

医学遗传学讲座

□ 夏家辉 主编

编者：(按章顺序排列)

邓汉湘 夏家辉 阮庆国

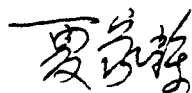
杨新平 黄 蕾 唐冬生

李麓芸 刘春宇

□ 湖南科学技术出版社

前言

第一线研究人员，他们对相关领域有比较全面的了解，并尽力收集有关方面的最新进展，不少信息是国际互联网上信息资源库的实时动态信息，如孟德尔遗传病的基因克隆及其他人类基因组计划等信息。故本书反映了医学遗传学热门领域的最新动态，不仅可作为医学研究生和本科生的教材，也可供从事医学遗传学研究、医疗工作的研究人员和临床医务人员参考。



1998年7月于长沙

目 录

第一讲 概 论·····	邓汉湘(1)
第二讲 染色体病及其研究进展 (1) ·····	夏家辉(7)
第三讲 染色体病及其研究进展 (2) ·····	夏家辉(16)
第四讲 单基因病的研究进展—— 人类基因组计划和基因克隆·····	阮庆国(31)
第五讲 单基因病的研究进展—— 单基因病的基因克隆及典型疾病·····	夏家辉 阮庆国(40)
第六讲 多基因病及其研究进展 ·····	夏家辉 杨新平 (100)
第七讲 线粒体病及其研究进展 ·····	黄 蕾 (114)
第八讲 基因治疗研究策略与研究进展 ·····	唐冬生 (123)
第九讲 遗传咨询与产前诊断 ·····	夏家辉 李麓芸 (133)
第十讲 基因组研究信息资源及其利用 ·····	刘春宇 (139)
本书缩略语英汉对照 ·····	唐冬生 (148)

第一讲 概论

生命现象是物质运动的一种最高形式。每一生命个体和自然界中其他独立存在而又相互联系的千千万万种事物一样，都有一个发生、发展、昌盛、衰落和消亡的过程。尽管这一过程受诸多因素的影响，然而，存在于生物体脱氧核糖核酸（DNA）中的遗传信息则是制约和编排这一程序性过程的决定因素。遗传信息在生命过程中的这一决定性作用主要是通过程序性基因活动并界定生物体对环境条件变化的反应范围来实现的。基因的异常，包括数目异常（如染色体畸变）和结构异常（如基因突变），均可干扰和破坏正常的程序性基因活动以及使生物体对环境条件的变化产生不利于个体生存的反应，或者由于环境条件的变化超越了个体遗传信息所能反应的范围，而导致生命活动的异常，即医学上所称之为的疾病。在不同的疾病中或同一疾病的不同类型中，遗传因素所起作用大小不一，但没有一种疾病的产生、发展和转归是脱离遗传因素的影响和制约而独立存在的。不同的生物种族之间由于其遗传构成的差异较大，它们可具有不同的易感或易患疾病，如禽类对某些病毒高度易感，而人类对这些病毒则不感；反之人类对某些病高度易感，而禽类则不感。属于同一种族的生物体，虽然具有相同的易感和易患疾病，但由于个体间的遗传学差异，使其对同一疾病的易感性或易患性各不一样。对于某些疾病，单个基因的突变足以引起表型改变；而对于另一些疾病，单个基因的改变尚不足以引起表型改变，还需要有其他基因和/或环境因素的共同作用。

医学遗传学是研究人类的生物学或遗传学变异及其与健康 and 疾病关系的学科。严格地说，作为一个独立的学科，医学遗传学是自 1956 年人类的染色体数目被确定以来，才建立并发展起来的。

一、遗传的物质基础

1865 年，Mendel 遗传物质的颗粒特性的发现，是遗传学发展史上的里程碑，被誉为遗传学研究的起点。但他的这一重大发现，也许超出了时代所能理解的极限，其重要意义并没有被人们所认识；同时，由于当时染色体的存在以及减数分裂的现象尚未被发现，使 Mendel 本人也不能为其理论作出合理的解释，以至“融合遗传”理论一直居统治地位。

1882 年，Flemming 在人类的肿瘤细胞中发现了染色体的存在。随后，有丝分裂和减数分裂现象也相继被发现。这些重大突破，使 Weismann 等人推想到染色体可能就是遗传物质的载体。1900 年，Mendel 定律被重新发现。Sutton 和 Boveri 等人于 1903 年正式提出了遗传的染色体理论，认为遗传物质存在于染色体上。这一理论经 Morgan 等人的连锁与交换分析，将某一特定的基因与某一特定的染色体直接联系起来而得到了证实，并由此创立了遗传的基因理论。Mendel 和 Morgan 等人的工作，使遗传学从过去的猜想、假设，发展成为一门

