



高等医药院校规划教材

分子医学遗传学

Fenzi Yixue Yichuanxue

◎ 主编 韩骅 舒青 张萍



第四军医大学出版社

分子医学遗传学

主 编 韩 骅 舒 青 张 萍

编 者 (按姓氏笔画排序)

王淑红 王耀春 孔 辉

吴强强 张 萍 胡兴斌

秦鸿雁 梁 亮 韩 骅

舒 青 雷晓英 颜 真

魏亚宁

第四军医大学出版社·西安

图书在版编目(CIP)数据

分子医学遗传学/韩骅,舒青,张萍主编. —西安:第四军医大学出版社,2009.1
ISBN 978-7-81086-492-3

I. 分… II. ①韩… ②舒… ③张… III. 医学-分子生物学-医学遗传学
IV. Q7-R394

中国版本图书馆 CIP 数据核字(2009)第 008092 号

分子医学遗传学

主 编 韩 骅 舒 青 张 萍

责任编辑 徐文丽 土丽艳

责任校对 黄 璐

出版发行 第四军医大学出版社

地 址 西安市长乐西路 17 号(邮编:710032)

电 话 029-84776765

传 真 029-84776764

网 址 <http://press.fmmu.sn.cn>

印 刷 陕西龙源包装印刷材料有限公司

版 次 2009 年 1 月第 1 版 2009 年 1 月第 1 次印刷

开 本 787×1092 1/16

印 张 18

字 数 405 千字

书 号 ISBN 978-7-81086-492-3/Q·21

定 价 36.00 元

(版权所有 盗版必究)

前 言

随着医学研究向分子水平深入,作为医学领域里最为活跃的学科之一,医学遗传学也已进入分子时代,可称为分子医学遗传学。它运用分子生物学技术,从DNA水平、RNA水平及蛋白质水平对遗传性疾病或疾病的遗传因素进行研究,揭示基因突变与疾病发生的关系,建立在分子水平上对遗传性疾病等疾病的诊断方法,进一步实现对遗传病等疾病的基因治疗,达到从根本上治愈遗传病的目的。由于先进的分子生物学技术不断涌现,新的研究成果层出不穷,尤其是在人类基因组计划完成之后,整个医学遗传学的面貌已经发生了深刻的变化,医学遗传学也被赋予了新的使命。它不但要从公共卫生的层面,识别疾病的易感人群,从而采取教育和控制措施以预防疾病的发生,更要在临床医学的层面,为个体化的疾病治疗提供科学的依据。因此,医学遗传学的未来,将不仅是一种理论的完善和知识的传授,更是基于现代生物学和生物技术的医学服务的实践。在此背景下,为了反映学科的最新发展动态,我们结合本教研室的科研实践,编写此书供大家参考和使用。

本书是我教研室成员通过参阅大量国内外先进教材、专著和文献编写而成的。在内容的编排上,在涵盖医学遗传学的经典内容的基础上,注重介绍分子医学遗传学的最新发展动态和研究成果,以求将经典医学遗传学和现代分子遗传学有机地融合在一起。全书的编写力求具备两个特点,一是密切联系遗传医学临床实践,二是简明扼要、适合教学。全书共分为十四章,包括绪论、遗传的物质基础、遗传信息的传递、遗传性状的表达、遗传性状的控制、变异的分子基础、DNA损伤与修复系统、染色体和染色体病、单基因遗传病、多基因遗传病、线粒体遗传病、药物遗传学与药物基因组学、肿瘤遗传学、基因诊断和基因治疗等章节。为了便于理解和复习,每章后均附有思考题,全书的末尾列出了主要的参考文献,可供进一步的查阅。本书面向全日制医学院校的本科生,可作为医学遗传学课程的教材或者参考书,也可供医学和生物学各专业的教师和科研

工作者参考。

最后,我对参编本书的各位编者的辛勤劳动表示衷心地感谢。他们在繁重的科研工作间隙积极投身此书的编写,在著述和修改过程中表现出高度的负责态度。我也感谢第四军医大学出版社的土丽艳编辑和黄璐编辑在审校此书过程中所表现出的专业精神,没有他们的共同努力,就没有此书的付梓。然而,由于我们的水平有限,书中难免有不妥之处,衷心希望使用本教材的师生们能够不吝指正,以便进一步地改进和提高。

韩 骅

2008 年 12 月 26 日

目 录

第一章 绪论	(1)
第二章 遗传的物质基础——基因组	(11)
第一节 核酸是遗传的物质基础	(11)
第二节 遗传的基本单位——基因	(12)
第三节 基因组概述	(15)
第四节 功能基因组学	(21)
第三章 遗传信息的传递	(27)
第一节 DNA 复制	(27)
第二节 细胞分裂	(44)
第三节 遗传的基本规律	(49)
第四章 遗传性状的表达——真核生物基因的转录和翻译	(53)
第一节 真核生物基因的转录	(53)
第二节 真核生物基因的翻译	(64)
第五章 遗传性状的控制——真核基因的表达调控	(72)
第一节 真核基因表达的表观遗传学调控	(72)
第二节 转录水平的基因表达调控	(76)
第三节 转录后水平的基因表达调控	(85)
第四节 翻译水平的基因表达调控	(88)
第五节 翻译后水平的基因表达调控	(92)
第六章 变异的分子基础	(98)
第一节 遗传重组	(98)
第二节 基因突变	(114)
第七章 DNA 损伤与修复系统	(117)
第一节 DNA 损伤	(117)
第二节 DNA 修复	(121)

