表现为一个常染色体显性蜂窝状掌跖角化病和弥漫性鱼鳞病样皮肤病, 在 II3 和 III7 两个患者的小指上可见轻度的环状缩窄, II3 有趾断症(pseudoainhums), 自发性离断(autoamputations) ^[19]。病理切片显示过度角化, 颗粒层轻度缩小。因此, 我们认为是一个存在不同的临床表现的常染色体显性遗传的鱼鳞病样和掌跖角化的疾病, LK 的两个特征是蜂窝状掌跖角化病和弥漫性鱼鳞病样皮肤病。

兜甲蛋白角化病都是由于单个碱基插入, 使其后的氨基酸发生改变, 终止密码子延迟出现, 引起蛋白质功能的改变。有学者认为, 角质化是一个凋亡过程。Loricrin 是富含甘氨酸的蛋白质, 序列包含许多编码甘氨酸的 GGX 密码子。

单个碱基插入使其发生阅读框的移码突变, 本身编码甘氨酸的 GGX 密码子改变阅读框成为编码精氨酸的 CGG 和 AGG^[19]。突变的 Loricrin 因此富含精氨酸, 其包含一个可能的二连的核定位信号区(nuclear localization signals, NLSs)^[11]。在皮肤角化病患者皮肤中, 角质层固有核凋亡标记 TUNEL 阳性。Suga 也证实转基因小鼠的突变的 Loricrin 的羧基端存在一个核定位信号区^[21]。推测聚集在核中的突变 LOR 可能通过干扰核的凋亡过程影响角质形成细胞(keratinocyte, KC)的终末分化^[19]。突变的 Loricrin 运输到细胞核并蓄积, 在角质化细胞终末分化时可能会延迟细胞凋亡的进程。突变 Loricrin 的活性很可能是 LK 表型的诱发因素^[15,20]。但也有学者认为, LOR 的插入突变使它通过胞外域结构使部分角质化包膜支架产生缺陷, 从而引起皮肤屏障功能的损害, 具体表现为脆性增加、经表皮水分丢失 (TEWL) 增多、角质层水合作用减少、pH 值降低等变化^[25]。

由相同的 Loricrin 基因 730insG 突变在中国鱼鳞病家系与其它 LK 家系中导致的临床表型不同。目前这种多样性的原因仍不清楚。可能与 Loricrin 表达强度有差异或 Loricrin 基因多态性有关^[19]。

已建立 Loricrin 突变的小鼠模型用于研究 Loricrin 表型和 LK 的皮肤病理学之间的关系。在鼠的 Loricrin 编码序列中插入一个核苷酸的 Loricrin 突变的转基因小鼠表现为持续的典型人类 LK 的表现(如广泛性鱼鳞, 在尾部有压缩带, 和足垫增厚)^[21]。然而, 人类 Loricrin 基因在转基因小鼠中过度表达, 未显示任何表型^[22]。Loricrin 基因敲除小鼠显示红斑的, 有光泽的, 半透明的皮肤表现, 这种表现在出生后的 4-5 天消失^[23]。尽管缺乏特征性的包含角质透明蛋白粒的 Loricrin, 角化细胞(corneocytes)的细胞膜表现上正常。皮肤表型的正常化是由于一些小的富含脯氨酸的蛋白质补偿性表达, 这些小的富含脯氨酸类似于 Loricrin, 也富含甘氨酸和丝氨酸^[24]。Loricrin 角化病皮肤表现为角质层(SC)的水合, 结合的功能异常, 和显著的通透性屏障缺陷。这种屏障异常与角质化包膜(cornified envelope, CE)支架缺陷有关, 其导致了细胞外的通透性增强。在 Loricrin 角化病的角质层外的角质化包膜支架局部正常, 这也许可解释其它角质化包膜蛋白质的补偿性的过度表达^[25]。这些结果揭示存在一种代偿机制, 诱导尝试恢复表皮屏障作用, 然后避免更严重的皮肤表型。

兜甲蛋白是分化后期的标志, 仅特异性地表达于基底层上(颗粒层)表皮终末分化型角质化细胞中。已证实有一些调控元件和转录因子可特异性地调控特殊基因表达的分化。钙离子和佛波酯上调兜甲蛋白的表达, 维 A 酸抑制其表达。在体外, 兜甲蛋白的转录需要一些蛋白因子结合到保守的 AP-1 位点上, 这种进化保守的 AP-1 位点位于接近兜甲蛋白启动子的区域, 该区域对兜甲蛋白启动子活性是必需的^[26]。位于小鼠 Loricrin 基因的第三个内含子有一个负调控序列, 其包

含一个假定的反向的 Yin-Yang1(YY1)结合域。Yin-Yang1 负调控分化-在未分化角质化细胞中,小鼠 Loricrin 基因的特殊转录本^[27]。在培养的人类角质化细胞中 Loricrin 表达是被一个复杂的相互作用的 Sp1, CREB, AP1 和 AP2 家族的转录因子所控制的^[28]。可能是由于在调控元件和转录因子上存在某个多态,导致 Loricrin 基因的差异表达使 LK 患者的临床表型的出现不同差异。

综上所述,我们报道一个因 Loricrin 基因的 730insG 突变的兜甲蛋白角化病家系。因 Loricrin 突变导致不同的临床表型的原因,需在今后的研究中探索清楚。

第二部分 结 论

- 1.通过连锁分析将兜甲蛋白角化病家系的致病基因定位于 D1S2726-D1S305 之间约 15cM 范围内。
2. LOR 基因的 730insG 突变是本兜甲蛋白角化病家系的致病突变。
3. 兜甲蛋白角化病的表型存在多样性。兜甲蛋白角化病的两个特征是蜂窝状掌跖角化病和弥漫性鱼鳞病样皮肤病。

参考文献

- [1] Wells RS, Kerr CB. (1966) Clinical features of autosomal dominant and sex-linked ichthyosis in an English population. *Br Med J* 1:947-950.
- [2] Lei G, Zhang Y, Hu Y. (1992) Investigation on the prevalence of Ichthyosis in Sichuan province *Chin J Dermatol* (in Chinese) 25:105-106.
- [3] Smith FJ, Irvine AD, Terron-Kwiatkowski A, et al. Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. *Nat Genet*, 2006, 38(3):337-342.
- [4] Anton-Lamprecht, I, Hofbauer, M (1972) Ultrastructural distinction of autosomal dominant ichthyosis vulgaris and X-linked recessive ichthyosis. *Humangenetik* 15: 261–264
- [5] Sybert, VP, Dale, BA, Holbrook, KA (1985) Ichthyosis vulgaris: identification of a defect in synthesis of filaggrin correlated with an absence of keratohyaline granules. *J Invest Dermatol* 84: 191–194.
- [6] Manabe, M, Sanchez, M, Sun, TT, Dale, BA (1991) Interaction of filaggrin with keratin filaments during advanced stages of normal human epidermal differentiation and in ichthyosis vulgaris. *Differentiation* 48: 43–50.
- [7] Compton JG, DiGiovanna JJ, Johnston KA, et al.(2002) Mapping of the associated phenotype of an absent granular layer in ichthyosis vulgaris to the epidermal differentiation complex on chromosome 1. *Exp Dermatol* 11:518-526.
- [8] Zhong W, Cui B, Zhang Y, et al. (2003) Linkage analysis suggests a locus of ichthyosis vulgaris on 1q22. *J Hum Genet* 48(7):390-392.
- [9] Sandilands A, O'Regan GM, Liao H, et al. Prevalent and rare mutations in the gene encoding filaggrin cause ichthyosis vulgaris and predispose individuals to atopic dermatitis. *J Invest Dermatol*, 2006, 126(8):1770-1775.
- [10] Maestrini E, Monaco AP, McGrath JA et al. A molecular defect in Loricrin, the major component of the cornified cell envelope, underlies Vohwinkel's syndrome. *Nat Genet* 1996; 13:70–7.
- [11] Korge BP, Ishida-Yamamoto A, Punter C et al. Loricrin mutation in Vohwinkel's keratoderma is unique to the variant with ichthyosis. *J Invest Dermatol* 1997; 109:604–
- [12] Takahashi H, Ishida-Yamamoto A, Kishi A et al. Loricrin gene mutation in a Japanese patient of Vohwinkel's syndrome. *J Dermatol Sci* 1999; 19:44–7.

- [13] Matsumoto K, Muto M, Sekin S et al. Loricrin keratoderma: a cause of congenital ichthyosiform erythroderma and collodion baby. *Br J Dermatol* 2001; 145:657–60.
- [14] O'Driscoll J, Muston GC, McGrath JA et al. A recurrent mutation in the Loricrin gene underlies the ichthyotic variant of Vohwinkel syndrome. *Clin Exp Dermatol* 2002; 27:243–6.
- [15] Gedicke MM, Traupe H, Fischer B, Tinschert S, Hennies HC. Towards characterization of palmoplantar keratoderma caused by gain-of-function mutation in Loricrin: analysis of a family and review of the literature. *Br J Dermatol*. 2006;154(1):167-71.
- [16] Armstrong DKB, McKenna KE, Hughes AE. A novel insertional mutation in Loricrin in Vohwinkel's keratoderma. *J Invest Dermatol* 1998; 111: 702-4.
- [17] Ishida-Yamamoto A, McGrath JA, Lam H, et al. (1997) The molecular pathology of progressive symmetric erythrokeratoderma: a frameshift mutation in the Loricrin gene and perturbations in the cornified cell envelope. *Am J Hum Genet* 61:581-589.
- [18] Richard, G.; Brown, N.; Smith, L. E.; Terrinoni, A.; Melino, G.; MacKie, R. M.; Bale, S. J.; Uitto, J. (2000) The spectrum of mutations in erythrokeratodermias – novel and de novo mutations in GJB3. *Hum. Genet* 106: 321-329.
- [19] Ishida-Yamamoto A (2003) Loricrin keratoderma: a novel disease entity characterized by nuclear accumulation of mutant Loricrin. *J Dermatol Sci* 31:3-8.
- [20] Ishida-Yamamoto A, Kato H, Kiyama H, et al. (2000) Mutant Loricrin is not crosslinked into the cornified cell envelope but is translocated into the nucleus in Loricrin keratoderma. *J Invest Dermatol* 115: 1088-1094.
- [21] Suga Y, Jarnik M, Attar PS, Longley MA, Bundman D, Steven AC, Koch PJ, Roop DR. (2000) Transgenic mice expressing a mutant form of Loricrin reveal the molecular basis of the skin diseases, Vohwinkel syndrome and progressive symmetric erythrokeratoderma. *J Cell Biol* 151:401-12.
- [22] Yoneda K, Steinert PM. (1993) Overexpression of human Loricrin in transgenic mice produces a normal phenotype. *Proc Natl Acad Sci USA* 90:10754-8.
- [23] Koch, P.J., P.A. de Viragh, E. Scharer, D. Bundman, M.A. Longley, J. Bickbach, Y Kawachi, Y. Suga, Z. Zhou. M. Huber, D. Hohl, T. Kartasova, M. Jarnik, A.C. Steven, and D.R. Roop (2000) Lessons for Loricrin-deficient mice: compensatory mechanisms maintaining skin barrier function in the absence of a

- major cornified envelope protein. *J. Cell Biol.* 151:389–400.
- [24] Jarnik M, de Viragh PA, Scharer E et al. (2002) Quasi-normal cornified cell envelopes in Loricrin knockout mice imply the existence of a Loricrin backup system. *J Invest Dermatol* 118:102–9.
- [25] Schmuth M, Fluhr JW, Crumrine DC, et al. (2004) Structural and functional consequences of Loricrin mutations in human Loricrin keratoderma (Vohwinkel syndrome with ichthyosis). *J Invest Dermatol* 122:909-922.
- [26] Rossi, A., Jang, S-I., Ceci, R., Steinert, P. M., and Markova, N. G. (1998) Effect of AP1 transcription factors on the regulation of transcription in normal human epidermal keratinocytes. *Invest. Dermatol* 110, 34-40.
- [27] Xu X, Kawachi Y, Nakamura Y, Sakurai H, Hirota A, Banno T, Takahashi T, Roop DR, Otsuka F. (2004) Yin-yang 1 negatively regulates the differentiation-specific transcription of mouse Loricrin gene in undifferentiated keratinocytes. *J Invest Dermatol* 123:1120-6.
- [28] Jang SI, Steinert PM (2002) Loricrin expression in cultured human keratinocytes is controlled by a complex interplay between transcription factors of the Sp1, CREB, AP1, and AP2 families. *J Biol Chem* 277:42268–42279

综述

识别人类耳聋候选基因的生物信息学方法

Bioinformatics methods for identifying candidate genes of human hearing loss

摘要：随着基因组学和功能基因组学相关信息的飞速发展，人类和小鼠耳蜗 cDNA 文库的建立。如今数据库成为选择候选疾病基因的必要方法。疾病的定位信息和功能信息的结合是选择候选疾病基因常见的策略。许多候选疾病基因的生物信息学方法已形成。这些区分优先次序的方法选择最有可能的候选基因进行分析，这将加速发现疾病基因的过程。在本文中概述这些方法及其可及性。