严谨的实验科学。他们以新的实验结果为基础,创造新理论,建立新学说,而使新的理论和学说客观地反映事物内在的本质的联系。是科学发展史上成功的范例。

遗传的染色体理论确立了遗传物质位于染色体上。染色体主要由 DNA 和蛋白质两大物质构成。决定 DNA 是遗传物质的实验是由 Griffith 和 Avery 等于 1928 年和 1944 年逐步完成的。他们通过光滑型和粗糙型肺炎双球菌的转化试验证明只有 DNA 能将粗糙型肺炎双球菌转化为光滑型;而蛋白质不具备这种转化功能,从而证明遗传物质是 DNA 而不是蛋白质。随后的一系列研究工作均证明在有 DNA 的生物中, DNA 是遗传物质;在没有 DNA 而只有 RNA 的生物中,如某些 RNA 病毒, RNA 是遗传物质。1953 年, Watson 和 Crick 的 DNA 双股螺旋结构的发现,以及随后遗传密码的破译,使人们认识到决定生物性状的遗传信息存在于 DNA 的碱基顺序之中;基因即为一段具有遗传学功能的 DNA 序列。DNA 由 A/T 和 G/C 互补,方向相反的两股多核苷酸键构成。人类的体细胞含有 46 条染色体(女性 23 种,男性 24 种染色体),每一染色体代表一条 DNA 链。人类单倍体基因组约为 3×10^9 bp,估计含有约 5 万~10 万个基因。每一个功能基因在一个单倍体组中通常只有一个拷贝。但结构和功能相似的几个基因可构成一个基因家族(gene family)。这些单拷贝基因和他们相关的调节序列一起约占基因组的 70%,其余 30% 属于重复序列。重复序列的功能尚不清楚,可能与基因表达的调控,染色体的运动以及染色体的稳定性有关。重复序列分为串联重复(tandem repeats)和散置重复(interspersed repeats),串联重复又叫卫星 DNA(satellite DNA)因为他们在密度梯度离心时,表现为单拷贝 DNA 主带以外的 DNA 带。串联重复约占基因组的 10%,又可分为微卫星 DNA(microsatellite DNA),小卫星 DNA(minisatellite DNA)和大卫星 DNA(macrosatellite DNA)。微卫星重复通常小于 1kb,最常见的重复序列有 CA, GA, NAA, NAAA(N 代表任意核苷酸)等。CA 通常可重复 10~60 次,在基因组中大约每 30~40 kb 可发现一次这样的重复,而这种重复的次数在人群中具有多态性,故在遗传病的连锁分析和法医学上起着重要作用。小卫星重复常在 1~30 kb 之间,重复序列的区段比微卫星重复长,重复次数在人群中亦有多态性。大卫星重复可大至数百万碱基对,他们通常存在于染色体的着丝粒和端粒区。散置重复约占基因组的 20%,他们通常在一段 DNA 中存在一个单拷贝,但在基因组中有很多这样分散存在的 DNA 序列。散置重复根据其重复片段的大小分为短散置重复和长散置重复。短散置重复通常小于 500 bp,最常见的为 Alu 重复,Alu 重复大约 300 bp,含有限制性内切酶 Alu 的酶切位点。在人类基因组中,约 6% 的 DNA 属于 Alu 重复序列,该重复序列只存在于人类和其他高等灵长类的基因组中。长散置重复通常 0.5~5 kb,如 Kpn 重复序列等。

DNA 有两个主要功能,即自身复制并传递给下代和指导蛋白质的合成,其指导蛋白质合成的功能是通过 mRNA 实现的。在基因组中,约有 80% 的基因只在特定的时间和特定的细胞中表达,如胰岛素基因只在胰岛的 B 细胞中表达。其余 20% 的基因在所有的细胞中均有表达以满足细胞进行基本代谢和执行基本功能的需要,该类基因叫家政基因(housekeeping gene),家政基因的表达约占细胞中基因表达总量的 90% 或更多。

在人类基因组中,基因的大小相差很大,有的基因很小,如胰岛素基因仅 1.4 kb,编码 51 个氨基酸;而有的基因很庞大,如进行性肌营养不良(DMD)基因长达 240 余万个碱基对(2400 kb),编码 3000 多个氨基酸。但多数基因界于 10~40 kb 之间,编码 200~400 个氨基酸。

一个功能基因常具有下列几个主要组成部分:即启动子(promoter),外显子(exon),

内含子 (intron)。启动子位于基因的 5' 端, 是 RNA 多聚酶和转录调节因子的结合区。启动子约几百个碱基对, 通常含有一个共同的碱基序列即 TATA (又叫 TATA 盒) 和/或一个 CAAT (又叫 CAAT 盒), 该序列可结合转录因子。RNA 的转录可有多位点, 但通常有一个主要转录起始位点。整个转录过程一直延续到转录终止位点。转录完成后, 经过 mRNA 修饰过程 [即 5' 加 G^m 帽, 3' 加 A 尾和 RNA 剪接 (RNA splicing)] 而形成成熟的 mRNA。在 mRNA 剪接过程中, 大部分的非蛋白质编码区被剪除而不存在于成熟的 mRNA 中。这些存在于基因中被转录而不存在于成熟 mRNA 中的部分叫内含子; 既转录又存在于成熟 mRNA 中的基因部分叫外显子。成熟的 mRNA 只含有基因的外显子部分。在成熟的 mRNA 中, 编码蛋白质的部分叫编码区 (coding region), 以起始密码 AUG 开始, 至终止密码 UAA, UAG 或 UGA 结束。起始密码 5' 端的非编码部分叫 5' 非编码区 (5' untranslated region 或 5' UTR) 或引导序列 (leader), 而终止密码 3' 端的非编码部分叫 3' 非编码区 (3' untranslated region 或 3' UTR) 或尾随序列 (tailer)。外显子和内含子交界处的 DNA 序列具有一定的特异性, 是剪接酶的识别位点。内含子通常以 GT (5' 供体位点) 开始, 以 AG (3' 受体位点) 结束。少部分基因不具有内含子叫无内含子基因 (intronless gene)。启动子、外显子以及外显子/内含子交界处 DNA 序列的改变, 将会影响 mRNA 的转录、稳定性、阅读框架和编码, 而影响蛋白质的结构和功能, 进而导致遗传病。

二、遗传病

人类作为最高级的生命体, 经过漫长的进化过程, 它与环境之间以及个体内部各器官、组织、细胞、细胞内的各亚细胞结构之间均达到了一种高度的和谐与统一。这种与外界和内部达成和谐统一的信息储存于遗传物质中, 并不断向下代传递。遗传物质在数量和质量上的改变, 均可干扰正常的生命活动过程而引起疾病。尽管人类所有的生物学性状 (包括各种疾病) 均与遗传密切相关, 但我们通常所谓的遗传病是指主要由于遗传物质的改变而引起的疾病。近年来, 随着医学遗传学研究方法和手段的提高与完善, 遗传病的概念也在不断扩大。一些遗传因素并不起主要作用的疾病但由于遗传因素的作用显而易见, 也被纳入遗传病的范畴, 如载脂蛋白的变异型 E4 与早老性痴呆症的关系等。

遗传病根据遗传物质改变的类型、大小、部位、作用以及时间, 大致可分为染色体病 (chromosomal disorder), 单基因病 (single gene disorder), 多因子或多基因病 (multifactorial disorder), 线粒体病 (mitochondrial disorder) 和体细胞遗传病 (somatic cell genetic disorder), 见表 1-1。