第八章 染色体和染色体病	(130)
第一节 人类正常染色体的基本特征	(130)
第二节 染色体畸变	(138)
第三节 染色体病	(145)
第九章 单基因遗传病	(157)
第一节 单基因病的遗传方式	(158)
第二节 常见单基因病的发病机制	(169)
第十章 多基因遗传病	(189)
第一节 数量遗传的多基因假说	(189)
第二节 多基因遗传病的遗传学特点	(192)
第三节 多基因遗传病的研究技术和研究进展	(198)
第十一章 线粒体遗传病	(202)
第一节 线粒体的遗传特点	(202)
第二节 线粒体基因突变和常见的线粒体遗传病	(206)
第十二章 药物遗传学与药物基因组学	(209)
第一节 药物遗传学	(209)
第二节 药物基因组学	(211)
第三节 基因组时代的个体化医疗	(215)
第十三章 肿瘤遗传学	(218)
第一节 染色体不稳定综合征和遗传性肿瘤综合征	(219)
第二节 染色体畸变与肿瘤	(227)
第三节 癌基因与肿瘤	(230)
第四节 肿瘤抑制基因与肿瘤	(234)
第五节 表观遗传学与肿瘤	(238)
第六节 端粒与肿瘤	(244)
第七节 肿瘤发生的遗传学说	(247)
第十四章 基因诊断和基因治疗	(251)
第一节 基因诊断	(251)
第二节 基因治疗	(263)
参考文献	(280)

第一章 绪 论

随着医学的进步和人类生活水平、医疗水平的提高,早先严重危害人类生命的传染性疾病已趋于绝迹或基本得到控制。与之相比,遗传病及与遗传因素密切相关的人类疾病在发病中所占比例日益突出,对人类的危害也更为明显。此外,由于化学工业以及社会其他方面的发展所带来的弊端和污染未能及时治理,也使得遗传病的发病率有增高的趋势。目前,已发现的遗传性疾病已经超过 6437 种,每 100 个新生儿中就有 3 ~ 10 名患有遗传性疾病。医学遗传学(medical genetics)作为临床医学与遗传学相互结合形成的一门科学,主要任务在于揭示各种遗传性疾病的遗传规律、发病机制、诊断和防治措施,已经成为医学中极为重要的方面,也是遗传学领域发展最快的学科之一。随着医学研究向分子水平深入,作为医学领域里最为活跃的学科之一,医学遗传学也已进入分子时代,称为分子医学遗传学,它运用分子生物学技术,从 DNA 水平、RNA 水平及蛋白质水平对遗传性疾病或疾病的遗传因素进行研究,揭示基因突变与疾病发生的关系,建立在分子水平上对遗传性疾病等疾病和疾病的遗传因素进行诊断,进一步实现对遗传病等疾病的基因治疗,达到从根本上治愈遗传病的目的。分子医学遗传学的研究对象不仅包括遗传性疾病,而且还包括其他获得性体细胞遗传疾病,如肿瘤、心血管疾病等。

一、健康与疾病的遗传学基础

疾病与遗传及环境的关系不仅是医学遗传学而且也是整个遗传学的基本问题。其实答案很简单,任何疾病都是遗传与环境两者相互作用所致。不同的疾病中,两者所起的作用不同。人类的其他性状如智力、身高等也是由遗传和环境所共同作用的。遗传性疾病主要是由遗传因素起作用,传染性疾病主要是外部环境因素所致。单纯由某一种因素决定的疾病极为罕见,如我国西南较为常见的一种遗传病——蚕豆病是由于编码葡萄糖-6-磷酸脱氢酶(G-6-PD)的基因缺陷所致,是一种典型的遗传病;如果该患者不食用蚕豆等食品,那么他的表型是正常的,这说明该病的表现与环境作用相关。同样,一些传染病过去被认为与遗传无关,如肝炎,虽然没有肝炎病毒的感染不会得病,但是同样的感染,不同的人表现不一样:有些感染得病,有些成为病毒携带者,还有些能够依靠自身的免疫系统清除所感染的病毒。即使感染得了肝炎,同样治疗后的结果也大不相同,这些差别中遗传的因素不容忽视。因而,不能说传染病与遗传无关。再者,许多病毒性传染病通过改变细胞的遗传物质,外部环境因素最终通过遗传而发生作用。肿瘤、心血管疾病及神经系统疾病大多为多因子疾病,是典型的遗传与环境两者作用的结果。