关键词：耳聋，生物信息学，候选基因，识别

识别人类耳聋致病基因的关键在于获取对耳聋发病机制的理解。连锁分析和突变筛选已较早的并成功地运用于染色体定位和鉴定耳聋基因，经识别的耳聋基因数目飞速增多。遵守孟德尔遗传的单基因病的致病突变的识别主要依赖于有好的病人和家系。通过连锁分析的遗传定位是人类遗传学研究的主要依据。虽然定位信息减少了致病候选基因的数目，但这种减少经常不足以迅速克隆疾病基因。在克隆耳聋位点的致病基因有时是困难的：定位区间过大，包含数百个可能的候选基因，在寻找到致病性突变前有时需要分析大量的基因，使得实验室方法费时和费钱。此外，从临床方面的研究寻找患者个体在基因组上的单核苷酸多态性 (SNP) 将产生大量的可能的候选基因^[1]。显然，需要有运用计算机技术的高通量分析方法来识别最应优先识别出的候选基因来做进一步研究。如今，随着基因组学和功能基因组学相关信息的飞速发展，可获得的数据激增，新的分子生物学技术的发展，新的识别疾病基因方法得以发展。在2003年人类基因组序列完成，到目前为止人类基因数目有20,000–25,000个^[2,3]。区分优先次序选择候选基因的方法其目的是选择那些最有可能导致疾病的基因进行突变分析。如在定位区间内有上百个不同的基因，选择候选基因时用这种方法尤其适合。运用这些区分优先次序的方法来选择最有可能的候选基因进行分析，将加速耳聋致病基因的发现过程。下面将识别人类耳聋候选基因的生物信息学方法和策略分述如下。

一. 耳聋概述

耳聋(deafness)是各种不同程度听力损失(hearing loss)的总称。外耳和中耳传音结构、内耳感音器官及听觉神经通路上任何部位及附近的病变都可能引起耳聋。耳聋是一种严重影响人类生活质量的常见先天性疾病，它可以由单一基因突

变或不同基因的复合突变引起,也可由环境因素(如医疗因素,环境暴露,创伤,药物等)或基因和环境两者共同作用而致。听力损失是最常见的感觉障碍。在世界范围内,每 1000 名新生儿中就有 1 名先天性耳聋患儿,50%患儿的耳聋与遗传因素有关^[4,5]。据中国残疾人联合会网站统计,中国 6000 万残疾人口中 2100 万为听力残疾者,居残疾人总数的首位。据人口调查统计,每年出生的新生儿中,重度听力障碍者约占 1‰~3‰。另外,随着全球人口老龄化加速,耳聋的绝对发病率正在显著增高。随年龄增长听阈差距增大,不同程度听力障碍者在人群中的比例,45~64 岁为 14%,65~75 岁为 30%,75 岁以上者高于 50%^[4,6]。尽管近年来遗传性耳聋分子病因学研究领域有许多重要进展,但中国耳聋人群的遗传学研究仍相对贫乏,大量的问题亟待解决。如何降低聋病在人群中的发病率和及时发现、早期治疗,是全社会共同的责任。

遗传性耳聋是指由来自亲代的致聋基因或新发生的突变致聋基因所导致的耳部发育异常或代谢障碍,以致出现听力障碍。遗传性耳聋根据是否伴有耳外组织的异常或病变可分为综合征性耳聋(syndromic hearing loss, SHL)和非综合征性耳聋(nonsyndromic hearing loss, NSHL),其中 SHL 占 30%左右,而 NSHL 约占 70%,非综合征耳聋是最常见的感音神经性聋。在 NSHL 中,根据遗传方式不同将遗传性耳聋分为常染色体显性(DFNA, 15%~20%)、常染色体隐性(DFNB, 80%)、性连锁(DFN X-linked, 1%)和线粒体遗传性耳聋(线粒体遗传所占的比例在不同的人群中从 1~20%之间不等)四类。

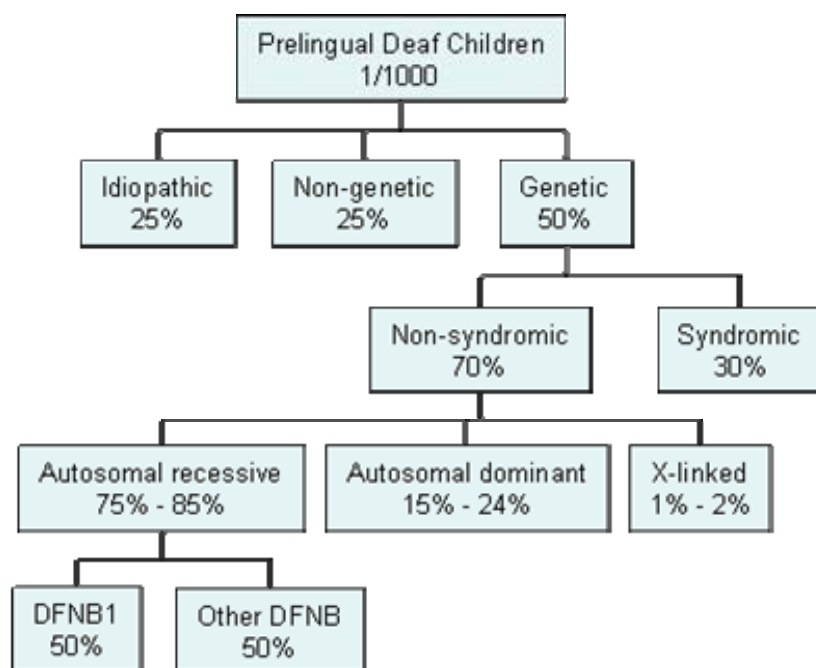


图 1 在儿童中学语前聋的病因分类

遗传性耳聋在 16 世纪就有文献记载,但由于内耳部位深,体积小,研究手段受到限制,因此认识和理解控制听觉系统的基因研究进展很慢。近 10 多年来,随着现代科学技术的发展及分子生物学、遗传学的飞速发展,人们对遗传性聋的认识不断加深,并取得了显著进步,与耳聋相关的基因的定位、分离、克隆及基因的突变得到了深入的研究。

遗传性听力损害具有高度的遗传异质性,根据目前的研究表明与 NSHL 相关的位点有 120 多个,并根据遗传方式和定位的先后进行命名,以 DFNA、DFNB、DFN 和 DFNY 分别代表常染色体显性、常染色体隐性、X 连锁和 Y 连锁 NSHL 位点。有人估计,非综合征型耳聋相关的位点可能超过 100 个,而综合征型耳聋的位点可能多达 400 个^[7]。截止 2007 年 9 月,已定位了 77 个非综合征型耳聋位点,其中常染色体显性遗传 50 个,常染色体隐性遗传 67 个,X-连锁遗传 5 个,Y-连锁遗传 1 个。并且克隆了 22 个常染色体显性遗传 NSHL 基因,25 个常染色体隐性遗传 NSHL 基因,2 个 X-连锁遗传 NSHL 基因^[7](Hereditary Hearing Loss Homepage, <http://webhost.ua.ac.be/hhh/>)。

二. 耳聋基因的定位克隆策略及候选耳聋基因筛选策略

遗传性耳聋基因定位克隆策略包括功能克隆、位置克隆、候选基因克隆、位置候选基因克隆 4 种方法。功能克隆是通过分析从组织中提取的蛋白质或 RNA 的信息进行。由于它们包含了表达基因所具有的大部分信息尤其是功能信息,从而可以预测相应的基因序列。位置克隆是通过连锁分析或者关联分析等遗传统计学方法,分析致病基因与多态性标记之间的关系,从而将基因定位和克隆的策略。候选基因克隆是分析与疾病可能相关的生理生化特征以及基因的表达模式,进一步确定相关的候选基因,再在这些候选基因上寻找与疾病突变表型共分离的突变,以此确定致病基因的方法。被认为最有前途和发展前景的克隆策略是位置候选基因克隆法,即通过连锁分析先将致病基因定位于染色体的某一区域内,分析定位区域内所有基因的功能与表达特征,或者模式动物染色体同源区基因的功能信息,选择适合的候选基因再进行突变筛查检测。以上是利用遗传学手段确定疾病基因的主要策略。虽然功能克隆可以较快地找到与疾病有关的基因,但对于耳聋基因来讲却并不现实;定位克隆则依靠于家系和疾病的群体样本,即使对基因功能和生化产物的了解较少时也可以找到与疾病相关的基因,但对于家系质量要求较高,需要遗传背景单纯,在家系中有多个患者和正常家属成员,最好是地理位置隔离的大家系。而定位候选克隆克服了经典的定位克隆纯粹依靠连锁分析进行染色体定位的繁冗而缓慢的弊端,大大加快了遗传性疾病易感基因的克隆工作的进程。总之,定位候选克隆策略(positional candidate cloning)是传统定位克隆

(positional cloning)的改进和发展,随着人类基因组计划的完成,定位候选克隆在克隆疾病相关基因中的有效性而日益受到重视,已成为一种有效的克隆疾病相关基因的方法。人类基因组计划已确定了一系列分布于全基因组 24 条染色体上的分子标记。根据这些分子标记的序列和定位信息,结合杂交或 PCR 的方法可以快速检测染色体上与遗传性疾病相关的热点位置,进而从中克隆疾病相关基因。

目前我们定位耳聋基因时常用的技术路线是: a.获取高信息量的非综合征遗传性耳聋家系; b.确定耳聋家系的遗传方式(显性、隐性、X-连锁或线粒体母系遗传); c.根据确定的遗传方式选择候选的耳聋基因的 DNA 多态标记(在一种遗传方式中可以有多种表型,即发病时间、听力损失程度、累及频段、病程发展、伴随症状的不同形式,应根据所拥有的耳聋家系的表型进行候选 DNA 多态标记的筛选); d.应该连锁分析软件确定或排除候选的耳聋基因位点; e.如与候选的耳聋基因位点均无连锁关系,则应用基因组每隔 20cM 或 10cM,或 5cM 的图距单位选择一个 DNA 标记,对家系成员作全基因组 DNA 标记座位的基因型分析,所有家系成员的基因型和表型刻录输入给计算机作连锁分析或区间定位分析,从而判断这个家系的表型性状的基因在基因组上的位置。通过单体型分析受累个体中观察到的交换事件,将该疾病基因定位于某一特异的染色体亚区,然后采用一些方法识别合适的候选基因(candidate genes),然后逐个检查是否具有疾病基因的确切证据,以克隆其疾病基因。

用来识别疾病基因的方法最后得到的是候选基因(candidate genes),这些基因必须逐个进行检验,以确定与疾病的相关性。但是有二种疾病基因识别的方法是直接从假设的候选基因开始的,即先假设某一基因为疾病位点,然后再检查该基因与疾病的相关性。这种候选基因的入围可能依据其产物的特征与疾病发病过程一致。显然,在未知某种疾病的基因定位也可以提出其候选基因,这称为不依赖定位的候选基因法。如果候选基因与疾病基因染色体的定位一致,则对该候选基因的信心就更大些。这种定位性候选基因法近年来大获成功。随着疾病基因和人类基因愈来愈多地在染色体上定位,该方法已渐渐成为疾病基因定位的主要方法。

人类耳聋候选疾病基因的选择可从以下几个方面考虑。

(1).依据适当的表达方式和功能:

根据有些疾病的发病机理可直接提出具有合适表达方式或功能的疾病候选基因。例如听觉缺陷可能涉及发育胚胎的耳蜗形成之前或期间表达的基因,选择耳聋候选基因的一个标准是它所对应的 mRNA 在内耳细胞中选择性表达。如果某一基因在此发育阶段时在耳蜗表达则为一个候选基因。从内耳特异组织筛选出可能参与听力形成的基因,在耳聋的家系中进行突变检测,从而克隆耳聋基因。

例如, COCH 基因编码的蛋白 cochlin, 它的 cDNA 克隆最初应用减除杂交法发现 cDNA 中非常丰富, 在内耳表达, 不存在于其它器官中。通过 Northern-blot 证实了 cochlin 的 mRNA 在内耳中特异表达。COCH 基因定位在 DFNA9 的连锁区域内。患者的基因组 DNA 测序分析发现 DFNA9 是由于 COCH 突变所造成的¹。

基因表达分析在研究内耳疾病时是一种非常重要的生理基因组工具, 基因表达分析的候选基因通常选自内耳 cDNA 文库。尽管缺乏从内耳标本中获取的材料延缓了 cDNA 文库的构建, 但仍有一些很好的内耳 cDNA 文库报道。然而大多数基因病不只在内耳中表达, 在其它组织中也有表达, 仅有一小部分局限在内耳中。