表 1-1 遗传病的类型和发病率

类型	出生时发病率	成年后发病率	终生发病风险
染色体病 ¹⁾	1.8/1 000	2/1 000	3.8/1 000
单基因病 ²⁾	3.6/1 000	16.4/1 000	20/1 000
多基因病 ³⁾	46.4/1 000	600/1 000	646.4/1 000
体细胞遗传病 ⁴⁾	—	240/1 000	240/1 000

1) 该栏发病率未包括无症状的平衡易位携带者频率。

2) 该栏发病率包括线粒体病发病率在内。

3) 该栏发病率中的先天畸形, 仅包括已经证实遗传因素起有重要作用的畸形, 对成年后的慢性疾病, 发病率代表一个总体估计数。

4) 该栏不包括遗传性单基因癌症综合征, 但假设所有的癌症均是由于累积性基因突变的结果。

1. 染色体病

在人群中的发病率较低，但因引起该类疾患的遗传物质的改变较大，通常为数百万个碱基对以上，累及数个甚至上百个基因，故染色体病通常表现为累及多器官、多系统的临床综合征。除少数性染色体异常以外，大多数的染色体病均表现有不同程度的智力低下。新生儿染色体异常的发生率（包括不具有临床表现者）约为 5‰~7‰。据夏家辉等人统计，我国新生儿染色体异常的发生率为 7.2‰。但实际上，这仅仅只是染色体异常胚胎的很小一部分。据分析，在流产或死亡的胚胎中染色体异常的发生率达 32%~42%，在早孕流产胚胎中，染色体异常的发生率更高。由于染色体异常所涉及到的基因甚多，故绝大部分染色体异常的个体不能发育至孕末期而中途流产。小部分能发育至分娩者多为性染色体异常和较小的常染色体异常。如 45, X 的发生率约占整个妊娠的 1%，但 98% 的 45, X 胚胎均中途流产。配子的染色体分析表明，10% 的精子 and 25% 的卵子具有染色体异常。由此可见，胚胎发育过程也是一个进化选择过程。

2. 单基因病

由于仅涉及到单个基因的突变，故较染色体病而言，单基因病在临床上的表现较简单而更具特征性。根据基因在个体的发育过程中所起的作用，单基因病的发病年龄相差很大。有的发病很早，在新生儿期即可表现出来，如数种类型的糖原储积症；有的在儿童期发病，如进行性肌营养不良症；有的成年后甚至到老年才表现出来，如肌萎缩性侧索硬化症、震颤麻痹等。在不同的人群中，单基因病的发病率有很大的差异，如位于赤道附近的非洲，镰刀状红细胞贫血基因携带者的频率高达 30%；在白种欧洲人中，囊性纤维化是发病率最高的遗传病；南非的 Afrikaner 人 Variegate porphyria 和家族性高胆固醇血症比其他人群高得多。这些主要与环境选择、奠基者效应 (founder effect) 或基因漂流 (genetic drift) 有关。毫无疑问，随着科学和技术水平的提高，将会有越来越多的疾病基因被鉴别。理论上，如果某一基因的突变不具有早期致死效应，而该基因又是个体发育的某一时期所必需，那么，该基因的突变（失去原有功能或获得新的功能）就可能引起生理活动的异常而导致临床上可见到异常表型即疾病。故人类遗传性疾病的位点数可高达基因组中的基因总数。另一方面，由于某些基因具有两种或两种以上的遗传学功能，故同一基因不同位置的突变可影响同一基因的不同功能而引起两种或两种以上的不同的遗传病。如着色素性干皮病 D 组蛋白质 (xerodermapigmentosum group D protein, XPD) 至少与两种不同的功能相关，即核苷酸切除修复和基础转录功能。不同区段的突变可影响不同功能而引起不同的疾病。如突变影响核苷酸的切除修复功能可引起着色素性干皮病；如突变影响基础转录功能则可导致缺硫脆发症 (trichothiodystrophy)。另外，不同基因的突变可引起同一类型的疾病，如至少有四个不同的基因突变可引起多发性外生性骨疣，而引起脊髓小脑型共济失调的致病基因至少有 8 种以上。这种不同的基因变异引起类似或同一表型或疾病的现象称作遗传异质性 (genetic heterogeneity)。

3. 多基因病

由于多基因病的发病涉及到多种基因和多种环境因素，疾病的发生不表现为孟德尔遗传，故其遗传学的研究要相对困难得多。目前用于多基因病分析的常用方法有连锁分析法、同胞配对研究法和候选基因的突变分析法，但这些方法的应用仍有很大的局限性，故多基因病也就成为了解得最少的病种。由于多基因病发病率最高，对人群健康的危害最大，对该类疾病的研究也就成为未来医学遗传学研究不可避免的重要课题。至目前为止，比较肯定的大

约有 20 余种常见的疾病属于多基因病，其中的先天畸形包括唇裂和腭裂、先天性心脏病、神经管缺陷以及幽门狭窄等；常见的成年人疾病包括类风湿性关节炎、癫痫、精神分裂症、躁狂抑郁症、多发性硬化、糖尿病、高血压等。多基因病虽然病种不多，但每种疾病在人群中的发病率甚高，比一般单基因病高出几千倍甚至几万倍，故总体发病约占先天畸形的一半左右，而占成年人疾病的一半以上。值得注意的是在这些常见疾病中，有部分类型又属于单基因病，即某一个基因的突变便足以引起该病表型的产生。如癫痫中的进行性肌阵挛性癫痫就可由于 cystatin B 的基因突变而引起。