二、遗传病的定义及分类

遗传病 (genetic disease) 是由于遗传物质改变所导致的疾病。一种疾病是否有遗传病因, 有时不易确定, 因为即使感染性疾病也还存在遗传素质问题。下列事实提示某种疾病存在遗传病因:

1. 患者有特征性表现型 (通常伴有智力障碍) 和染色体异常, 有或无同一疾病或有关疾病的家族史。
2. 一卵双生的同病率高于二卵双生。
3. 在未知有触发因素的情况下, 患者有特征性的发病年龄和病程。
4. 在排除环境因素的前提下, 亲属中有一定比例的患病者。
5. 在非血缘成员 (配偶) 中不出现患者。

遗传性疾病与先天性疾病及家族性疾病不同。遗传性疾病常具有先天性、终生性和家族性的特点, 但先天性疾病并不都是遗传病, 在胎儿发育过程中, 由于环境因素或母体条件的影响, 可出现非遗传性先天性疾病。如胎儿在宫内感染天花造成出生时脸上有瘢痕, 母亲怀孕早期感染风疹病毒致使胎儿患有先天性心脏病, 孕妇服用反应停 (酞咪哌啶酮) 引起胎儿先天畸形等。遗传性疾病也常表现为晚发, 一些致病基因的作用仅在个体发育达到一定年龄后才表现出来。尤其是那些由遗传基因和环境条件两种因素支配着表现型的遗传性疾病, 如肌营养不良症到儿童期发病, Huntington 舞蹈病一般发病于 25 ~ 45 岁, 痛风好发于 30 ~ 35 岁。因此, 晚发的疾病仍可以是遗传性疾病。

遗传性疾病与家族性疾病也不同。遗传性疾病常可表现为家族性, 这是因为同一家系中的成员可共同具有某一致病基因。但同一家系的成员也处于相似的生活条件和环境中, 由相似环境条件所引起的非遗传性疾病, 有时也具有家族性。一个家族有多个成员患同一疾病 (如结核病、肝炎) 可能系由共同的生活环境所引起。夜盲也常有家族性, 但它显然并非遗传病, 而仅为维生素 A 缺乏所致。而一些遗传性疾病 (隐性遗传病) 仅在基因型为纯合子状态下才表现出来, 因而发病的概率小, 相似于散发性疾病。

(一) 常见的遗传病的分类

遗传病可分为五大类。在分析一种疾病的遗传基础时, 首先要明确它属于这五大类中的哪一类。

1. 染色体病

染色体病 (chromosome disorder) 是由于个体染色体发生数目或者结构畸变所导致的遗传性疾病。人类正常体细胞具有二倍体数 46 条染色体。如果在生殖细胞形成阶段或者受精卵早期发育过程中的细胞分裂时发生了同源染色体或者姐妹染色体单体的不分离, 或者染色体的异常断裂和错误拼接, 就会导致染色体发生数目或者结构畸变, 最终导致患者患上以种种先天发育异常为特征的染色体综合征。染色体病通常不在家系中传递, 但也有可传递的。已知的染色体病有 300 多种, 出生时染色体病发生率约为 7‰。在妊娠前 3 个月的自发性流产中, 染色体畸变约占一半。

2. 单基因病

如果一种遗传病的发病仅仅涉及一对等位基因,其导致的疾病称为单基因病(single gene disorder),这个基因称为主基因(major gene)。单基因病通常呈现特征性的家系传递格局。按照致病基因在染色体上的定位和致病基因控制的是显性性状还是隐性性状来区分,可将单基因病划分为五种类型,常染色体显性遗传,常染色体隐性遗传,X连锁显性遗传,X连锁隐性遗传和Y连锁遗传。目前已知的单基因遗传病有3000多种,单种单基因病均相对罕见,其发生率的上限约为2‰。但由于发现的病种越来越多,所以单基因病整体来说并非罕见。

3. 多基因病

多基因病(polygenic disorder)亦称复杂性疾病(complex disease),起因于遗传素质和环境因素的共同作用。与单基因病受控于一对等位基因不同,多基因病的发病受控于多对基因,但每对基因所起的作用都是微小的,称为微效基因。多基因病包括一些先天性发育异常性疾病和一些常见病,如糖尿病、高血压、精神分裂症和哮喘等。多基因病有家族聚集现象,但无单基因病那样明确的家系传递格局。

4. 线粒体遗传病

线粒体具有独立的细胞核外基因组,也即是线粒体基因组。线粒体基因组编码呼吸链部分肽链及线粒体蛋白质合成系统rRNA和tRNA。这些线粒体DNA的突变可致线粒体遗传病(mitochondrial genetic disorder)。由于受精卵中的绝大部分线粒体来自于卵子,精子的贡献很小,所以线粒体遗传病属于细胞质遗传,呈现母系遗传的传递格局。目前所知的线粒体遗传性疾病约有60多种。