尽管近年来在寻找耳聋基因方面已经取得了很大的成就, 但仍有大量的非综合征型听力损失(non-syndromic hearing loss, NSHL)基因有待发现。部分原因是由于许多 NSHL 家系较小, 没有足够的重组信息能够将 NSHL 基因定位在较窄的区段内。因此, 在许多家系中, 需要在基因组 DNA 几个 MB(megabases, MB)中进行分析, 为 NSHL 位点寻找候选基因。表达序列标签(expressed sequence tags, ESTs)分析在为人类疾病寻找定位候选基因方面是一种非常有用的工具。ESTs 提供短的核苷酸序列充当新基因和已知基因的标示, 并且能够用作探针在适合的 cDNA 文库中克隆基因。Skvorak 等^[8]构建了人类耳蜗 cDNA 文库, 并从其中产生了 4000 多个的 ESTs。收录在 GenBank ESTs 数据库中几个 EST 可以组装成一个连续的 cDNA 草图, 成为重叠群, 重叠群至少显示蛋白所对应的部分氨基酸序列, 应用基因组 DNA 的信息便可以将其补充完整。

(2).疾病模型动物表型同源基因:

如果某一动物疾病表型与某一人类疾病很相似, 则人类疾病有可能是由于相关的同源基因突变所致。如果动物的疾病基因已知, 就可以用以帮助人类疾病基因的识别和鉴定, 也可以用于检验该基因是否与人类疾病相关。例如, 小鼠内耳与人类内耳在结构和功能上相似, 小鼠是研究人类耳聋基因的首选模型动物。小鼠全基因组的突变率很低, 常规的基因敲除(knock-out)和转基因(transgenesis)突变规模小, 且研究的是已知基因的功能。近年, 采用化学诱导突变的表型筛选(phenotype screens)已成为日益关注的目标, 乙烷亚硝基脲(N-ethyl-N-nitrosourea, ENU)诱变的大规模表型筛选能获得大量新突变, 确定了一些新基因, 同时也发现了与听觉及平衡功能有关的突变和基因, 为内耳感觉系统的结构和发育提供了很多重要的信息^[9]。估计, 小鼠存在近 200 个与听力损害有关的突变基因。小鼠突变对于展现哺乳动物基因功能和模拟人类疾病具有关键性的作用, 对分析和探讨疾病的发病机制具有特别重要意义, 从而也推动了人类聋病基因的发现。与小

鼠耳聋相关的基因有一些基因在人体的同源基因已证实是耳聋基因,其余基因仍是克隆和鉴别新的耳聋基因的良好候选基因。将一个突变基因在小鼠图谱中定位后,可以观察定位在附近的表型异常是否有一种表型与定位基因引起的缺陷一致。如果发现合理的联系,就可以对有问题的基因进行分子鉴定以识别突变。对有听觉缺陷的小鼠进行近交建立听觉缺陷小鼠模型,再对这些小鼠进行缺陷基因的鉴别和克隆,克隆出来的小鼠耳聋基因的人类同源基因就是人类耳聋基因的良好候选基因。耳聋基因 MY07A 就是通过此方法克隆的。

(3).与某一导致相似人类疾病的基因同源或功能相关:

有时不同的基因位点突变可能产生同样的临床结果,因此不同基因引起的表型有时可以区别,有时则不能。如果已识别某一基因为这种疾病基因,那么序列或功能上紧密相关的其他基因可能是相似疾病的候选基因。例如所识别基因可能是多基因家族的成员之一,其他成员也许是相似疾病的候选者。若某一基因的功能与某已定位的具有相似表型疾病基因密切相关也可以作为候选基因。基因功能上的相关体现在它们产物功能上的相关,如编码受体和其相应配体,或编码同一代谢途径中或发育过程中的相互作用的蛋白。

最有力的候选基因是那些与疾病基因具有同样染色体区带定位的基因。定位克隆耗费大量人力,相反,检查已知基因是否具有病人特异的突变却相对容易。于是在疾病定位区域内的已知基因都是值得考虑的候选基因。越来越多的基因已经被定位在染色体区带上,那么经过数据库搜寻在适当的位置找到一个或更多的候选基因的机会也是很大的。分子病理学很复杂,单基因分离后常常证明原来有关的疾病基因的生化功能的预测是错误的。因此与相关疾病基因具有相同染色体定位的候选基因将最具有价值。

三. 耳聋候选基因的生物信息学分析方法

以前在定位区域内寻找致病基因非常困难,需要根据基因的表达模式或编码蛋白的功能来进行。当前,人类基因组序列完成以后,人类基因组研究的重心正在由“结构”向“功能”转移,一个以基因组功能研究为主要内容的所谓“后基因组时代”(post-genomics),也即功能基因组(functional genomics)时代已经到来。如何获取基因的功能信息,即与人类重大疾病和重要生理功能相关的基因信息,就摆在了我们面前。研究者可以直接调用连锁区域内的 DNA 序列。而且,人类基因组计划还提供了更多的信息,比如已知基因的位置和预测还没有被研究的新基因的存在。

识别人类耳聋候选基因的生物信息学方法与策略可从如下 4 个方面考虑: 1.

根据已克隆耳聋基因的表达情况和所致耳聋临床表现来选择同源或功能相关的候选基因；2.根据耳蜗基因表达相关数据库选择在耳蜗有表达的基因；3.根据小鼠耳聋基因选择候选基因；4.识别候选基因的生物信息学分析软件。

3.1. 根据已克隆耳聋基因的表达情况和所致耳聋临床表现来选择同源或功能相关基因

遗传性耳聋的基因型和表型之间存在着因果对应关系,而且由于高度遗传异质性的存在,这种对应关系显得尤为复杂,一个耳聋基因的不同的致病突变位点可能导致耳聋表型和发病年龄等的差异,随着更广泛、更深入的研究,这种基因型和表型的对应关系将更加明确,也就能更好地揭示耳聋的生理病理机制。

截止 2007 年 9 月,已定位了 77 个非综合征型耳聋位点,其中常染色体显性遗传 50 个,常染色体隐性遗传 67 个,X-连锁遗传 5 个,Y-连锁遗传 1 个。并且克隆了 22 个常染色体显性遗传 NSHL 基因,25 个常染色体隐性遗传 NSHL 基因,2 个 X-连锁遗传 NSHL 基因^[7](Hereditary Hearing Loss Homepage, <http://webhost.ua.ac.be/hhh/>) (如表 1 所示),其中 7 个基因突变既可以导致 DFNA 也可以导致 DFNB,另外还有 11 个 NSHL 基因突变也可以导致 SHL。对于一个特定的 NSHL 位点,可能不止一个耳聋致病基因,如 DFNA2(*GJB3*, *KCNQ4*)和 DFNA3(*GJB2*, *GJB6*)。这提示染色体上特定区域有成簇耳聋基因的存在。

众多的位点被定位也说明了 NSHL 具有相当高的遗传异质性。这些耳聋基因编码的产物分属不同的蛋白家族:包括细胞外基质蛋白,肌球蛋白和其他骨骼肌蛋白,间隙连接蛋白和离子通道,以及转录因子和一些功能未知的基因表达产物。它们涉及听力器官的胚胎期发育、功能单位的形成、正常功能的维持等多个方面。从 NSHL 基因产物的功能方面, NSHL 基因分为五类。

- 1.细胞外基质成分(*TECTA*, *COL11A2*, *OTOA*, *COCH*);
- 2.细胞骨架蛋白(*MYH9*, *MYH14*, *MYO1A*, *MYO3A*, *MYO6*, *MYO7A*, *MYO15*, *ACTG1*, *ESPN*, *CDH23*, *USH1C*, *PCDH15*, *HDIA1*, *STRC*, *OTOF*) ;
3. 离子内环境稳态中涉及的分子(*KCNQ4* *GJB2*, *GJB3*, *GJB6*, *GJA1*, *SLC26A4*, *CLDN14*, *TMPRSS3*, *WFS1*) ;
- 4.转录因子(*POU4F3*, *POU3F4*, *EYA4* 和 *TFCP2L3*) ;
- 5.未知功能的基因(*DFNA5*, *WFS1*, *TMIE*, *TMC1*, *DFNA5*)。

表 1 常染色体显性遗传及常染色体隐性遗传耳聋基因

序号	位点	基因	序号	位点	基因
1		CRYM ^[10]	1		PRES ^[34]
2	DFNA1	DIAPH1 ^[11]	2	DFNB1	GJB2 ^[14]
3	DFNA2	GJB3 ^[12]	3	DFNB1	GJB6 ^[35]
4	DFNA2	KCNQ4 ^[13]	4	DFNB2	MYO7A ^[36]
5	DFNA3	GJB2 ^[14]	5	DFNB3	MYO15 ^[37]
6	DFNA3	GJB6 ^[15]	6	DFNB4	SLC26A4 ^[38]
7	DFNA4	MYH14 ^[16]	7	DFNB6	TMIE ^[39]
8	DFNA5	DFNA5 ^[17]	8	DFNB7/DFNB11	TMC1 ^[31]
9	DFNA6/DFNA14	WFS1 ^[18, 19]	9	DFNB8/DFNB10	TMPRSS3 ^[40]
10	DFNA8/DFNA12	TECTA ^[20]	10	DFNB9	OTOF ^[41]
11	DFNA9	COCH ^[21]	11	DFNB12	CDH23 ^[42]
12	DFNA10	EYA4 ^[22]	12	DFNB16	STRC ^[43]
13	DFNA11	MYO7A ^[23]	13	DFNB18	USH1C ^[44, 45]
14	DFNA13	COL11A2 ^[24]	14	DFNB21	TECTA ^[46]
15	DFNA15	POU4F3 ^[25]	15	DFNB22	OTOA ^[47]
16	DFNA17	MYH9 ^[26]	16	DFNB23	PCDH15 ^[48]
17	DFNA20/26	ACTG1 ^[27]	17	DFNB24	RDX ^[49]
18	DFNA22	MYO6 ^[28, 29]	18	DFNB28	TRIOBP ^[50, 51]
19	DFNA28	TFCP2L3 ^[30]	19	DFNB29	CLDN14 ^[52]
20	DFNA36	TMC1 ^[31]	20	DFNB30	MYO3A ^[53]
21	DFNA44	CCDC50 ^[32]	21	DFNB31	WHRN ^[54]
22	DFNA48	MYO1A ^[33]	22	DFNB36	ESPN ^[55]
			23	DFNB37	MYO6 ^[56]
			24	DFNB59	PJVK ^[57]
			25	DFNB67/DFNB67	LHFPL5 ^[58]

一般来说,呈常染色体显性遗传的耳聋表现为学语后聋,听力损失或轻或重,可单独累及某个频率段或各个频率段,病程发展多呈进行性;而呈常染色体隐性遗传的耳聋表现为学语前聋,听力损失一般为中、重度以上,全频听力损失多见。

根据不同的 NSHL 相关基因导致耳聋的发病率、发病时间、听力损失程度、累及频段、病程发展、伴随症状和相关检查,等相关特征性临床情况判断相应的耳聋基因,此耳聋基因序列或功能上紧密相关的其他基因可能是耳聋位点的候选基因(见图 2)。所识别耳聋基因可能是多基因家族的成员之一,其他成员也许是相似临床表现的耳聋的候选基因。若某一基因的功能与某已定位的具有相似表型

疾病基因密切相关也可以作为候选基因。基因功能上的相关体现在它们产物功能上的相关,如编码受体和其相应配体,或编码同一代谢途径中或发育过程中的相互作用的蛋白。

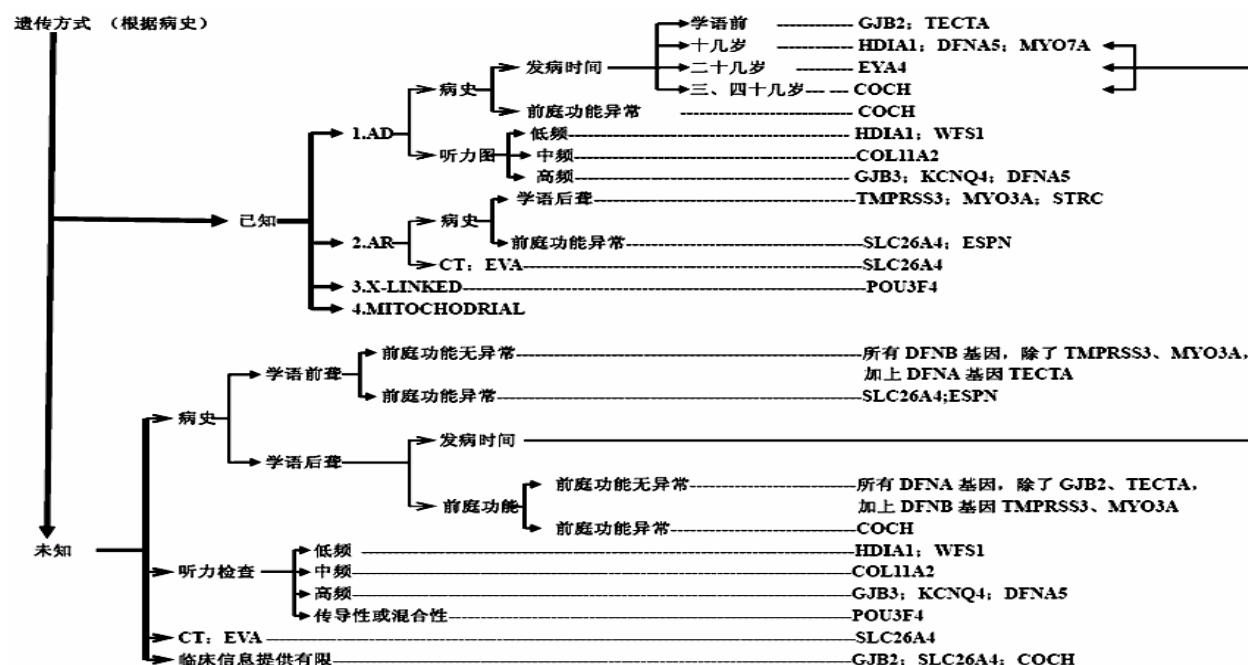


图2 非综合征型耳聋相关基因及相关特征性临床情况

3.2. 根据耳蜗基因表达数据库选择在耳蜗有表达的基因

通常, 基因的表达情况可通过基因表达相关数据库查询, 如 UniGene、Gene Expression Omnibus(GEO)数据库、TIGR Human Gene Index、SAGE Genie(The Cancer Genome Anatomy Project)、GeneNote、Gene Expression Atlas、Stanford Microarray Database 等。为鉴定在耳蜗的基因表达情况和识别耳聋的候选基因, 通过构建人类和小鼠的耳蜗 cDNA 文库, 一些耳聋基因表达数据库已产生, 促进了对耳蜗发育, 内环境稳定性和听力下降的发病机理的理解^[9,59,60]。通过分析耳基因表达数据库, 可加快耳聋候选基因的识别过程, 这些耳基因表达数据库包括 Morton 耳蜗 EST 数据库^[61] (Morton Cochlear EST Database)¹, 耳发育过程中基因表达情况表^[62] (table of gene expression in the developing ear), 及 NEIBank 的耳文库^[63] (ear libraries on NEIBank)。