4. 线粒体病

线粒体病实质上属于单基因病，但也带有多基因病和体细胞遗传病的特点。由于在受精时卵子带有上万个线粒体，而精子含有的线粒体数极少(10 个左右)，且受精后，精子所带的线粒体被逐渐排除，所以，线粒体是严格母性遗传的，故由线粒体基因突变而引起的疾病也就表现为母性遗传。正常细胞含有几百个线粒体，每个线粒体中含有 10 个左右的线粒体 DNA。线粒体 DNA 是一个闭环分子，大小为 16 569 bp。除编码 12S、16S rRNA 和 22 种 tRNA 外，线粒体 DNA 还编码另外 13 个多肽链，是线粒体氧化磷酸化复合物的的重要组成部分。氧化磷酸化由 5 个多聚合物构成。复合物 I ~ IV 构成电子转移链，复合物 V 为 ATP 合成酶。除复合物 II 全部由细胞核基因编码的多肽链构成外，其他 4 个复合物均含有线粒体基因编码的多肽链，如在构成复合物 V 即 ATP 合成酶的 16 条多肽链中，细胞核基因编码 14 条，线粒体基因编码 2 条。线粒体基因的突变可影响其氧化磷酸化功能而产生疾病。Leber's 遗传性视神经病(Leber hereditary optic neuropathy)是一种急性或亚急性、无痛性中央视觉丧失导致出现中央盲点。有的可早在儿童期发病，有的则成年后才发病，大多数在 20 ~ 30 岁出现临床症状，呈母性遗传，现已发现线粒体 DNA 中多个基因的突变均可导致本病。由于在同一个细胞中，有几百个线粒体，故线粒体病的临床表现除与基因突变的位置和类型有关外，还与突变线粒体 DNA 的数目与正常线粒体 DNA 数目的比例有关，故线粒体病带有多基因病的特点，即症状轻重的连续性。另外由于线粒体基因的突变频率比核基因高出 10 倍以上，故在个体的发育成熟过程中，不断地有新的突变产生并积累，当这种积累的突变达到一定的比例，影响氧化磷酸化的功能，即可引起临床表型，所以线粒体病的这一特征又与体细胞遗传病很相似。

线粒体病的临床表现还与不同的细胞对能量代谢的依赖性有关。由于神经系统、心肌、骨骼肌、肾脏以及内分泌腺等组织中的细胞对能量代谢的依赖性比其他细胞更大，故线粒体病时，这些器官更易受累。至今为止，已确定的线粒体病有 60 种。

5. 体细胞遗传病

该病又称累积性遗传病(cumulative disorder)，是由于体细胞的数个基因突变的累积作用而引起的。这种情况最常见于癌症。癌症发生的本质是由于基因突变而导致细胞的无控制增殖。大部分的癌症是散发性的，约有 5% 的病例呈家族遗传性。这主要是由于引起癌症的基因突变是一个多步骤的过程，当某一细胞同时具有两个或更多的基因突变，才能导致细胞的无控制性增殖。家族性遗传性的癌症第一个基因突变发生在生殖细胞中，带有核突变的个体的每一个体细胞都带有同样的突变，如该个体在其生长发育过程中再发生一次体细胞突变，则可产生癌症，故这种具有生殖细胞突变的个体通常发生癌症的年龄较早，双侧发病者较多，且并发其他癌症者较多见。以乳腺癌为例，大多数散发的乳腺癌为单侧发病，但若是 I 型乳腺癌基因(BRCA1)的携带者，则到 50 岁时，发生双侧乳腺癌的风险为 50%，70 岁时

双侧病者达 80%。同时 BRCA1 的携带者并发卵巢癌的风险也大为增高,到 50 岁时卵巢癌的发生率为 23%;70 岁时,发生卵巢癌者达 63%。另外, BRCA1 携带者也有较高的大肠癌和前列腺癌的发病风险。引起癌症的基因突变常发生在下列三组基因中,即肿瘤抑制基因(tumor suppressor gene)、癌基因(oncogene)和 DNA 修复基因(DNA repair gene)。由于绝大部分的引起癌症的突变是在个体的生长发育过程中发生的,故该类突变只局限于相关的肿瘤细胞中,且发生肿瘤的风险随着年龄的增大而增加。除癌症外,自身免疫性疾病以及老化可能也属于体细胞遗传病的范畴。

6. 医学遗传学的研究展望

人类的疾病作为人类的一种特殊性状,其发生和转归的本质,是遗传信息在一定的环境条件下表达的结果。如果对疾病发生和转归追根寻源,最终将要追溯到与疾病相关的基因或基因群。而要对疾病的发生作有效的预防或对疾病的转归作有效的干预进而达到治疗的目的,归根到底是要抑制异常基因的表达(如显性遗传病或癌基因突变引起的癌症),或增强正常基因的表达(如隐性遗传病或肿瘤抑制基因和 DNA 修复基因突变引起的癌症)或通过药物对抗或模拟基因活动或改善环境条件以调节基因活动。基于这一认识,本世纪 80 年代一批处于世界科学研究前沿的医学家、生物学家和计算机科学家联合提出了一项与 60 年代“登月计划”相似的“人类基因组计划”(Human Genome Project)。这一计划的目的,在于确定人类基因组中大约 5 万~10 万个基因在人类 24 条染色体上的位置,即基因的染色体定位或基因制图,并确定整个人类基因组大约 30 亿对核苷酸的碱基序列。这一计划自 1990 年实施以来,为医学遗传学的发展带来了革命性的变化,一大批直接引起疾病的所谓“致病基因”被克隆,而且这一克隆的速度在不断加快。如 1990 年,克隆的主要疾病基因仅 3 种,1993 年达 13 种,而 1995 年达 60 余种。可以预期在未来 5~10 年内对基因结构(即基因克隆)的研究是医学遗传学研究的主要任务。基因结构的阐明,必将引导出一个更加宏伟的计划,即基因的作用和作用机制的研究。较人类基因组计划而言,基因的作用和作用机制的研究将更直接地为人类的健康服务。毫无疑问这一研究必将开创一个药物与生物预防和治疗的新时代。

参考文献

- [1] Adolph, Kenneth W. Human molecular genetics. Academic Press, 1996
- [2] Shaw, Duncan J. Molecular genetics of human inherited disease. J Wiley & Sons, 1995
- [3] Wong, Dominic WS. The ABCs of gene cloning. Chapman & Hall, 1997
- [4] 美国国家科学研究会. 人类基因组的作图与测序. 上海: 上海科学技术出版社, 1990