5. 体细胞遗传病

并不是所有的遗传物质的突变都遗传自父母。人出生后的一生中体细胞会发生数十亿次有丝分裂,在每次分裂过程都有可能发生碱基突变,DNA拷贝数的变化或者染色体的异常分离等现象。体细胞中的突变累积到一定程度便会导致疾病的产生。最典型的体细胞遗传病(somatic cell genetic disorder)便是肿瘤。癌基因和抑癌基因的累积突变常可导致肿瘤的发生。有些先天畸形亦属体细胞遗传病。

(二) 遗传病对我国人群的影响

随着医学的进步,急性传染病和流行病逐渐得以控制,遗传病对人类的危害已经变得越来越明显,这从以下几方面的事实可以看出其概貌:

1. 据统计,我国每年约出生1500万新生儿。在全部妊娠中,自然流产约占15%,其中约有50%是染色体畸变所造成的,也即是单纯有染色体畸变就可造成112万例的自然流产。在每年约1500万新生儿中,约有1.3%有严重的出生缺陷或先天畸形,其中约有70%~80%涉及遗传因素。在新生儿中,除一部分有出生缺陷外,在出生后,由于携带的致病基因的表达还可出现多种遗传病,例如杜氏肌营养不良症、甲型血友病等。

2. 从人群中的患病率来估计,约有3%~5%的人患有某种单基因病,15%~20%的人患有某种多基因病,约1%的人患染色体病。此外,体细胞遗传病中的恶性肿瘤构成了我国不

同地区人群中死亡原因第一位或第二位。

3. 即使未受遗传病所累的人,也并非与遗传病无关。据估计,人群中平均每个人都携带有 5~6 个隐性有害基因,他们虽然未患遗传病,却可将这些有害基因向后代传递,所以称为致病基因的携带者。每个人都是 5~6 种有害基因的携带者,这就是人群的遗传负荷。人群的遗传负荷对于我们后代的健康是不利的,是一个需要及时面对的问题。

三、医学遗传学的发展简史

分子医学遗传学是医学遗传学进入分子医学时代的产物,因此在这部分,我们首先介绍一下医学遗传学的发展简史,它的发展可分为缓慢发展期,快速发展期和飞跃发展期三个阶段。

(一)缓慢发展期

人们对遗传学的最初步的认识就包含了医学遗传学的内容,如:一母生九子、子子不相同,兄妹不宜通婚等。在医学遗传学的萌芽时期,人们根据感性认识总结了一些医学遗传学知识。古希腊希波克拉底时代之前,人们就已经认识到某些疾病可能在家庭中传递。大约 1500 年之前,犹太教法典就有对“易出血者”的某些男性家属免除割礼的规定,证明人们已经认识了血友病的遗传规律。此外,人们发现了一些比较明显的遗传病。18 世纪 Maupertuis 研究了多指(趾)及皮肤和毛发缺乏色素者(白化症)的家系,指出这两种症状有各自不同的遗传方式。1859 年 Boedeker 首先确诊尿黑酸尿症,这是最早报道的先天性代谢病。

遗传学的发展推进了医学遗传学的进步。1866 年,现代遗传学的奠基人奥地利的僧侣 Mendel 发表了豌豆杂交实验的结果,提出了遗传性状是由成对的遗传因子决定的,在生殖细胞形成时,成对的遗传因子要分开,分别进入两个生殖细胞中去,并且在生殖细胞形成时,不同对的遗传因子可以自由组合。这被后人称为 Mendel 分离律和自由组合律。这两个定律是孟德尔遗传定律的中心内容。1900 年, Garrod 与 Baton 首次运用孟德尔遗传定律解释了一些罕见遗传病的遗传方式。1901 年, Garrod 描述了四个尿黑酸尿症家系,在遗传学家 Bateson 的提示下, Garrod 推测尿黑酸尿症属隐性性状。1903 年, Farabee 指出短指(趾)为显性性状,这是人类显性遗传的第一例。此后发现人体许多遗传性状都符合 Mendel 遗传。

1903 年 Sutton 和 Boveri 分别注意到 Mendel 遗传因子的行为跟生殖细胞形成和受精过程中染色体的行为完全平行,于是两人分别提出,遗传因子就在染色体上。这就是染色体遗传学说。1909 年 Johannsen 将遗传因子改称为基因(gene)。1910 年左右,美国哥伦比亚大学的 Morgan 研究果蝇性状的遗传方式时,发现果蝇的性状可分为四个遗传连锁群,这恰和果蝇的染色体对数相一致,提出了遗传的连锁律和交换律。连锁律(law of linkage)是指染色体是遗传的传递单位,因此一条染色体上的连锁基因也是一起传递给子代的。但是连锁不是绝对的,在生殖细胞形成过程中,同源染色体之间有时可交换一个片段,使连锁基因发生重新组合,这就是交换律(law of crossing-over)。