1. Morton 耳蜗 EST 数据库^[61] (Morton Cochlear EST Database)

隶属于 Brigham 妇女医院和哈佛大学的 Morton 实验室致力于各种哺乳动物内耳的分子生物学的研究, 研究重点是识别耳聋基因, 了解在内耳中的各种蛋白质的功能和研究正常或听觉缺陷的耳中的基因表达情况。耳蜗 EST 数据库包括两个部分: Human Cochlear ESTs 和 Inner Ear cDNA Microarrays。

人类耳蜗 ESTs(Human Cochlear ESTs): Brigham 妇女医院病理科收集的人类耳蜗用于生成 RNA 和 cDNA 文库。这个文库捐赠给 IMAGE 协会, 建立表达序列标志 ESTs(dbEST Library ID.371)。①人类胚胎耳蜗 cDNA 文库 EST 数据(Morton Human Fetal Cochlea cDNA Library EST Data): 共列出了 15000 多个 ESTs。这个数据库提供了 ESTs 的 GenBank 登录号和识别出的相应的基因(若是已知的基因)。②人类耳蜗 EST 消减文库(Subtracted Human Cochlear Library EST Data): 有少部分的 cDNA 文库来自于人类胚胎 mRNA(dbEST Library ID 1822)。根据选择性扩增差异表达基因(Clontech PCR-Select cDNA Subtraction) 的方法将人类胚胎的耳蜗与脑, 睾丸, 肾, 子宫和纤维原细胞的 cDNA 标准化, 建立此文库, 共列出 3880 条 ESTs。

内耳 cDNA 微阵列(Inner Ear cDNA Microarrays): Morton 实验室建立了两个不同的内耳相关的 cDNA 微阵列。①Human Cochlea: 人类耳蜗 cDNA 微阵列来源于 Morton 胚胎耳蜗 cDNA 文库的克隆。芯片包含 5000 多个 cDNA 克隆, 与阳性和阴性对照一样, 附加的耳聋基因也加入此芯片中。②Mouse Organ of Corti: 小鼠 Corti 器微阵列来源于 NIH 的 Kachar 实验室的小鼠 Corti 器 cDNA 文库。此 cDNA 芯片包含 9000 多个 cDNA 克隆, 与阳性和阴性对照一样, 附加的耳聋基因也加入此芯片中。

2. 耳发育相关基因表达数据库^[62] (Table of gene expression in the developing ear, <http://www.ihr.mrc.ac.uk/hereditary/genetable/>)

此数据表由英国 MRC 听力研究所和欧委会建立及支助的。此数据库共包含 766 条数据, 最近更新日期为 2002 年 4 月。数据库收集了大量的在不同的动物种类的内耳发育过程中在不同时期有表达的基因。数据库表格的第一部分是所有染色体位置已知的小鼠(资料来自于 MGD 数据库)和人类(资料来自于 OMIM 数据库) 基因的列表。这些基因已根据它们在小鼠染色体上的位置进行了升序排序, 即从 1 号染色体的着丝粒附近开始起始位置。例如, 当唯一的人类基因的位置已知, 将此基因插入数据表中与小鼠对应的保守同线性(conserved synteny)的位置上(见 <http://www.ncbi.nlm.nih.gov/Homology>)。数据表的第二部分列出染色体位置未知的基因和蛋白质, 其按照字母顺序排序。

检测基因或蛋白质表达的方法有利用特殊 RNA 探针的原位杂交(in-situ hybridisation)、或用抗体检测蛋白质表达的免疫定位(immunolocalisation)。其它方法还包括 Northern blotting, 逆转录酶 PCR(reverse transcriptase polymerase chain reaction, RTPCR), 从耳组织中提取蛋白质及 mRNA 来产生 cDNA 和报道基因。

3. Ear Libraries of NEIBank^[63]

(<http://neibank.nei.nih.gov/cgi-bin/libList.cgi?tissue=ear>)

由美国国立卫生研究院的眼科研究所(National Institutes of Health, National Eye institute)建立的小鼠 Corti 器(OC)的表达序列标志(EST)数据库, 有 20000 个独立的克隆经过分离, 分析和聚类成 8690 个唯一的基因族。大量的 Corti 器相关的新基因被识别出来。在序列比对时经常发现已知基因的选择性剪接本与 Corti 器功能相关。通过生物信息学分析将小鼠 Corti 器的 ESTs 找到对应的同线性的与人类常染色体显性和隐性耳聋相关的人类染色体位置, 这为识别人类疾病提供了有用的新的定位候选基因。经过此资源既可查看在听觉过程上在 Corti 器中基因的表达谱情况, 又是识别新的在听觉过程中起重要作用基因的工具。

数据库包括四部分: ①NbLib0011: dbEST human cochlea unnormalized; ②NbLib0053: NIDCD mouse organ of corti unnormalized (gi); ③NbLib0123: NIDCD mouse taste cell subtracted (ia); ④NbLib0010: NIDCD human fetal cochlea subtracted (n)。

3.3. 根据小鼠耳聋基因识别候选基因^[64]

随着小鼠基因组测序的完成, 小鼠耳聋基因克隆正在逐年增加。由于小鼠繁殖周期短, 易于选择交配, 小鼠和人类的基因组及内耳结构非常相似, 是识别新的耳聋基因及研究基因在体内的功能的理想动物模型, 从而也推动了人类聋病基因的发现。现在已经发现了数百个小鼠耳聋和 / 或平衡功能缺陷模型, 许多与人类耳聋直接相关, 为内耳感觉系统的结构和发育提供了很多重要的信息。估计, 小鼠存在近 200 个与听力损害有关的突变基因。与小鼠耳聋相关的基因有一些基因在人体的同源基因已证实是耳聋基因, 其余基因仍是克隆和鉴别新的耳聋基因的良好候选基因。

在两个小鼠耳聋相关的网站上, 可查找到小鼠耳聋的相关资料。

1. 遗传性听力损害相关的小鼠品系和突变(Hereditary Hearing Impairment in Mice, <http://www.jax.org/hmr/>)

由国立耳聋和其它交际障碍研究所(National Institute on Deafness and other Communication Disorders, NIDCD)支助, Jackson 实验室通过大规模地表型筛选识别出与遗传性听力损害相关的小鼠品系和突变。根据国际上最新的研究文献, 更新一些数据, 如人类听力障碍的小鼠模型、与听觉及平衡功能有关的小鼠基因和突变、耳聋位点在染色体上的图位、与采用小鼠模型的听觉研究的网络资源的链接。其主要研究重点是将小鼠作为人类遗传性听力障碍的动物模型。数据包括的内容有:

①人类听力障碍的小鼠模型(Mouse Models of Human Hearing Disorders): 列出人类非综合征及综合征耳聋基因和对应的小鼠基因和突变。

②近亲交配小鼠品系中老年性聋潜在的遗传因素(Genetic factors that underlie age-related hearing loss in inbred mouse strains): 数据表列出在近亲交配的小鼠品系中会导致老年性聋(age-related hearing loss,AHL)的遗传因素, 其遗传位置图谱, 基因定位的所采用的方法, 已知的易感和抗性位基因的品系, 和相关参考文献。

③耳聋的遗传修饰因子和二基因遗传 (Genetic modifiers and digenic inheritance of hearing loss): 数据表列出已报道的在人和鼠的遗传性疾病中可改变耳聋表达的遗传修饰因子位点和二基因遗传的案例。

④已报道导致内耳畸形或异常的小鼠突变(Mouse mutations reported to cause inner ear malformations or dysfunction): 数据表包括 200 多个单个基因位点。各个条目提供了以下信息: 基因符号, 基因名称, 染色体图位置, 突变符号, 突变名称, 突变类型, 相关耳部症状, 听觉测定, 前庭功能, 和相关参考文献。通过基因符号将条目按字母顺序排序。

2. 小鼠耳聋突变(Deaf Mouse Mutants, <http://www.sanger.ac.uk/PostGenomics/-mousemutants/deaf/>)

Sanger 研究列出的小鼠耳聋突变数据, 数据表包括已知涉及耳聋和/或平衡功能障碍的主基因或位点。内容包括 9 个部分: 中耳或耳廓缺陷的小鼠突变; 内耳畸形的小鼠突变; 伴神经上皮缺陷(neuroepithelial defects)的小鼠突变; 伴内淋巴内环境稳定缺陷的小鼠突变; 伴盖膜或耳石膜(otolithic membrane)缺陷的小鼠突变; 伴外周神经缺陷的小鼠突变; 伴中枢听觉系统缺陷的小鼠突变; 伴听阈增高但病理未知的小鼠突变; 可作为耳聋的候选基因但无听觉系统相关信息的小鼠突变。

3.4. 识别候选基因的生物信息学分析软件

在一些数据库中收集了目前已知的人类致病疾病基因的突变及导致症状有关信息, 如 OMIM (Online Mendelian Inheritance in Man)^[65], LocusLink^[66]和人类基因突变数据库^[67]。这些信息和序列数据全来源于人类和其它物种的基因组, 使识别疾病相关基因的内在本质的成为可能, 并可更深入了解。

数据库已经成为现在基因搜索的核心来源。检索系统如国家生物技术信息中心的Entrez^[68], 序列检索系统(the Sequence Retrieval System)^[69]和Maarten检索系统^[70]提供了简便和快速地获得常用数据库所收集到的信息。这些检索系统的主要

用于获取一组符合用户查询要求的数据库条目。当结合了定位信息与功能途径,识别疾病基因才有最可能成功。如California大学的圣克鲁斯(Santa Cruz)基因组浏览器和Ensembl^[71,72]基于基因组注释数据的整合,根据选择染色体位置,基因表达和基因语义学(gene ontology, GO)^[73]来提取数据,通过步进式界面(如EnsMart^[74])生成结果。使用这些数据库的界面来选择疾病候选基因,很大程度上依靠研究者操作技能。

随着基因组学和功能基因组学相关信息的飞速发展,可获得的数据激增新的识别疾病基因方法得以发展。把候选基因区分优先次序方法的目的是选择那些最有可能导致疾病的基因进行突变分析。这些区分优先次序的识别疾病有关的候选基因的生物信息学方法使用了不同的数据资源和策略,其目的是选择最有可能的候选基因进行分析。

1. Genes to Diseases (G2D)

(<http://www.bork.embl-heidelberg.de/g2d/>; http://www.ogic.ca/projects/g2d_2/)

Perez-Iratxeta 等发展了 Genes2Diseases (G2D)方法^[75,76]。G2D 系统是基于生物医学文献检索和结合病理学条件的基因语义学术语,通过 GO 注释和疾病相关基因的同源性分析识别出候选基因。用主题词 C(表型特征)和主题词 D(化学功能)在 Medline 中检索摘要^[77]。如 Medline 摘要中主题词 C 与 D 一起出现,就认为化学目标和表型术语有关联。RefSeq 数据库的序列用于关联从 Medline 链接 GO 功能注释的主题词 D 的相同序列。通过化学术语实现了表型术语与功能术语关联的可能性。疾病有关的文献用于筛选主题词 C 术语,其用于决定在疾病和有 GO 注释的基因之间的关联。预测结果的得分是经过比较了 100 种致病基因已知的疾病。Perez-Iratxeta 等平均测试了 30Mb 的候选区间。假定 20,000 至 25,000 个人类基因和平均每 120kb 一个基因,用这个方法计算其丰度为 8-31 倍(fold)。G2D 方法扩展了表达序列标记的数据。表型输入的系统也改进,减少先前要求输入临床知识。