第二讲 染色体病及其研究进展(1)

染色体数目或结构异常所致的疾病称为染色体病。

染色体是组成细胞核的基本物质,是基因的载体。人类细胞遗传学研究的主要对象是染色体,在1956年以前,由于人类染色体数目未弄清楚,故根本谈不上有可靠的研究资料。1956年美籍华人蒋有兴等查明,人类染色体数目为46条。相继于1960年在丹佛、1963年在伦敦、1966年在芝加哥召开了三次人类细胞遗传学国际会议,制定并不断修改了国际命名体制,方便了国际间的交流。但是,由于方法学上的局限性,在1970年以前的十多年间,世界上仅仅记载了十余种染色体异常。自Caspersson等于1970年发表第一张人类染色体显带照片,以及1971年巴黎国际命名会议以来,现已发现人类染色体数目异常和结构畸变20 000余种,染色体病综合征100余个,其中除携带者和少数性染色体异常者外,智力低下和生长发育迟缓几乎是染色体异常者的共同特征。通过对流产、死产及新生儿和一般人群的调查表明,染色体异常占流产胚胎的50%,占死产婴的8%,占新生儿死亡者的6%,占新生活婴的5%~10%,占一般人群的5%;在一般人群中,平衡易位占1.9%(其中约有15%生育力降低),不平衡易位占0.5%。不平衡易位通常引起严重疾患而夭折,平衡易位由于没有遗传物质的丢失,本人虽不表现畸形,但可以遗传,而且生育染色体异常患者的机率高达50%~100%,给后代带来严重的灾难。因此,在70年代,通过羊水穿刺获取胎儿脱落细胞,进行胎儿染色体分析的宫内诊断技术得以迅速地发展。1981年人类细胞遗传学国际命名委员会发表的人类染色体高分辨550~850条带的模式图即“ISCN(1981)”,标志着人类染色体病的研究进入第三个发展阶段。基于分子遗传学的发展,已使人们达到了能测定基因中单个核苷酸的改变,而能检出染色体微小改变的细胞遗传学高分辨技术则能识别这个基因是否易位到了能异常表达的位置。一些遗传学家认为:细胞遗传学,特别是染色体高分辨技术与染色体绘画(FISH)技术相结合的分子细胞遗传学,已进入了一个研究人类染色体基因定位、人类染色体病图谱、人类肿瘤基因激活原理,从而开展染色体病甚至某些肿瘤等的宫内逆向诊断,以提高人口素质的黄金时代。

一、染色体的结构

组成染色体的基本物质是脱氧核糖核酸(DNA)分子,一条染色体是一个DNA分子,每个DNA分子具两条互补的多核苷酸链,每条多核苷酸链均由腺嘌呤(A)、鸟嘌呤(G)、胞嘧啶(C)、胸腺嘧啶(T)四种单核苷酸所组成,两条互补的多核苷酸链之间,A与T,G与C各借氢键相互配对。

染色体的基本结构单位是核小体(nucleosome),每个核小体是由组蛋白(H2A、H2B、H3

及 H4 各两个分子)组成的八聚体,其外面盘绕1.75圈的 DNA (含 140 碱基对),两个核小体间以 60 个碱基对(bp)的 DNA 双螺旋与组蛋白 H1 形成的细丝相连接,每六个核小体绕成一圈而成空心螺线管,以此为基础,再进一步折叠和多级螺旋化而成为在长度上被压缩了近十万倍的中期染色单体。中期染色体则是由两条碱基排列完全相同的 DNA (姐妹染色单体)链以着丝粒相联所组成的。

二、染色体的遗传

父母传给子女什么东西?父亲提供精子,母亲提供卵子,在母体的输卵管内精卵结合成合子,合子在子宫内分裂、分化发育成胎儿。据此,可以肯定父母的遗传信息必然是由精子和卵子带给子女的。

近代细胞遗传学查明:在人体各类组织中,细胞的各种成分及其结构,除细胞核中的染色体外,随着其组织的功能不同,差别较大,但其细胞核内的染色体则含量是恒定的,在活细胞中它以 DNA 的分子状态存在于细胞核内,在光学显微镜下呈均质的染色质,在电子显微镜下它以松散的螺旋状态满布于细胞核内。在细胞分裂时,它经多级螺旋折叠形成在光学显微镜下可见的 46 条染色体;按其长短、形态进行排列,可分为 7 组 23 对。由于同对的两条染色体一条源自父方,一条源自母方,而且在形态、功能上相同,故称为同源染色体。在 23 对染色体中,有 22 对男女相同的染色体称为常染色体,编号为 1~22 号;一对男女有别者称之为性染色体。女性的两条等长,介于 6~8 号染色体之间,命名为 X 染色体;男性的一条为 X,另一条与 21 号大小相近,命名为 Y 染色体。

仔细观察染色体在亲、子代细胞和亲子代个体之间的遗传,在体细胞的有丝分裂过程中,46 条染色体有规律地复制一次,而后被均等地分给两个子细胞,所以子细胞也保持了 46 条染色体。在性细胞的成熟分裂(减数分裂)中,染色体有规律地复制,并在同源染色体间发生有规律的配对、交换与分离,染色体仅复制一次而细胞连续分裂了两次,从而形成仅含有 23 条染色体的精子或卵子,再通过精、卵结合又恢复 46 条(23 对),从而保持了其亲子代个体间染色体的恒定性。常染色体和性染色体在亲子代之间的遗传是子代的成对染色体,一条来自父方,一条来自母方,父亲的 X 染色体仅能传给女儿,通过女儿传给外孙和外孙女;父亲的 Y 染色体仅能传给儿子,儿子的 X 染色体仅能从母亲来。估计人类约有 10 万个基因,这些基因就是染色体上一段有遗传功能的 DNA 片段,同一条染色体上的基因随着这条染色体而连锁遗传。人类 23 对染色体,在减数分裂中,通过同源染色体间的分离和非同源染色体间的自由组合,则可形成 2^{23} 即 8 388 608 种类型的精子或卵子,精卵随机结合则可形成 $2^{23} \times 2^{23}$ 即 70 368 744 177 664 种不同的合子。以上就是经典遗传学的分离、自由组合、连锁、交换遗传规律的细胞学基础;也就是亲子代之间之所以相像而又千差万别的遗传学根据。