这个时期,人们还在群体水平对遗传病进行调查并对不同的遗传病进行分类、描述及规律总结。Adams 于 1914 年发表论文,讨论到先天性疾病、家族性疾病和遗传性疾病的差别,

遗传病发病年龄、环境因素、近亲婚配之间的关系等,其中涉及有关遗传病的一些基本问题,并逻辑地推论到遗传的一些基本原理。

(二)快速发展期

20 世纪 30 年代以后,细胞遗传学、生物化学和免疫学实验技术等实验技术在医学中的应用加速了医学遗传学的发展,使之进入快速发展期,并使医学遗传学在多个领域里取得了重大进展,形成和建立了一些分支,如生化遗传学、药物遗传学、生态遗传学、细胞遗传学和免疫遗传学等。

遗传性疾病中有相当多的部分是由于酶或蛋白质缺陷所引起的代谢性疾病,随着生物化学的发展,人们对一些遗传代谢性疾病有了进一步认识。所发现的病种由本世纪初的少数几种如尿黑酸症、白化病等扩大到目前的数百种。包括电泳技术、免疫技术肽链和氨基酸分析技术和酶促反应动力学研究技术等生化技术的应用,从生化水平上揭示了血红蛋白病、G-6-PD 缺乏症、苯丙酮尿症、尿黑酸症和高胆固醇血症等疾病的病理机制。如 Garrod 在研究尿黑酸尿症时提出患者体内的尿黑酸是酪氨酸的降解产物,由于先天性酶缺乏而不能被进一步氧化,因而在体内累积并随尿排出这样的假设。此后,1958 年,La Du 等发现患者活检肝组织中缺乏尿黑酸氧化酶,从而证实了 Garrod 的假设。这一时期,在理论研究上,医学遗传学最引人注目的进展是对血红蛋白的研究。Pauling 等(1949 年)在研究镰形红细胞贫血时发现患者有一种异常血红蛋白分子 HbS,其电泳性质不同于正常的 HbA,从而提出分子病(molecular disease)的概念。1954 年 Ingram 创立“指纹法”,查明 HbS β 链第 6 位氨基酸是缬氨酸而不是正常的谷氨酸,由此展开了对血红蛋白分子病的深入研究。在实际应用上,这一时期开辟了治疗某些遗传病的有效途径。苯丙酮尿症的治疗标志着这方面的重大进展。1953 年 Bickel 等发现通过控制新生儿的苯丙氨酸摄入量,能有效地防止苯丙酮尿症的发展。此外科学家们还发现乙酰胆碱敏感是由于血清胆碱酯酶缺乏所致,伯氨喹啉引起药物性溶血是由于 G-6-PD 缺乏所致。这些发现说明,药物反应有受遗传控制的代谢基础。1959 年 Vogel 提出药物遗传学(pharmacogenetics)一词。进入 70 年代后,这一概念进一步扩展,不仅药物反应要考虑遗传基础,而且对一切环境因子的反应,包括食物反应在内,也必须考虑遗传基础。1971 年 Brewer 提出生态遗传学(ecogenetics)这一术语,对食物、药物和毒物的反应都必须从遗传基础上加以认识。

遗传性疾病中约有 10% 的疾病表现为染色体畸变。在 1956 年 J. H. Tjio(蒋有兴)和 Levan 明确了人体染色体数目为 46 条后,染色体分析技术即被迅速应用于临床。细胞遗传学系列研究明确了缺失、倒位、易位、重复等染色体畸变的机制,发现了一大批染色体异常综合征。1959 年就有三个发现:Lejeune 等发现 Down 综合征患者有 47 条染色体,即多一条小型近端着丝粒染色体(第 21 号染色体);Ford 发现 Turner 综合征妇女只有 1 条 X 染色体;Jacobs 和 Strong 发现 Klinefelter 综合征男子的性染色体是 XXY。1960 年美国费城(Philadelphia)研究小组在慢性粒细胞性白血病(chronic myelogenous leukemia, CML)患者的细胞里第一次发现了特定的染色体结构畸变,称之为费城染色体或 Ph1 染色体。随后又发现了其他许多染色体综合征和一些肿瘤的标记染色体。20 世纪 60 年代末 Caspersson 等发现,植物染

染色体在用荧光染料氮芥喹吖因(quinacrine mustard)染色时,染色体的不同区域显示强弱不等荧光。1970 年将此技术应用于人体染色体,揭示了人体各条染色体独特的荧光带型。同时随着染色体显带技术,姐妹染色单体互换技术和高分辨显带技术的应用和发展,使人们能够越来越精细地识别染色体及其区带,为研究各种遗传病及遗传相关疾病的染色体畸变提供了有力工具,由此提高了染色体分析的精确性,并发现了不少新的染色体综合征。