2. POCUS

Turner等发展的POCUS方法使用功能域注释和基因表达数据来区分候选疾病基因的优先次序^[78]。这种方法假定涉及相同疾病的基因必共有GO注释,蛋白质域和相同的表达模式。评分系统包含这些资源允许一个基因在候选疾病区间内的得分。POCUS搜索相同疾病的位点之间的功能注释的代表性。大的候选区间很有可能享有类似性,并较不可能产生基因联系。这种方法用29种疾病进行测试,丰度达到12-42倍。这种方法不能在线使用,但POCUS程序可下载。

3. 特定基因的特征用于候选疾病基因的识别

人类/真核基因的序列分析显示长链的氨基酸的人类蛋白质比短的蛋白质通常与遗传性疾病有关^[79]。Lopez-Bigas 等发现涉及遗传性疾病的蛋白质通常趋向于较长, 保守和无近旁系同源(close paralogues)^[80]。疾病基因在其它种属里也经常是保守的, 但这也可能是因已知疾病基因的优先排序而产生的偏差。使用这些序列特征的疾病基因预测系统可在线获得(<http://cgg.ebi.ac.uk/services/dgp/>)。DGP 是一个人类基因涉及可能与遗传性疾病有关的概率的数据库。这种概率是用有参数的数据挖掘算法计算出来的, 这些参数是保守性, 种系发生范围内的, 蛋白质长度和旁系同源。

4. PROSPECTR

Adie等用交互决策树分析序列特性^[81]。他们发现任意基因和疾病基因之间差别有许多特征, 包括: 基因/cDNA/蛋白质/3'非翻译区长度, 外显子数目, 邻近基因的距离, 在小鼠中高度保守水平, 编码的信号肽和5'CpG岛。候选基因识别测试显示丰度为 2-25 倍。数据可使用 PROSPECTR 网页服务(<http://www.genetics.med.ed.ac.uk/prospectr/>)获得。使用者可在任一条目或基因组区域中将涉及疾病的可能性的基因进行排序。这种方法用GO术语、内在功能域及基因表达数据扩展。SUSPECTS网页服务使用PROSPECTR, 能以基因与疾病相关的可能性排序(<http://www.genetics.med.ed.ac.uk/suspects/>)。Smith等使用比较分析, 发现在疾病基因和非疾病基因之间有类似的差异。使用判别式分析, 他们发现这些差别可有助于预测人类疾病基因^[82]。然而, 这些方法不能在线获得。

5. eVOC

Tiffin 等发展一种使用解剖学的语义学(ontology, eVOC) 的方法^[83], 它结合文献和人类基因的表达数据^[84]。这种方法运用 eVOC 为控制的词汇来发现在PubMed摘要中与疾病解剖有关的术语。解剖有关的术语排序, 候选基因的选择使用最高排序的术语。选择的候选基因有基因表达模式, 其匹配疾病关联或受累的组织。使用了当前存在的 EST 和 SAGE 数据, 及混合的微点阵数据进行分析。这种方法的丰度达到 1.5-3.0 倍, 发现的正确基因占总数的 85%。数据和程序可在线获得(<http://www.evoontology.org/>)。

6. GeneSeeker

遗传性疾病其受累组织和器官与候选基因表达谱之间的联系已在利用^[85,86]。GeneSeeker 使用人类和小鼠的表达和表型有关各种数据库

(<http://www.cmbi.ru.nl/GeneSeeker/>)。这些信息结合了定位区间内两个种属的基因数据。系统使用不同的在线数据库,而不是本地数据,因此能及时进行数据挖掘。GeneSeeker 方法不同于其它运用跨种属的区分候选基因优先次序的方法。模式生物的知识使比较候选基因的选择更合理。一个基因在两个不同种属中导致相类似的表型,在这种情况下才运用这种方法。但是,在人类和模式生物之间的直接比较会因解剖学上的差异而变得复杂。运用 GeneSeeker 测试十种疾病的候选基因其丰度达到 7-25 倍^[85]。

7. GFINDER

Masseroli 等发展了一种 GFINDER 系统。这个网页工具允许研究者通过给定的一组的序列标识符从一些数据库中挖掘注释信息^[87]。在这个系统中人工设置过滤参数来选择疾病基因和执行统计分析。GFINDER 系统(<http://www.bioinformatics.polimi.it/GFINDER/>)也整合了 OMIM 的临床概述。为获得两个控制的词汇以适合计算的目的,表型数据进行标准化和结构化。缺乏预定义的策略使系统的效率严重依赖于研究者操作技能。作者仅列出了他们的方法选择的几个例子,不能评估这种方法的丰度。

8. MimMiner

van Driel 等设计一种方法来比较来源于 OMIM 数据库的表型信息,其通过自动化的全文分析技术使用文本相似性测量,而不是预定义术语分类,和分析表型与基因型之间关系^[88]。他们发现表型类似性得分,其通过自动化而量化,确实与许多基因功能量度有联系,包括蛋白质序列,共享蛋白质基序的相似性,功能性注释和直接的蛋白质-蛋白质相互作用。这些数据支持表型的关系的想法,可用作在基因和蛋白质水平中生物学的功能的相互作用的指示器。这种表型-表型排序评分可在线查询(<http://www.cmbi.ru.nl/MimMiner/>)。这种方法可用于研究在基因型水平的表型关系,通过其表型成为一种功能基因组学工具^[89]。

9. TOM

依靠高并列装置如基因表达芯片产生大量的生物学数据使新的分子遗传学方法成为可能。通过查询大量的表达数据而推断其可能的生物学意义。基于这个原理, Rossi 等设计的网页工具 TOM <http://www-micrel.deis.unibo.it/~tom/>, 能有效地提取遗传性疾病的候选基因^[90], 主要针对多基因疾病的候选基因预测。服务要求有先前的至少另一个基因导致这种疾病和连锁区域的知识,或其它两个疾病关联的遗传区间。算法使用的信息是公共数据资源,包括空位,表达和功能数据

库。给予查询, TOM 将选择和列出一个或多个候选基因。这种方法允许遗传学家节省费用和通过整个家系追踪遗传标记的时间, 并可提高识别出疾病基因的机率, 尤其是罕见疾病。TOM 允许用户结合功能的和定位的数据, 可有效地寻找遗传性疾病新的候选基因。

10. 共同径路扫描(Common Pathway Scanning ,CPS)和共同模块谱(Common Module Profiling, CMP)

George 等提出两种区分候选基因优先次序的方法用于实验研究: 共同径路扫描(Common Pathway Scanning ,CPS)和共同模块谱(Common Module Profiling, CMP)^[91]。CPS 基于假设—共同的表型伴随蛋白质方面的异常, 参与相同的复合体或途径。CPS 应用源于蛋白质-蛋白质相互作用(PPI)的网络数据和途径数据库来识别基因之间的关系。CMP 运用功能域依赖序列相似性方法识别可能的候选基因, 基于假设—相似功能的基因的分离将导致相同的表型。两种算法使用两种输入数据的格式: 输入疾病基因或多个疾病位点。当使用已知疾病基因作为输入, 方法的灵敏度为 0.52, 特异性为 0.97, 可减少候选列表 13 倍(fold)。运用多个位点, 这种方法成功识别为疾病基因的灵敏度为 0.84, 特异性为 0.63。

这些识别人类疾病候选基因的不同方法覆盖了不同的概念。它们采用功能有关资料和文献资料, 基因特殊性特征, 基于解剖学的基因/蛋白质表达数据或表型比较分析。根据要达到不同的丰度水平而用不同的方法, 使之相互自我补充。

这些在讨论中的结果提示表型是一重要的资源, 揭示了生物学功能, 而要特别注意的是表型描述的需要进行标准化。各种方法系统描述表型数据已计划出, 等待进一步的发展^[92,93]。将候选疾病基因的识别方法改进的本质在于建立标准词汇表, 能跨数据库和跨种属使用。进一步的挑战是制定, 精练和整合这些方法成一个有助于阐明和理解(复杂性)疾病的机理的系统。

展望

总之, 遗传性耳聋表型的多样性和遗传的异质性使克隆出全部的耳聋相关基因面临着巨大的挑战。随着基因组学和功能基因组学相关信息的飞速发展, 相关生物信息学数据库激增, 各种新思路新方法层出不穷。利用生物信息学方法和相关数据库, 可大大提高了获得有价值候选基因的可能性。定位克隆方法的周期也将不断缩小, 为人类最终克隆各种疾病基因提供了可能性。

参考文献:

- [1] Smyth,D.J., Cooper,J.D., Bailey,R., Field,S., Burren,O., Smink,L.J., Guja,C., Ionescu-Tirgoviste,C., Widmer,B., Dunger,D.B. et al. A genome-wide association study of nonsynonymous SNPs identifies a type 1 diabetes locus in the interferon-induced helicase (IFIH1) region. *Nature Genet.* 2006 , 38, 617–619.
- [2] International Human Genome Sequencing Consortium (2004), Finishing the euchromatic sequence of the human genome, *Nature* ,2004,Vol. 431, pp. 931–945.
- [3] Larsson, T.P., Murray, C.G., Hill, T. et al. Comparison of the current Ref Seq, Ensembl and EST databases for counting genes and gene discovery, *FEBS Lett.* 2005,Vol. 579, pp. 690–698.
- [4] Morton NE. Genetic epidemiology of hearing impairment. *Ann NY Acad Sci* 1991;630:16–31.
- [5] Marazita ML, Ploughman LM, Rawlings B, Remington E, Arnos KS. Genetic epidemiological studies of early-onset deafness in the US school-age population. *Am J Med Genet* 1993;46:486–91
- [6] Liu XZ, Xu LR, Hu Y, Nance WE, Sismanis A, Zhang SL, Xu Y. Epidemiological studies on hearing impairment with reference to genetic factors in Sichuan,China. *Ann Otol Rhinol Laryngol* 2001;110:356–63.
- [7] Van Camp G SR, Hereditary Hearing Loss Homepage <http://dnalab-www.uia.ac.be/dnalab/hhh/>. Mar,2007.
- [8] Anne B. Skvorak, Z. W., Andrew J. Yee, Nahid G. Robertson and Cynthia C. Morton, Human cochlear expressed sequence tags provide insight into cochlear gene expression and identify candidate genes for deafness. *Human Molecular Genetics*, 1999, Vol. 8, No. 3 8: 439-452.
- [9] Avraham KB. Mouse models for deafness: lessons for the human inner ear and hearing loss. *Ear Hear.* 2003 Aug;24(4):332-41.
- [10] Abe S, Katagiri T, Saito-Hisaminato A, et al., Identification of CRYM as a candidate responsible for nonsyndromic deafness, through cDNA microarray analysis of human cochlear and vestibular tissues. *Am J Hum Genet*, 2003, 72(1): 73-82.
- [11] Lynch ED, Lee MK, Morrow JE, et al., Nonsyndromic deafness DFNA1 associated with mutation of a human homolog of the *Drosophila* gene *diaphanous*. *Science*, 1997, 278(5341): 1315-1318.
- [12] Xia JH, Liu CY, Tang BS, et al., Mutations in the gene encoding gap junction

- protein beta-3 associated with autosomal dominant hearing impairment. *Nat Genet*, 1998, 20(4): 370-373.
- [13] Kubisch C, Schroeder BC, Friedrich T, et al., KCNQ4, a novel potassium channel expressed in sensory outer hair cells, is mutated in dominant deafness. *Cell*, 1999, 96(3): 437-446.
- [14] Kelsell DP, Dunlop J, Stevens HP, et al., Connexin 26 mutations in hereditary non-syndromic sensorineural deafness. *Nature*, 1997, 387(6628): 80-83.
- [15] Grifa A, Wagner CA, D'Ambrosio L, et al., Mutations in GJB6 cause nonsyndromic autosomal dominant deafness at DFNA3 locus. *Nat Genet*, 1999, 23(1): 16-18.
- [16] Donaudy F, Snoeckx R, Pfister M, et al., Nonmuscle myosin heavy-chain gene MYH14 is expressed in cochlea and mutated in patients affected by autosomal dominant hearing impairment (DFNA4). *Am J Hum Genet*, 2004, 74(4): 770-776.
- [17] Van Laer L, Huizing EH, Verstreken M, et al., Nonsyndromic hearing impairment is associated with a mutation in DFNA5. *Nat Genet*, 1998, 20(2): 194-197.
- [18] Bernalova IN, Van Camp G, Bom SJ, et al., Mutations in the Wolfram syndrome 1 gene (WFS1) are a common cause of low frequency sensorineural hearing loss. *Hum Mol Genet*, 2001, 10(22): 2501-2508.
- [19] Young TL, Ives E, Lynch E, et al., Non-syndromic progressive hearing loss DFNA38 is caused by heterozygous missense mutation in the Wolfram syndrome gene WFS1. *Hum Mol Genet*, 2001, 10(22): 2509-2514.
- [20] Verhoeven K, Van Laer L, Kirschhofer K, et al., Mutations in the human alpha-tectorin gene cause autosomal dominant non-syndromic hearing impairment. *Nat Genet*, 1998, 19(1): 60-62.
- [21] Robertson NG, Lu L, Heller S, et al., Mutations in a novel cochlear gene cause DFNA9, a human nonsyndromic deafness with vestibular dysfunction. *Nat Genet*, 1998, 20(3): 299-303.
- [22] Wayne S, Robertson NG, DeClau F, et al., Mutations in the transcriptional activator EYA4 cause late-onset deafness at the DFNA10 locus. *Hum Mol Genet*, 2001, 10(3): 195-200.
- [23] Liu XZ, Walsh J, Tamagawa Y, et al., Autosomal dominant non-syndromic deafness caused by a mutation in the myosin VIIA gene. *Nat Genet*, 1997, 17(3): 268-269.
- [24] McGuirt WT, Prasad SD, Griffith AJ, et al., Mutations in COL11A2 cause