三、研究染色体的方法

细胞从前一次分裂完成到下一次分裂结束,这样一个周而复始的过程称为细胞增殖周期,简称为细胞周期(cell cycle)。在细胞周期中,染色体以不同的形态出现。在细胞间期核中,它以分子状态散布在细胞核内;在进行有丝分裂或减数分裂的细胞中,随着细胞分裂周期从 S 期(DNA 合成期)的末期向 M 期(中期)延续, DNA 有规律地多级螺旋化,从而形成在光学显微镜下能清楚辨认的染色体。因此,人体的任何一种体内或离体的处于旺盛有丝分

裂或减数分裂的细胞群体,按图 2-1 程序进行细胞学处理,都可作为研究人类染色体的材料。

四、染色体畸变

染色体畸变是染色体病形成的根本原因。它包括数目畸变和结构畸变两大类。前者是细胞有丝分裂机制被损伤的结果;后者是染色体断裂和变位重接的产物。在细胞周期的不同时期和个体发育的不同阶段,所致染色体畸变的遗传效应是不完全相同的。

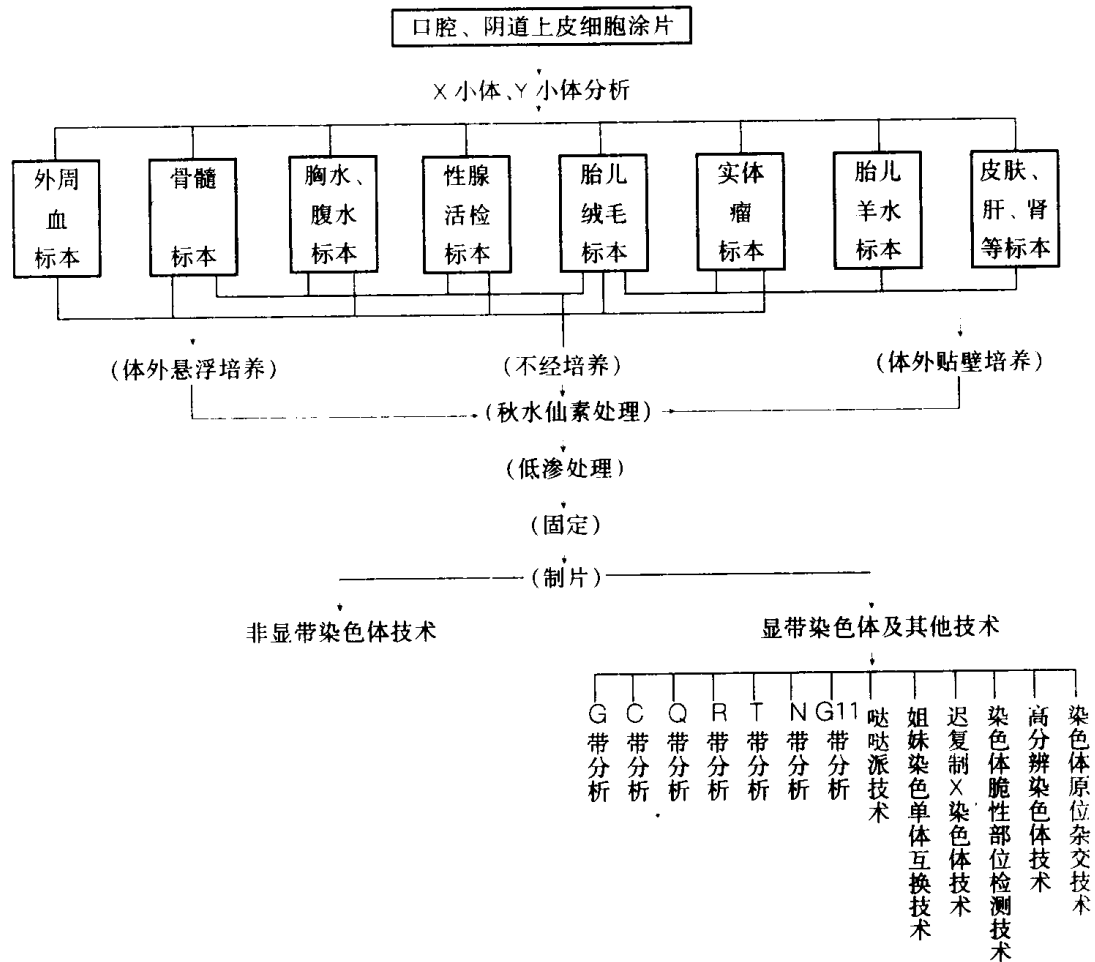


图 2-1 染色体标本制作的基本程序

(一)在细胞周期的不同时期所致染色体畸变的特点

人体细胞分裂包括有丝分裂和减数分裂两种类型,前者是体细胞增殖的唯一方式,后者是配子形成的必经阶段,它们的每个周期均包括 G_1 期(DNA 合成前期)、S 期(DNA 合成期)、 G_2 期(DNA 合成后期)和 M 期(分裂期,包括前、中、后、末 4 期)。在各个时期发生的染色体畸变,既有其相似性,又有其独特性。现分述如下:

1. 有丝分裂

在 G_1 期和早 S 期发生的畸变是染色体型的畸变,而在晚 S 期、 G_2 期和分裂期发生的畸变是染色体单体型的畸变。如果畸变仅涉及到一条染色体或者形成的畸变染色体仅具有一个着丝粒,或者虽有两个或两个以上的着丝粒,但只有一个具活性,其余都是失活的着丝粒,

这些畸变的染色体在细胞有丝分裂中能稳定地传给子细胞，故称为稳定型染色体畸变。无着丝粒片断在有丝分裂的后期因不能定向运动而丢失；具有两个或两个以上有活性的着丝粒的染色体，在有丝分裂的后期，常常形成染色体桥而导致细胞死亡或产生新的畸变，这两种情况统称为非稳定型染色体畸变。