免疫学实验技术的发展扩大了遗传病的概念,并为疾病防治带来了新的方法。1900 年 Landsteiner 发现了 ABO 血型。在这以后的半个世纪中,利用红细胞凝集试验先后发现了十几个血型系统,为临床输血配型奠定了基础。1941 年 Levine 等提出,胎儿红细胞增多症(新生儿溶血症)系由胎-母红细胞抗原不相容引起同种免疫(alloimmunization)所致。1952 年, Dausset 和 Nenna 在多次输血患者的血液发现了白细胞凝集素,1958 年检出第一个白细胞抗原 Mac (HLA - A2 + A28)。1964 年 Terasaki 和 McClelland 设计微量淋巴细胞毒试验以取代白细胞凝集试验。此后,免疫遗传学研究揭示了人体高度多态性的 HLA 系统,使器官移植供、受体配型有了可能。现已揭示了一系列遗传方式各异、临床表现多种多样的遗传性免疫缺乏病。

(三)飞跃发展期

自 1980 年以来,分子生物学的迅猛发展促进了医学遗传学的飞跃发展,使医学遗传学走向 21 世纪的大发展,也使得医学遗传学逐渐进入分子时代。

这一时期,在遗传性疾病致病基因的识别上,由于分子生物学的发现所导致的多种新的遗传学标记的发现,加速了遗传性疾病的致病基因的克隆的进展。1983 年,第一个疾病基因得到定位, Gusella J. F 等人发现定位于 4 号染色体的一个遗传学标记与 Huntington 舞蹈病高度连锁。然而 10 年后,1993 年这种疾病的致病基因 huntingtin 才被由来自 6 个研究组的 58 位科学家组成的合作小组克隆并精确定位于 4p16.3。1986 年,来自于 Boston Children's Hospital 的研究小组采用定位克隆的方法于 Xp21 分离到了慢性肉芽肿病的致病基因,这是首次通过定位克隆方法获得的致病基因。定位克隆是一种在对基因所编码的蛋白一无所知的情况下,完全通过 DNA 连锁分析鉴定致病基因的方法。定位克隆方法的出现体现了反求遗传学(reverse genetics)的新趋势。这就是在不知道某种遗传病蛋白质异常的情况下,直接寻找致病的 DNA 变异,进而揭示这种 DNA 变异所导致的蛋白质异常,使遗传学研究从表现型到基因型这条经典路线转变成为从基因型到表现型的反求路线。采用这种方法,人们很快找到了 Duchenne 肌营养不良(Duchenne muscular dystrophy, DMD)、视网膜母细胞瘤及囊性纤维化(cystic fibrosis, CF)的致病基因。近年来,在反求遗传学思路指导下,遗传学家们对一些不明原因的遗传病进行了 DNA 标记连锁分析,使 DNA 标记逐渐逼近,最终找到了多种单基因病的致病基因。

关于多种遗传性疾病的病因学的明确推进了遗传性疾病的诊断学的发展,同时限制性内切酶的发现及 DNA 分子杂交技术等分子生物学技术的建立,为遗传性疾病的诊断提供了新的手段。1976 年 Y. W. Kan 等、1978 年 Wong 等、1979 年 Dozy 等利用人群中限制性内切酶片段长度多态性(restriction fragment length polymorphism, RFLP)分析患者疾病位点之间的

连锁,发明了获取胎儿羊水细胞 DNA 进行地中海贫血和镰形红细胞贫血症出生前诊断的方法。随后不久,苯丙酮尿症、血友病等都已能在 DNA 水平作出诊断。

对于致病基因的明确直接导致了针对致病基因的基因治疗方法的出现。20 世纪 90 年代初,基因治疗(gene therapy)进入了临床试验阶段。所谓基因治疗就是将某个正常基因导入患者体内细胞中使之表达,对患者缺乏的或异常的某种蛋白质提供其正常表达产物,从而起到治疗作用。由腺苷脱氨酶(adenosine deaminase,ADA)缺乏引起的严重联合免疫缺乏症(severe Combined immunodeficiency,SCID)和由凝血Ⅸ因子缺乏引起的血友病 B,基因治疗的临床试验都已得到初步的治疗效果,但还存在着安全性、有限性和可调控性方面的一系列问题。致病基因的明确也推进了靶向致病基因的药物开发。2001 年,FDA 批准 Bcr - Abl 酪氨酸激酶抑制剂 - imatinib (Gleevec) 格列卫用于治疗慢性粒细胞白血病。

这一时期最为重要的进展便是人类基因组计划的实施与完成。1986 年,诺贝尔奖金获得者 Dulbecco 提出,如果我们希望对肿瘤有更多的了解,这就必须集中于细胞的基因组,最有用的是从细胞的基因组测序开始。经过学术界几年的争论,1990 年美国国会批准 15 年(1991 ~ 2005)拨款 30 亿美元的人类基因组计划(human genome project,HGP)。计划通过三部曲,即连锁图(遗传图)、物理图和基因组测序,揭示人类基因组 DNA 30 亿碱基对的全序列。由于 HGP 意义重大而影响深远,它引起各国政府高度重视,纷纷投入大量资金推进 HGP 研究,使其研究进展一再超前。2000 年 6 月 26 日美国总统克林顿和英国首相布莱尔宣布人类基因组序列工作草图诞生。2004 年 10 月 21 日,《自然》杂志公布了人类基因组的完成序列,宣布了人类分子遗传学的登峰之作的诞生。