- non-syndromic hearing loss (DFNA13). *Nat Genet*, 1999, 23(4): 413-419.
- [25] Vahava O, Morell R, Lynch ED, et al., Mutation in transcription factor POU4F3 associated with inherited progressive hearing loss in humans. *Science*, 1998, 279(5358): 1950-1954.
- [26] Lalwani AK, Goldstein JA, Kelley MJ, et al., Human nonsyndromic hereditary deafness DFNA17 is due to a mutation in nonmuscle myosin MYH9. *Am J Hum Genet*, 2000, 67(5): 1121-1128.
- [27] Zhu M, Yang T, Wei S, et al., Mutations in the gamma-Actin Gene (ACTG1) Are Associated with Dominant Progressive Deafness (DFNA20/26). *Am J Hum Genet*, 2003, 73(6): 1082-1091.
- [28] Melchionda S, Ahituv N, Bisceglia L, et al., MYO6, the human homologue of the gene responsible for deafness in Snell's waltzer mice, is mutated in autosomal dominant nonsyndromic hearing loss. *Am J Hum Genet*, 2001, 69(3): 635-640.
- [29] van Wijk E, Krieger E, Kemperman MH, et al., A mutation in the gamma actin 1 (ACTG1) gene causes autosomal dominant hearing loss (DFNA20/26). *J Med Genet*, 2003, 40(12): 879-884.
- [30] Peters LM, Anderson DW, Griffith AJ, et al., Mutation of a transcription factor, TFCP2L3, causes progressive autosomal dominant hearing loss, DFNA28. *Hum Mol Genet*, 2002, 11(23): 2877-2885.
- [31] Kurima K, Peters LM, Yang Y, et al., Dominant and recessive deafness caused by mutations of a novel gene, TMC1, required for cochlear hair-cell function. *Nat Genet*, 2002, 30(3): 277-284.
- [32] Modamio-Hoybjor S, Mencia A, Goodyear R, Del Castillo I, Richardson G, Moreno F, Moreno-Pelayo MA. A Mutation in CCDC50, a Gene Encoding an Effector of Epidermal Growth Factor-Mediated Cell Signaling, Causes Progressive Hearing Loss. *Am J Hum Genet*, 2007, 80:1076-1089.
- [33] Donaudy F, Ferrara A, Esposito L, et al., Multiple mutations of MYO1A, a cochlear-expressed gene, in sensorineural hearing loss. *Am J Hum Genet*, 2003, 72(6): 1571-1577.
- [34] Liu XZ, Ouyang XM, Xia XJ, et al., Prestin, a cochlear motor protein, is defective in non-syndromic hearing loss. *Hum Mol Genet*, 2003, 12(10): 1155-1162.
- [35] del Castillo I, Villamar M, Moreno-Pelayo MA, et al., A deletion involving the connexin 30 gene in nonsyndromic hearing impairment. *N Engl J Med*, 2002,

- 346(4): 243-249.
- [36]Liu XZ, Walsh J, Mburu P, et al., Mutations in the myosin VIIA gene cause non-syndromic recessive deafness. *Nat Genet*, 1997, 16(2): 188-190.
- [37]Wang A, Liang Y, Fridell RA, et al., Association of unconventional myosin MYO15 mutations with human nonsyndromic deafness DFNB3. *Science*, 1998, 280(5368): 1447-1451.
- [38]Li XC, Everett LA, Lalwani AK, et al., A mutation in PDS causes non-syndromic recessive deafness. *Nat Genet*, 1998, 18(3): 215-217.
- [39]Naz S, Giguere CM, Kohrman DC, et al., Mutations in a novel gene, TMIE, are associated with hearing loss linked to the DFNB6 locus. *Am J Hum Genet*, 2002, 71(3): 632-636.
- [40]Scott HS, Kudoh J, Wattenhofer M, et al., Insertion of beta-satellite repeats identifies a transmembrane protease causing both congenital and childhood onset autosomal recessive deafness. *Nat Genet*, 2001, 27(1): 59-63.
- [41]Yasunaga S, Grati M, Cohen-Salmon M, et al., A mutation in OTOF, encoding otoferlin, a FER-1-like protein, causes DFNB9, a nonsyndromic form of deafness. *Nat Genet*, 1999, 21(4): 363-369.
- [42]Bork JM, Peters LM, Riazuddin S, et al., Usher syndrome 1D and nonsyndromic autosomal recessive deafness DFNB12 are caused by allelic mutations of the novel cadherin-like gene CDH23. *Am J Hum Genet*, 2001, 68(1): 26-37.
- [43]Verpy E, Masmoudi S, Zwaenepoel I, et al., Mutations in a new gene encoding a protein of the hair bundle cause non-syndromic deafness at the DFNB16 locus. *Nat Genet*, 2001, 29(3): 345-349.
- [44]Ouyang XM, Xia XJ, Verpy E, et al., Mutations in the alternatively spliced exons of USH1C cause non-syndromic recessive deafness. *Hum Genet*, 2002, 111(1): 26-30.
- [45]Ahmed ZM, Smith TN, Riazuddin S, et al., Nonsyndromic recessive deafness DFNB18 and Usher syndrome type IC are allelic mutations of USH1C. *Hum Genet*, 2002, 110(6): 527-531.
- [46]Mustapha M, Weil D, Chardenoux S, et al., An alpha-tectorin gene defect causes a newly identified autosomal recessive form of sensorineural pre-lingual non-syndromic deafness, DFNB21. *Hum Mol Genet*, 1999, 8(3): 409-412.
- [47]Zwaenepoel I, Mustapha M, Leibovici M, et al., Otoancorin, an inner ear protein restricted to the interface between the apical surface of sensory epithelia and their

- overlying acellular gels, is defective in autosomal recessive deafness DFNB22. *Proc Natl Acad Sci USA*, 2002, 99(9): 6240-6245.
- [48] Ahmed ZM, Riazuddin S, Ahmad J, et al., PCDH15 is expressed in the neurosensory epithelium of the eye and ear and mutant alleles are responsible for both USH1F and DFNB23. *Hum Mol Genet*, 2003, 12(24): 3215-3223.
- [49] Khan SY, Ahmed ZM, Shabbir MI, Kitajiri SI, Kalsoom S, Tasneem S, Shayiq S, Ramesh A, Srisailpathy S, Khan SN, Smith RJ, Riazuddin S, Friedman TB (2007) Mutations of the RDX gene cause nonsyndromic hearing loss at the DFNB24 locus. *Hum Mutat*. 28(5):417-423
- [50] Riazuddin S., Khan S. N., Ahmed Z. M., Ghosh M., Caution K., Nazli S., Kabra M., Zafar A. U., Chen K., Naz S., Antonellis A., Pavan W. J., Green E. D., Wilcox E. R., Friedman P. L., Morell R. J., and Friedman T. B. (2006). Mutations in TRIOBP, which encodes a putative cytoskeletal-organizing protein, are associated with nonsyndromic recessive deafness. *Am J Hum Genet* 78: 137-43.
- [51] Shahin H., Walsh T., Sobe T., Abu Sa'ed J., Abu Rayan A., Lynch E. D., Lee M. K., Avraham K. B., King M. C., and Kanaan M. (2006). Mutations in a novel isoform of TRIOBP that encodes a filamentous-actin binding protein are responsible for DFNB28 recessive nonsyndromic hearing loss. *Am J Hum Genet* 78: 144-52.
- [52] Wilcox ER, Burton QL, Naz S, et al., Mutations in the gene encoding tight junction claudin-14 cause autosomal recessive deafness DFNB29. *Cell*, 2001, 104(1): 165-172.
- [53] Walsh T, Walsh V, Vreugde S, et al., From flies' eyes to our ears: mutations in a human class III myosin cause progressive nonsyndromic hearing loss DFNB30. *Proc Natl Acad Sci USA*, 2002, 99(11): 7518-7523.
- [54] Mburu P, Mustapha M, Varela A, et al., Defects in whirlin, a PDZ domain molecule involved in stereocilia elongation, cause deafness in the whirler mouse and families with DFNB31. *Nat Genet*, 2003, 34(4): 421-428.
- [55] Naz S, Griffith AJ, Riazuddin S, et al., Mutations of ESPN cause autosomal recessive deafness and vestibular dysfunction. *J Med Genet*, 2004, 41(8): 591-595.
- [56] Ahmed ZM, Morell RJ, Riazuddin S, et al., Mutations of MYO6 are associated with recessive deafness, DFNB37. *Am J Hum Genet*, 2003, 72(5): 1315-1322.
- [57] Delmaghani S, del Castillo FJ, Michel V, Leibovici M, Aghaie A, Ron U, Van Laer L, Ben-Tal N, Van Camp G, Weil D, Langa F, Lathrop M, Avan P, Petit C.

- Mutations in the gene encoding pejvakin, a newly identified protein of the afferent auditory pathway, cause DFNB59 auditory neuropathy. *Nat Genet.* 2006 Jul;38(7):770-8.
- [58] Shabbir, M.I., Ahmed, Z.M., Khan, S.Y., Riazuddin, S., Waryah, A.M., Khan, S.N., Camps, R.D., Ghosh, M., Kabra, M., Belyantseva, I.A. and Friedman, T.B. (2006). Mutations of human TMHS cause recessively inherited nonsyndromic hearing loss. *J Med Genet* 43: 634-40.
- [72] Anne B. Skvorak, Z. W., Andrew J. Yee, Nahid G. Robertson and Cynthia C. Morton, Human cochlear expressed sequence tags provide insight into cochlear gene expression and identify candidate genes for deafness. *Human Molecular Genetics*, 1999, Vol. 8, No. 3 8: 439-452.
- [59] Luijendijk Mw, V. D. P. T., Van Duijnhoven G, Den Hollander Ai, Ten Caat J, Van Limpt V, Brunner Hg, Kremer H, Cremers Fp. Cloning, characterization, and mRNA expression analysis of novel human fetal cochlear cDNAs. *Genomics* , 2003, 82: 480-490.
- [60] Pompeia C, H. B., Belyantseva ia, Noben-Trauth K, Beisel K, Gao J, Buchoff P, Wistow G, Kachar B, 2004 Gene expression profile of the mouse organ of Corti at the onset of hearing. *Genomics* 83: 1000-1011.
- [61] Resendes BL, R. N., Szustakowski JD, Resendes RJ, Weng Z, Morton CC. , 2002 Gene Discovery in the Auditory System: Characterization of Additional Cochlear-Expressed Sequences. *J Assoc Res Otolaryngol.* 3: 45-53.
- [62] HOLME RH, B. T., STEEL KP, 2003 Table of gene expression in the developing ear. [<http://www.ihr.mrc.ac.uk/Hereditary/genetable>].
- [63] Pompeia C, H. B., Belyantseva ia, Noben-Trauth K, Beisel K, Gao J, Buchoff P, Wistow G, Kachar B, 2004 Gene expression profile of the mouse organ of Corti at the onset of hearing. *Genomics* 83: 1000-1011.
- [64] Ahituv N, Avraham KB. Mouse models for human deafness: current tools for new fashions. *Trends Mol Med.* 2002;8(9):447-51.
- [65] Hamosh, A., Scott, A.F., Amberger, J., Valle, D. and McKusick, V.A. Online Mendelian Inheritance in Man (OMIM). *Hum. Mutat.* 2000, 15, 57-61.
- [66] Pruitt, K.D., Katz, K.S., Sicotte, H. and Maglott, D.R. Introducing RefSeq and LocusLink: curated human genome resources at the NCBI. *Trends Genet.* 2000, 16, 44-47.
- [67] Krawczak, M., Ball, E.V., Fenton, I., Stenson, P.D., Abeyasinghe, S., Thomas, N. and