在有丝分裂的中期，由于纺锤体机制被部分破坏，可造成姐妹染色体不分离，就会形成具有三体或单体型的子细胞；若某一分裂中期细胞的纺锤体机制被完全破坏，则可造成所有姐妹染色单体的不分离，从而进一步形成内复制而产生多倍体。

上述改变如发生在体细胞中则影响其子细胞，如发生在性细胞中则可导致减数分裂的异常。

2. 减数分裂

减数分裂与有丝分裂的差别在于前者是一个细胞经过两次连续的分裂而染色体只复制一次、中间经历了一个同源染色体的配对、交换和分离的过程，从而导致不同类型的染色体畸变将产生不同的遗传效应。如相互易位染色体在减数分裂中，经过同源染色体间的配对、交换和分离，不再产生新的结构重排的染色体，这类畸变染色体由于是原发性重排的产物，在减数分裂中不再产生新的重排，故谓之衍生染色体(der)；如插入或倒位染色体在减数分裂中，由于在移位片段和正常位置的相应片段间发生了交换，从而产生了由新的片段组成的染色体，由于它是继发性重排的产物，在减数中发生了新的重排，故叫重组染色体(rec)。

作为衍生染色体的一个例子如 46, XX, t(2;5)(q21;q31)在减数分裂中，通过同源染色体间的配对，将形成四射体(图 2-2)，不论在哪一位点发生互换，通过邻近-1 和邻近-2 分离，以及 12 种可能的 3:1 分离，至少可形成 18 种合子，其中仅 1 种为正常者，1 种为表型正常的易位携带者，其他均为部分三体或部分单体患者，但没有任何 1 种类型出现新的结构重排，故称为衍生染色体，可记述为 der(2)和 der(5)(表 2-1)。

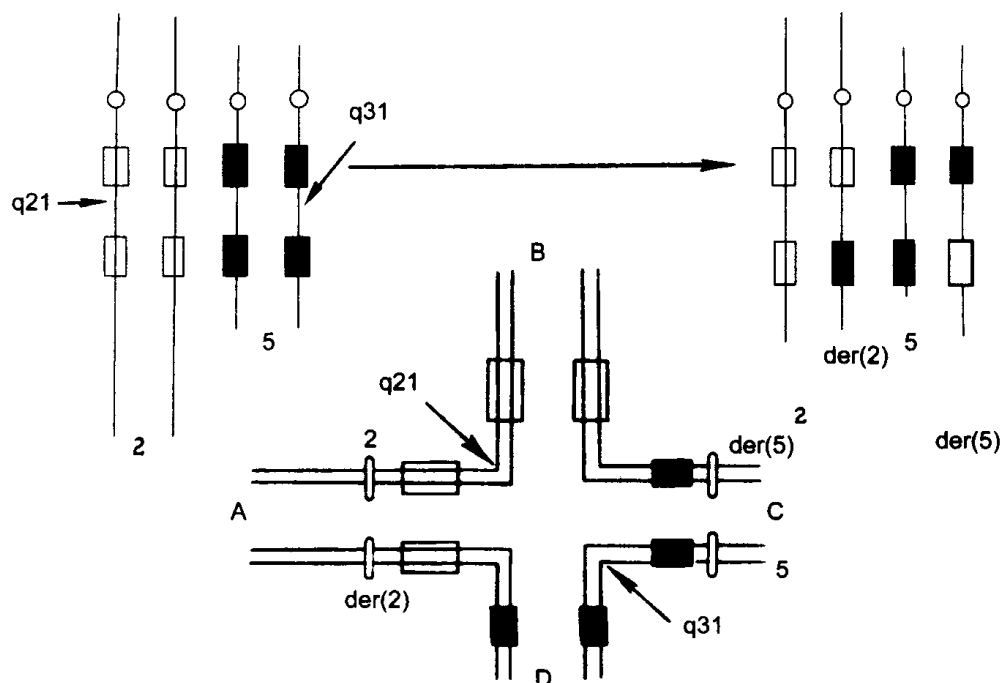


图 2-2 衍生染色体在减数分裂中形成的四射体

表 2-1 相互易位携带者可能产生的 18 种类型的配子及其与正常配子结合可能形成的合子类型

染色体分离后配子类型	与正常配子可能形成的合子核型
AB CB	46, XX 或 XY, -5, + der(5)t(2;5)(q21;q31)
AD CD	46, XX 或 XY, -2, + der(2)t(2;5)(q21;q31)
AB CD	46, XX 或 XY
AD CB	46, XX 或 XY, -2, -5, + der(2), + der(5), t(2;5)(q21;q31)
AB AD	46, XX 或 XY, -5, + der(2)t(2;5)(q21;q31)
CB CD	46, XX 或 XY, -2, + der(5)t(2;5)(q21;q31)
AB AB ¹⁾	46, XX 或 XY, +2, -5
CD CD ¹⁾	46, XX 或 XY, -2, +5
CB CB ¹⁾	46, XX 或 XY, -2, -5, + 2der(5)t(2;5)(q21;q31)
AD AD ¹⁾	46, XX 或 XY, -2, -5, + 2der(2)t(2;5)(q21;q31)
AB CB CD	47, XX 或 XY, + der(5)t(2;5)(q21;q31)
AD	45, XX 或 XY, -2, -5, + der(2)t(2;5)(q21;q31)
CB CD AD	47, XX 或 XY, -2, + der(2), + der(5), t(2;5)(q21;q31)
AB	45, XX 或 XY, -5
CD AD AB	47, XX 或 XY, + der(2)t(2;5)(q21;q31)
CB	45, XX 或 XY, -2, -5, + der(5)t(2;5)(q21;q31)
AD AB CB	47, XX 或 XY, -5, + der(2), + der(5)t(2;5)(q21;q31)
CD	45, XX 或 XY, -2