医学遗传学发展至今,已经形成由细胞遗传学、分子遗传学、生化遗传学、药物遗传学、免疫遗传学、行为遗传学、生态遗传学、体细胞遗传学、群体遗传学、临床遗传学等约 15 个分支学科所组成的涉及基础与临床的综合性科学。

四、展望分子医学遗传学的发展

人类基因组计划完成之后,疾病相关基因的定位、克隆和鉴定依然是分子医学遗传学研究的核心组成部分。这是因为数千种单基因遗传病和多种大范围危害人群健康的多基因疾病的致病基因和相关基因代表了全部人类基因中至关重要的那一部分。疾病相关基因的发现,可直接促进基因诊断和靶向治疗的发展。展望分子医学遗传学的发展,将具有以下三大趋势:

(一)疾病尤其是复杂性疾病的相关基因的识别步伐大大加快

基因组在结构或者表达上的异常可能是人类大部分疾病的根本原因。人类基因组全序列的最直接的应用就是促进了疾病相关基因的识别。人类基因组计划的完成将疾病相关基因的识别策略从传统的定位克隆方法推进到定位候选克隆方法。传统的定位克隆方法建立在遗传连锁分析原理基础上,首先采用大量的 DNA 多态标志对受累家系进行连锁分析,初步定位与疾病相关的染色体区段,然后克隆出该区段的 DNA 片段,从中寻找突变,最终确定疾病相关基因。新的定位候选克隆方法是采用连锁分析的方法将某个疾病位点定位于某一

染色体区段,由于人类基因组计划的完成明确了染色体的各个区域中已知的人类基因分布,这时即可从局部的染色体区域中遴选出结构、功能相关的基因,也即是候选基因进行分析,因而大大加快了疾病相关基因的识别速度。目前约 3000 种人类单基因遗传病中一半以上的疾病基因已经被识别,而且新发现的疾病基因正在以大约每周 5 个的速度增长。

由于多基因病不遵循 Mendel 遗传表型——基因型的规律,难以用一般的家系遗传连锁分析取得突破。多基因病已成为后基因组时代疾病基因组学研究的重点。目前,结合 MS 和 SNP 等高度多态性标志,已发展对受累同胞进行分析、关联分析和连锁不平衡分析等多种分析,对复杂性疾病家系和人群进行疾病相关位点在基因组中的定位。一般认为,隔离人群或较均一的人群可能是进行多基因性状定位研究较理想的群体。2005 年完成的单倍型图谱 (HapMap) 计划将有效地促进多基因疾病的研究。

另外,在人类全基因组中,编码蛋白质的基因只占人类基因组序列的 1.1% ~ 1.4%。疾病相关基因的概念扩展到蛋白编码基因之外的基因组中其他的功能元件,尤其是许多参与调控基因时空特异性表达的顺式调控元件。我们目前对基因组中的功能元件知之甚少,但最近启动的 ENCODE (ENCyclopedia Of DNA Elements) 计划致力于识别人类基因组中所有的功能元件,将为我们寻找疾病相关基因提供新的研究对象。

相比结构的异常,基因组在表达上的异常可能是人类大部分疾病的根本原因,因此另一类寻找疾病相关基因的方法不是基于反应基因组结构的基因序列的分析,而是通过研究疾病状态下基因表达的变化。首先,表观遗传学的异常可能是导致基因表达调控异常的重要原因。目前,一些大规模检测表观遗传型的方法,特别是检测 DNA 甲基化的方法已经成熟;人类表观基因组计划 (HEP) 将识别人类基因组中所有的甲基化可变位点。这些进展将使我们可以从表观遗传学水平寻找疾病相关基因。其次,一个基因编码的 mRNA 或蛋白质在疾病和正常组织的表达有显著的差异,可以作为该基因与疾病相关的间接证据。DNA 阵列技术可以同时检测上万个基因在疾病和正常组织之间的表达差异。应用这种技术,已经发现了许多复杂疾病(如癌症、糖尿病、精神分裂症等)的相关基因。人类基因组计划的完成使得芯片上的基因数目更多、分类更细致,借此平台,许多公司已可提供涵盖了所有人类已知基因的芯片,这些都将大大扩展我们对于疾病相关基因的识别。

(二) 基于基因组的诊断学蓬勃发展

在越来越多的疾病相关基因得以鉴定或者识别的基础上,遗传病的诊断技术将获得进一步蓬勃发展。目前人类单基因遗传病中的一半以上疾病基因已经被识别,它们将很快被用于疾病诊断,迅速增加的基因诊断方法使得许多疾病得以精确诊断。