- Cooper,D.N. Human gene mutation database-a biomedical information and research resource. *Hum. Mutat.* 2000, 15, 45-51.
- [68] Schuler, G.D., Epstein, J.A., Ohkawa, H. et al. Entrez: Molecular biology database and retrieval system, *Methods Enzymol.* 1996, Vol. 266, pp. 141–162.
- [69] Etzold, T. and Argos, P. SRS — An indexing and retrieval tool for flat file data libraries, *Comput. Appl. Biosci.* 1993, Vol. 9, pp. 49–57.
- [70] Hekkelman, M.L. and Vriend, G. MRS: A fast and compact retrieval system for biological data, *Nucleic Acids Res.* 2005, Vol. 33, pp. W766–W769.
- [71] Kent, W.J., Sugnet, C.W., Furey, T.S. et al. The human genome browser at UCSC, *Genome Res.* 2002, Vol. 12, pp. 996–1006.
- [72] Hubbard, T., Andrews, D., Caccamo, M. et al. Ensembl 2005, *Nucleic Acids Res.* 2005, Vol. 33, pp. D447–D453.
- [73] Ashburner, M., Ball, C.A., Blake, J.A. et al. Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium, *Nat. Genet.* 2000, Vol. 25, pp. 25–29.
- [74] Kasprzyk, A., Keefe, D., Smedley, D. et al. EnsMart: A generic system for fast and flexible access to biological data, *Genome Res.* 2004, Vol. 14, pp. 160–169.
- [75] Perez-Iratxeta, C., Wjst, M., Bork, P. and Andrade, M.A. G2D: a tool for mining genes associated with disease. *BMC Genet.* 2005, 6, 45.
- [76] Perez-Iratxeta, C., Bork, P. and Andrade, M.A. Association of genes to genetically inherited diseases using data mining, *Nat. Genet.* 2002, Vol. 31, pp. 316–319.
- [77] Wheeler, D.L., Barrett, T., Benson, D.A. et al. Database resources of the National Center for Biotechnology Information, *Nucleic Acids Res.* 2005, Vol. 33, pp. D39–D45.
- [78] Turner, F.S., Clutterbuck, D.R. and Semple, C.A. POCUS: Mining genomic sequence annotation to predict disease genes, *Genome Biol.* 2003, Vol. 4, p. R75.
- [79] Karlin, S., Brocchieri, L. and Bergman, A. (2002), ‘Amino acid runs in eukaryotic proteomes and disease associations’, *Proc. Natl. Acad. Sci. USA* Vol. 99, pp. 333–338.
- [80] Lopez-Bigas, N. and Ouzounis, C.A. Genome-wide identification of genes likely to be involved in human genetic disease, *Nucleic Acids Res.* 2004, Vol. 32, pp. 3108–3114.
- [81] Adie, E.A., Adams, R.R. and Evans, K.L. Speeding disease gene discovery by

- sequence based candidate prioritization, BMC Bioinformatics, 2005, Vol. 6, p. 55.
- [82] Smith, N.G. and Eyre-Walker, A. Human disease genes: Patterns and predictions, Gene, 2003, Vol. 318, pp. 169–175.
- [83] Kelso, J., Visagie, J., Theiler, G. et al. eVOC: A controlled vocabulary for unifying gene expression data, Genome Res. 2003, Vol. 13, pp. 1222–1230.
- [84] Tiffin, N., Kelso, J.F., Powell, A.R. et al. Integration of text and data-mining using ontologies successfully selects disease gene candidates, Nucleic Acids Res. 2005, Vol. 33, pp. 1544–1552.
- [85] van Driel, M.A., Cuelenaere, K., Kemmeren, P.P. et al. A new web-based data mining tool for the identification of candidate genes for human genetic disorders, Eur. J. Hum. Genet. 2003, Vol. 11, pp. 57–63.
- [86] van Driel, M.A., Cuelenaere, K., Kemmeren, P.P. et al. GeneSeeker: Extraction and integration of human disease-related information from web-based genetic databases, Nucleic Acids Res. 2005, Vol. 33, pp. W758–W761.
- [87] Masseroli, M., Galati, O. and Pincioli, F. GFINDER: Genetic disease and phenotype location statistical analysis and mining of dynamically annotated gene lists, Nucleic Acids Res. 2005, Vol. 33, pp. W717–W723.
- [88] van Driel, M.A., Bruggeman, J., Vriend, G. et al. A text-mining analysis of the human phenome, Eur. J. Hum. Genet. 2006, Vol. 14, pp. 535–542.
- [89] Brunner, H.G. and van Driel, M.A. From syndrome families to functional genomics, Nat. Rev. Genet. 2004, Vol. 5, pp. 545–551.
- [90] Rossi, S., Masotti, D., Nardini, C., Bonora, E., Romeo, G. et al. TOM: a web-based integrated approach for identification of candidate disease genes, Nucleic Acids Res. 2006, Vol. 34, pp. W285–W292.
- [91] George, R.A., Liu, J.Y., Feng, L.L., Robert J. Bryson-Richardson. et al. Analysis of protein sequence and interaction data for candidate disease gene prediction, Nucleic Acids Res. 2006, Vol. 34, pp. e130.
- [92] Freimer, N. and Sabatti, C. The human phenome project, Nat. Genet. 2003, Vol. 34, pp. 15–21.
- [93] Biesecker, L.G. Mapping phenotypes to language: A proposal to organize and standardize the clinical descriptions of malformations, Clin. Genet. 2005, Vol. 68, pp. 320–326.

附录 1:

DFNA52 候选基因列表(UniGene)

	Genes	Description	OMIM	UniGene	EST counts of Cochlea in Expression profile
1	ANKHD1	ankyrin repeat and KH domain containing 1		Hs.434219	362 广泛
2	ARHGAP26	Rho GTPase activating protein 26	605370	Hs.293593	0
3	BRD8	bromodomain containing 8	602848	Hs.519337	241 广泛
4	CDC23	CDC23 (cell division cycle 23, yeast, homolog)	603462	Hs.153546	120 广泛表达
5	CTNNA1	catenin (cadherin-associated protein)	116805	Hs.534797	241 广泛
6	EGR1	early growth response 1	128990	Hs.326035 Hs.645774	302 广泛/181
7	ETF1	eukaryotic translation termination factor 1	600285	Hs.483494	0
8	FGF1	fibroblast growth factor 1 (acidic)	131220	Hs.483635	0
9	GFRA3	GDNF family receptor alpha 3	605710	Hs.58042	181
10	GNPDA1	glucosamine-6-phosphate deaminase 1	601798	Hs.633853	181 广泛
11	H2AFY	H2A histone family, member Y	610054	Hs.420272 Hs.599225	60/0
12	HBEGF	heparin-binding EGF-like growth factor	126150	Hs.799	0
13	HDAC3	histone deacetylase 3	605166	Hs.519632	60 广泛
14	HNRPA0	heterogeneous nuclear ribonucleoprotein A0	609409	Hs.96996	302 广泛
15	HSPA9B	heat shock 70kDa protein 9B (mortalin-2)	600548	Hs.184233	181 广泛
16	IK	IK cytokine, down-regulator of HLA II	600549	Hs.421245	0
17	JMJD1B	jumonji domain containing 1B	609373	Hs.483486	60 广泛
18	KCTD16	potassium channel tetramerisation domain containing 16		Hs.161256	0
19	KIF20A	kinesin family member 20A	605664	Hs.73625	0
20	KLHL3	kelch-like 3 (Drosophila)	605775	Hs.96996	302 广泛
21	LARS	leucyl-tRNA synthetase	151350	Hs.432674	181 广泛
22	LOC492311	similar to bovine IgA regulatory protein		Hs.200938	60
23	LRRTM2	leucine rich repeat transmembrane neuronal 2		Hs.445981	0
24	MATR3	matrin 3	604706	Hs.268939	544 广泛
25	Myot	myotilin	604103	Hs.84665	0
26	NDFIP1	Nedd4 family interacting protein 1		Hs.9788	362 广泛
27	NEUROG1	neurogenin 1	601726	Hs.248149	0
28	NR3C1	nuclear receptor subfamily 3, group C, member 1	138040	Hs.122926	181 广泛
29	NRG2	neuregulin 2	603818	Hs.408515	0
30	PACAP	proapoptotic caspase adaptor protein	609447	Hs.409563	0
31	PAIP2	poly(A) binding protein interacting	605604	Hs.396644	181 广泛

		protein 2			
32	PCBD2	pterin-4 alpha-carbinolamine dehydratase/dimerization cofactor of hepatocyte nuclear factor 1 alpha	609836	Hs.527213	0
33	PCDH1	protocadherin 1 (cadherin-like 1)	603626	Hs.79769	0
34	PCDHAC1	protocadherin alpha subfamily C, 1	606320	Hs.199343	0
35	PCDHAC2	protocadherin alpha subfamily C, 2	606321	Hs.199343	0
36	PCDHB3	protocadherin beta 3	606329	Hs.674730	0
37	PCDHB5	protocadherin beta 5	606331	Hs.119693	120 表达量最高
38	PCDHB11	protocadherin beta 11	606337	Hs.283084	120 最强
39	PCDHGC3	protocadherin gamma subfamily C, 3	603627	Hs.368160	60 广泛
40	PITX1	paired-like homeodomain transcription factor 1	602149	Hs.84136	0
41	PPP2R2B	protein phosphatase 2 (formerly 2A), regulatory subunit B	604325	Hs.193825	423 广泛
42	SLC23A1	solute carrier family 23 (nucleobase transporters), member 1	603790	Hs.643467	0
43	SMAD5	SMAD, mothers against DPP homolog 5	603110	Hs.167700	120 广泛表达
44	SPOCK1	sparc/osteonectin, cwcv and kazal-like domains proteoglycan (testican) 1	602264	Hs.124611	362 广泛
45	TAF7	TAF7 RNA polymerase II, TATA box binding protein (TBP)-associated factor	600573	Hs.438838	0
46	TCERG1	transcription elongation regulator 1	605409	Hs.443465	241 广泛
47	TGFB1	transforming growth factor, beta-induced	601692	Hs.369397	362 广泛
48	TRPC7	transient receptor potential cation channel, subfamily C, member 7		Hs.591263	0
49	UBE2D2	ubiquitin-conjugating enzyme E2D 2	602962	Hs.108332	120
50	WNT8A	wingless-type MMTV integration site family, member 8A	606360	Hs.591274	\
51	YIPF5	Yip1 domain family, member 5		Hs.372050	60 广泛
52	ZMAT2	zinc finger, matrin type 2		Hs.350194	120 广泛

附录 2:

部分检测基因的外显子个数、引物名称和各引物序列

Genes	Exons	引物名称	引物序列	Tm	bp
Neurog1	exon1	neurog1-1F	ATTgTTgCgCCgggTACTTAA	62.3	
		neurog1-1R	ggCTgCCTgTTggAgTCTgg	62.7	434
	Exon2	neurog1-2F	AgCAGCggCAGTgACCTATC	60.2	
		neurog1-2R	gCCAgAgCCCAGATgTAGTTgT	60.8	386
	Exon3	neurog1-3F	CgACgACACCAAgCTCACC	59.2	
		neurog1-3R	CCCATCTATTgCCTgCTgACT	59.3	402
SMAD5	exon1	SMAD5-1F	ATCggAAGTCggCgTTggC	65.4	
		SMAD5-1R	gggTCgTTTCCgCTCCgTC	64.8	396
	Exon2	SMAD5-2F	TATTACCATATCAAaggAg	45.0	
		SMAD5-2R	gTTTTCCCAgAATTTTAgTT	48.6	331
	Exon3	SMAD5-3F	AggCACAAGAAAaggAAgC	52.3	
		SMAD5-3R	TgATgATgACAgTggAAgT	52.5	297
	Exon4	SMAD5-4AF	ATTgCCAgTgTCTTACTTgTTC	54.0	
		SMAD5-4AR	CACATTTACTCggCTgTCC	53.4	567
		SMAD5-4BF	ggCAGTTgATgCTTTggT	53.5	
		SMAD5-4BR	AATCCAATAgCCCTTCCA	52.5	454
	Exon5	SMAD5-5F	gAATCCTTTCTACCAACCCC	55.3	
		SMAD5-5R	ggCATgTAACAgTTCACCCA	56.3	430
	Exon6	SMAD5-6F	gTTCTCCTTTTCCCTTgC	51.7	
		SMAD5-6R	CTgTggTgCAAAACTTgg	52.1	447
	Exon7	SMAD5-7F	gggTTgggTAAAAgATA	47.2	
		SMAD5-7R	TgAgggAACATCATgTCTT	49.7	478
	Exon8	SMAD5-8F	CTgTCTAAAgACTgCTggAA	50.2	
		SMAD5-8R	TTAAATCAggCCAAAAGC	51.3	496
	Exon9	SMAD5-9F	TTggTTTAgAgTggTTgA	45.4	
		SMAD5-9R	CAATgTAAgCTCACAgTATC	44.9	293
PCDHGC3	exon1	PCDHGC3a-1F	gCTgTCggCCAATgAAgACg	63.7	
		PCDHGC3a -1R	CAGACACCACCCggAACCTg	63.1	431
		PCDHGC3a -2F	AgAAgggTTTCgCTgTgggC	63.8	

		PCDHGC3a -2R	gTCCAACgCCgTCAgCACTA	61.8	530
		PCDHGC3a -3F	gCggAgCTggTgTTggAgC	63.6	
		PCDHGC3a -3R	gAgCAAAgCgATgACAgTCCC		543
		PCDHGC3a -4F	CCAATCCCgAAggAgCACA	61.4	
		PCDHGC3a -4R	CCCggTTTCAGCTCCTTg		527
		PCDHGC3a -5F	AATgCTCggCTTCTTTCTTTC	60.0	
		PCDHGC3a -5R	AgTCTgCCTgggTgAATCTgT		465
		PCDHGC3a -6F	gggTggTAggCTgggACg	60.6	
		PCDHGC3a -6R	ggTAAAggTgCggCgACAT	60.7	498
		PCDHGC3a -7F	TCTAATCCTggTTTCTgTggg	56.8	
		PCDHGC3a -7R	CCCTgAAATgggACCACA	56.4	545
	Exon2	PCDHGC3-2F	ggggCTCgggTgACATTC	60.4	
		PCDHGC3-2R	gggCAGTCgCCTCTTCCTT	61.5	307
	Exon3	PCDHGC3-3F	AAgTgCCTCCTACCTTgCTgA	59.1	
		PCDHGC3-3R	ggCATggTggCTgTgAgAT	58.4	309
	Exon4	PCDHGC3-4F	TgCTCAGCgTggTCAAagg	58.1	
		PCDHGC3-4R	gACTgggCACTTgTTTCTgTTT	58.6	506
SPOCK1	exon1	SPOCK-1F	AgATgTgAATgACCACgTTCC	59.8	
		SPOCK-1R	ggAggTCCCAACTACTCCAAG	59.9	345
	exon2	SPOCK-2AF	AgCCTCCTCgTTCgTAgATg	59.4	
		SPOCK-2AR	TACTTgTCCCggTCgTACTg	58.6	371
		SPOCK-2BF	TgCTTCCTCCAAgTCgAgAg	60.6	
		SPOCK-2BR	AAgTTCCTTTCAggCAAgACC	59.7	333
	exon3	SPOCK-3F	CggTTTgAAAgACAggAATTTg	60.8	
		SPOCK-3R	AACTCAGCATgCAAggAAgg	60.4	318
	exon4	SPOCK-4F	ATggTggCgTgCACCTTAg	54.9	
		SPOCK-4R	ACgTCAAAGACTggAAACCCTC	53.8	416
	exon5	SPOCK-5F	CTTggCCTggTTggATAgTC	59.5	
		SPOCK-5R	AATgCAGCAGCTTCTCAAgg	60.6	373
	exon6	SPOCK-6F	TgAAgCAAAgggCACTgCTg	56.2	
		SPOCK-6R	CCAgTTTgTTCCTAgTgCCTgg	55.3	304

	exon7	SPOCK-7F	TgCACACAgAAGgCAGAAAg	60.1	
		SPOCK-7R	AAAgTATgTACCACAgCATgCTTAC	58.7	406
	exon8	SPOCK-8F	TCaggAAgggAgTCaggTgg	53.7	
		SPOCK-8R	TgTggCTCTCTgCCCAATTC	54.9	432
	exon9	SPOCK-9F	CACTCCCAGgTgTgTCATTg	60.0	
		SPOCK-9R	gCCCTAATTgAATAACCATTgC	59.7	264
	exon10	SPOCK-10F	TgTgAgCCCTCTTCCCTgTg	54.9	
		SPOCK-10R	TggAgTTACCTCTCAACATTCTTCTg	54.9	326
	exon11	SPOCK11-1F	CCCTACAATTCAggCAAAGg	59.5	
		SPOCK11-1R	TCgggTATAgAgAgCAACAATg	59.2	397
		SPOCK11-2F	gCCCACAAGAAAgAggACAC	59.7	
		SPOCK11-2R	gggAACCATgCTgTAATAACC	58.3	394
		SPOCK11-3F	gCCTCCTTCCTTTCTgTgAg	59.0	
		SPOCK11-3R	CAGggAAGTATCCAgACACC	57.0	380
		SPOCK11-4F	TCTggCTTCTATAAACAACAgC	59.8	
		SPOCK11-4R	TgATTTAggCAGACggAggT	59.6	387
		SPOCK11-5F	TgCACTAACAgTTggCTCTTg	59.1	
		SPOCK11-5R	TgCCCTTATCATgATTTgTCAg	59.9	384
		SPOCK11-6F	gCAATCCTgTCTgCCATAgg	60.6	
		SPOCK11-6R	ggCCAgCTTgATAgggAgAg	61.2	446
		SPOCK11-7F	CATCgTAATAACACATggAAgC	57.2	
		SPOCK11-7R	TTTgTgTggACAggCTgAAG	59.8	394
		SPOCK11-8F	CAGggTgTCCTAATCAGAAATCAG	60.0	
		SPOCK11-8R	TTTgCCTATggTCTgTCTTCC	59.1	375
		SPOCK11-9F	AgTCggCATTgCCTgAATAC	60.1	
		SPOCK11-9R	TCAgTggTTCACCCTgTTCC	60.9	425
		SPOCK11-10F	CCATTgTTACTATCTCCTTTgAgC	58.4	
		SPOCK11-10R	AATgTggCTCCATCgATTTg	60.8	330
		SPOCK11-11F	TggAgAAGAgTgCTCAATAgACAg	60.0	
		SPOCK11-11R	TggAggACTATTTCCAAGAAgAAg	60.1	295
		SPOCK11-12F	gCCTCACTTTgACATgACAgC	60.8	
		SPOCK11-12R	gggATTTAAGTTggATACCAAGg	59.9	358

		SPOCK11-13F	CCACTTCCTCATTTgATACAgg	58.5	
		SPOCK11-13R	gggCgTAAAggAggAgAAAC	60.0	265
		SPOCK11-14F	TCCACCATggTTAgTTATTTgC	58.8	
		SPOCK11-14R	TCAggATTTgATgggCTCTC	60.1	355
		SPOCK11-15F	CCACgCTgTACAgTAACATCC	58.1	
		SPOCK11-15R	gACATggCCTgCTgAATg	58.6	375
PCDHB5	exon1	PCDHB5-1F	gTAAATTAggACgCCTggTg	57.2	
		PCDHB5-1R	TggCCACAgAATAgCCACTT	60.6	400
		PCDHB5-2F	ggTCTgAACAATggAgACTgC	59.7	
		PCDHB5-2R	ggCATggTCATTTATATCTgTgAg	59.7	400
		PCDHB5-3F	ggCgACAgAACCCCTgTATATTg	60.7	
		PCDHB5-3R	AATgCgAATTgTggTggTC	59.3	400
		PCDHB5-4F	AATACCCAgAgCTggTgCTg	60.2	
		PCDHB5-4R	CCATCTgTggCTACAATTTCC	59.4	400
		PCDHB5-5F	TgCTCTATTCCAaggCgATg	61.2	
		PCDHB5- 5R	CggTgACAgTgATggTgATg	61.0	434
		PCDHB5- 6F	ggATgATTTgCTCCATCCAg	60.4	
		PCDHB5- 6R	TCgTAgtCCAgtCgACCTgAg	61.5	427
		PCDHB5- 7F	gACAgAgACTCaggCACCAAC	59.9	
		PCDHB5- 7R	TTgAgCagCTggTACgACAg	60.2	395
		PCDHB5- 8F	CTgTATCCgCTgCagAACg	60.5	
		PCDHB5- 8R	AAGaggAAGgCgACgACAC	59.6	425
		PCDHB5- 9F	CTTCTCCCAgCCCTACCTg	59.8	
		PCDHB5- 9R	CTATTCCggAaggCagCag	60.8	381
		PCDHB5- 10F	AggTgTgTTTgACCggAgAC	60.0	
		PCDHB5- 10R	ACAgCATgCATggATTTgg	60.5	357
		PCDHB5- 11F	CTTCATgCTgTCCAAGtTTg	59.3	
		PCDHB5- 11R	gAgTAATTTgTgTggCACCTg	58.1	547
TTID	exon1	TTID-1F	CATgTAATCCCACTCCTTAA	52.7	
		TTID-1R	ATgCCTTCCAgTCTCAAT	52.8	435

	exon2	TTID-2AF	CAATgTTATTgTTTAAgC	43.6	
		TTID-2AR	AgAAAATCTTTgCTCTgT	46.9	421
		TTID-2BF	gCCTCCTggACCAgAAAC	58.1	
		TTID-2BR	CCTCTTATTgCCAgTACCCT	55.9	473
	exon3	TTID-3F	gACTgTTATTAgggCAAT	47.2	
		TTID-3R	CTgTgAgCTAgAACTTCC	47.6	386
	exon4	TTID-4F	TgCTTCTTTgAAgTTCTg	49.4	
		TTID-4R	AAgCATTTAgAAAAGCag	48.2	273
	exon5	TTID-5F	TTCTTgCTAgAgTggTAg	45.6	
		TTID-5R	AAAAGgTATAgAgTAggC	43.2	342
	exon6	TTID-6F	gCAAAATgAAAAGCAATg	52.7	
		TTID-6R	gCTATgAAgTAggAAAAGa	47.0	
	exon7	TTID-7F	gCTCTTgCTTTgCATTTc	54.6	
		TTID-7R	AgTCTgCTgggCTTTTCT	55.1	426
	exon8	TTID-8F	TTCTTTCTgTTgCTgTTg	50.3	
		TTID-8R	TTAgTggAATCTggCTTA	48.9	476
	exon9	TTID-9F	TggCAATAAATAgATACAg	46.1	
		TTID-9R	TTTATACTCTgCTgggAT	47.7	346
	exon10	TTID-10AF	CAATTTggTTAgAACAgg	48.6	
		TTID-10AR	gCAAgAgTCAAaggATAA	48.3	384
		TTID-10BF	ACAgATTATggTTTTAATTAgg	50.0	
		TTID-10BR	ggTTTTgCTgAgTggAgT	53.4	435

致 谢

衷心感谢夏家辉院士！夏院士实事求是的科研作风，锲而不舍、追求真理、勇于创新的探索精神以及他那博学、敏锐、无私、无畏的科学家风范无时无刻不在激励我、鼓舞我。他的言传身教不仅将我带入了一个科学探索的殿堂，而且使我树立起一个献身科学事业的目标。这一切在我的心里留下了不可磨灭、深刻的印记，将永远激励和鞭策我在今后的科研道路上奋进。

衷心感谢陈凤琼老师对我生活的关心！

衷心感谢张灼华教授！他严谨缜密的科学思维、精益求精的科研态度无不凸现着深厚的学者素养，在给我以莫大启迪的同时也让我感受到最深切的触动。

衷心感谢夏昆主任！一直以来，他渊博的学识，宽广的胸襟，朋友般的鼓励和帮助都使我受益良多。感谢他的悉心指导和不断鼓励，感谢他在我成长路上给予我亦师亦友般的关怀。

衷心感谢中南大学湘雅医院耳鼻喉科冯永教授，胡杰对我工作的大力支持！

衷心感谢实验室梁德生主任、邬玲仟主任、潘乾主任、蔡芳主任、戴和平主任、龙志高主任、苏薇、刘春宇副教授在本课题进行中给与的大力支持和热心帮助！

衷心感谢生物信息学组卜枫啸、朱赞华、张瑞芳、姚凤霞、宋宁；基因克隆组胡正茂、谢志国、李薇、徐伟、莫晓云、李乾、龚惠勇、张静、滕祥云、胡艺俏！在这个团结协作的集体中，不仅对我的课题给与莫大的帮助，而且使整个课题组始终保持和谐的气氛和良好的风气，使我受益匪浅。

衷心感谢诊断组的朱海燕、胡浩、吴伟锋、胡兰萍、贾蓓；功能组的刘慕君、刘小平、杨眉、张宗磊、凌捷；治疗组的刘雄昊、薛志刚、张雅坤、王丽娜等各位同学的热情帮助！

衷心感谢陈湘、文娟、李娟、黄尚国、黄赛娟等实验室工作人员对我的热情帮助！

衷心感谢全室所有的同事！

最后，要深深的感谢我的妻子、父母和岳父母，感谢他们多年来从经济上和感情上给予了大力支持和无私帮助。他们的关爱和期待将永远是我前进的动力！

衷心感谢所有帮助我成长的人！

攻读博士学位期间主要的研究成果

一. 参与课题情况

1. 参与国家自然科学基金:精神疾病家系资源库的建设与数据库管理, 2005-2009。项目批准号:30671153
2. 参与国家自然科学基金项目: DFNA42 疾病基因的克隆, 2004-2006。项目批准号: 30371530

二. 发表论文:

1. Pan Qiong, Hu Zheng-mao, Yong Feng, Xia K et. Bioinformatics analysis candidate genes and mutation detection of 52 Genes for DFNA52 Locus. *Journal of Laryngology and Otology* (已接收)
2. Pan Qiong, Hu Zheng-mao, Xie Zhi-guo, Wu Ling-qian et al. A Ichthyosis vulgaris phenotype of Loricrin keratoderma caused by 730insG mutation in Loricrin gene. *Journal of Dermatological Science* (审稿中)