许多常见病,包括肿瘤、心血管病、糖尿病等,属于多基因病,其患病风险有一部分取决于遗传因素,另外一部分则取决于环境因素,例如吸烟、饮食和锻炼。它们的遗传因素来自于多个基因的共同作用,每个基因对于患病风险的贡献不足 5%,所以难以用单个基因进行诊断。在这种情况下,基因表达模式可被用于疾病的诊断。广泛应用的工具是 DNA 微阵列。它在早期曾被用于淋巴瘤和乳腺癌的分型,现在已经广泛应用于许多其他肿瘤,包括脑癌、卵巢癌、肺癌、结肠癌、肾癌、前列腺癌、胃癌和白血病等。另外,蛋白质组的研究,特别是

最近开展的人类蛋白质组计划(HPP)可能会提供在蛋白质水平上进行疾病诊断的方法。

(三)基于基因组的靶向治疗方兴未艾

人类基因组计划的完成提供了更多的潜在药物靶点,可供新药开发。首先从蛋白质药物靶点来说,目前只有不到 500 个蛋白质成为小分子药物的靶点。人类基因组包含 2 万 ~ 2.5 万个基因,而由于存在 mRNA 选择性剪接等机制,蛋白质的数量要远多于此,就算只有很少一部分蛋白质可以作为药物靶点,预期它们的数目会有数千个,这大大扩展了药物靶点的数目。此外,除蛋白质以外,各种 RNA 以及基因组内的调控元件都有可能成为药物的靶点。深入的功能基因组学研究可扩充新的药物靶点库,结合其他如高质量的化合物文库的储备以及大规模、高通量的药物筛选技术的建立必将大大加快药物开发的速度。其次,基因阵列技术用于确定新药的特异性和作用机制,可以更为精确地评价药物在分子水平的效果。FDA 已经联合制药和生物技术公司开始制定阵列数据发布的标准,目的是加快这些数据在药物开发上的应用。另外,由于基于疾病机制的治疗策略比原始的基于表型的治疗方法具有更好的效果,人类基因组对靶向治疗的另外一个贡献是可以产生更好的药物。如用于治疗乳腺癌的生长因子受体 Her2/Neu 的单抗 Trastuzumab (Herceptin),可以抑制多种骨髓瘤中的 NF κ B 途径的激活蛋白酶体的抑制剂 bortezomib (Velcade)等,这些药物都是靶向病因的治疗方案,它们在临床上都显示了良好的疗效。

目前的医学正在从“以疾病为中心”到“以健康为中心”过渡,针对人类基因组展开研究的贡献,不仅在于寻找到病因和指导治疗,还在于理解疾病发病机制,维护人类健康从而预防疾病的发生。通过研究人类基因组多态性,特别是大量单核苷酸多态性(single nucleotide polymorphism, SNP)的获得,一方面可以使连锁不平衡分析变得简单,从而大大加速确定常见疾病的易感基因的步伐;另一方面,这种多态性还可以阐明个体间在药物反应上存在差异的遗传基础,从而根据个体的遗传背景优化治疗方案,实现“个体化医学”。2005 年,旨在鉴定不同人群基因组中的多态性位点的 HapMap Project 1 期计划完成,共获得约 100 万个遗传多态性位点;2008 年 4 月 17 日,在《自然》杂志上,美国的科学家发表了首个利用新一代高速测序技术得到的全基因组,而接受这一全基因组测序的不是别人,正是“DNA 之父”James Watson。2008 年 11 月 6 日,《自然》杂志以封面文章形式发表了由深圳华大基因研究院完成的首个中国人基因组序列研究成果(定名“炎黄一号”)。这些成果标志着我们向个人化基因组这一目标又迈进了一步,也为个体化医疗的最终实现带来了新的曙光。展望不久的将来,医疗模式可能发生巨大的变化。基因芯片越来越多的进入临床服务,辅助进行高效分子诊断,并提供每个人基因组的表达格局信息,即是生物学密码。临床医生可以根据每个人的生物学密码,制定个人特异的治疗方案,避免药物的毒副作用,或根据每个人的生物学密码判断多基因复杂病的发病风险,通过改进生活方式,防止发病,使医疗服务从治病走向防病。高科技的基因组医学在造福病人的同时,无疑也存在个人遗传信息泄密或被误用的风险。因此,临床医生必须注意为病人保守秘密,维护病人知情同意的权益,反对基因增强和胎儿选择,以防止滑向反科学和反人道的新优生运动。综上,可以说,医学正在进行一场影响深远的遗传学革命。

分子医学遗传学在临床上的应用催生了遗传医学(genetic medicine)。遗传医学是为遗传病患者提供临床服务,包括遗传病的诊断、治疗、筛查、预防、咨询、随访等,以尽可能减少遗传病患者的痛苦,使他们尽可能享有幸福的人生。目前,我国大城市的医院没有遗传咨询门诊,临床各科的遗传医学服务也日益受到重视。临床医生必须掌握发展中的分子医学遗传学和医学伦理学,才能为病人提供更好的服务,从而为社会带来福祉。