1) 要产生这 4 种类型的配子, 必须在着丝粒和交换点之间的中间节段中发生交换。

作为倒位染色体的一个例子如 46, XX, inv(2)(p21q31), 在减数分裂中通过同源染色体间的配对, 将形成倒位环(图 2-3), 如果在倒位环内发生交换, 则在一个重组染色体中将导致 2p 重复, rec(2)dup p, inv(2)(p21q31), 而在另一重组染色体中, 将导致 2q 重复, rec(2), dup q, inv(2)(p21q31), 通过与正常配子结合则可形成 46, XX, rec(2), dup p, inv(2)(p21 q31) 和 46, XX, rec(2), 图 2-3 示倒位染色体在减数分裂中形成的倒位环 dup q, inv(2)(p21q31) 的合子, 在第 1 种核型中表示 2pter 到 2p21 的重复和从 2q31 到 2qter 的缺失, 在第 2 种核型中, 表明从 2q31 到 2qter 重复和从 2pter 到 2p21 的缺失。由于它是继发性重排的产物, 故称为重组染色体, 记述为 rec(2)。

(二) 在个体发育的不同阶段所导致染色体畸变的特点

人类个体发育从精卵发生开始可大致分为配子形成期、合子期、胚胎期、胎儿期及出生后各期等, 每个时期产生的染色

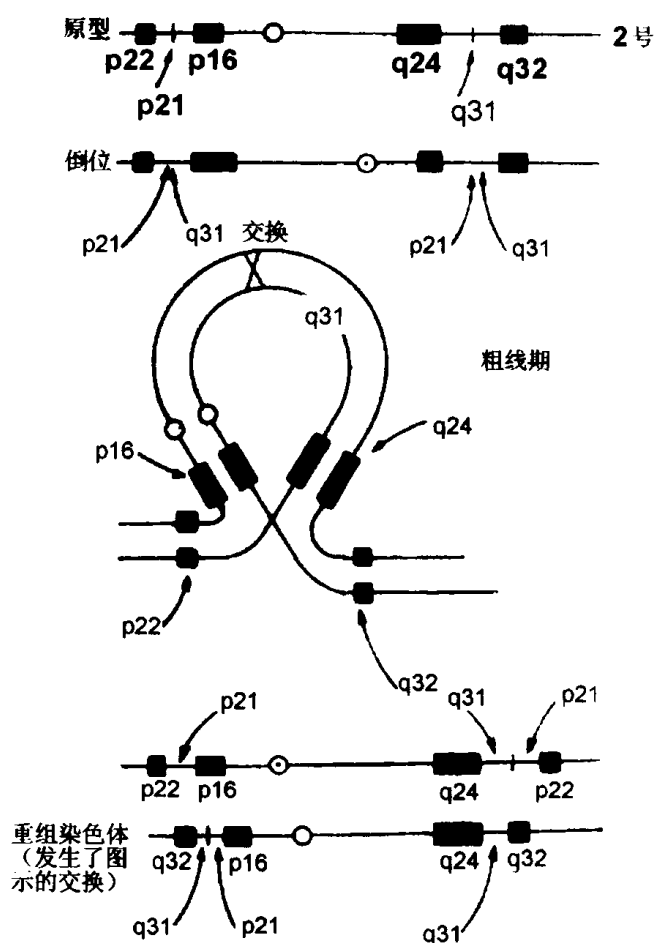


图 2-3 倒位环模式图

体畸变所致之遗传效应不尽相同。一般,在配子形成期或合子期(约在受精后的 24 小时内)发生的染色体畸变,将导致整个个体带有畸变的染色体,即形成各种常见的单体型、三体型、部分单体型、部分三体型、三倍体等各种类型的染色体病患者;在胚胎早期的卵裂及桑椹胚期(约受精后 3~4 天内)发生的染色体畸变,将导致某一个体发育为含有不同百分比核型的嵌合体的染色体病患者。其后由于三个胚层的分化,某一胚层发生的染色体畸变,将累及由此胚层发育而来的有关器官系统,目前研究甚少;胎儿出生后,在各种内外环境因素的影响下,可以导致各种类型的体细胞突变,它常常通过有丝分裂传给其子细胞而形成克隆,由于它往往导致恶性肿瘤的发生,故研究得较为深入。

五、从基因、染色体到临床特征

临床所见往往是智力低下、发育不良、五官不正、四肢畸形、心肾功能不全或各器官系统的异常所导致的病状和体征,形成这些临床特征的基础,常常可追溯到细胞形态的异常、代谢的异常、蛋白质的异常、酶的异常,等等,但其本质都是基因或染色体的异常,甚至包括某些传染病在内,如疟疾的易感性等,均可以在基因水平上找到答案。

分子遗传学证明,基因是组成染色体的 DNA 分子的一个区段, DNA 链上的核苷酸有一定的序列,这序列就携带了遗传信息。在分裂型细胞中, DNA 的双链拆开,以每条单链为模板,按照核苷酸互补原则进行半保留复制;在代谢型细胞中,以 DNA 双链中的一条(基因链)为模板,按互补的原则转录为信使 RNA (mRNA), mRNA 与 DNA 仅有一种核苷酸的差别(即尿嘧啶 U 代替了胸腺嘧啶 T),成熟的 mRNA 由细胞核进入细胞质中,在核糖体上指导合成相应的多肽,即以 mRNA 分子的三个相邻的核苷酸(又称遗传密码子)决定一个氨基酸的方式翻译为多肽,再加工成蛋白质或酶。各种氨基酸的遗传密码子见表 2-2。

表 2-2 各种氨基酸的 mRNA 三联体密码子

UUU } 苯丙氨酸	UCU } 丝氨酸	GCU } 丙氨酸	GAA } 谷氨酸
UUC } 苯丙氨酸	UCC } 丝氨酸	GCC } 丙氨酸	GAG } 谷氨酸
UUA } 亮氨酸	UCA } 丝氨酸	GCA } 丙氨酸	UGU } 半胱氨酸
UUG } 亮氨酸	UCG } 丝氨酸	GCG } 丙氨酸	UGC } 半胱氨酸
CUU } 亮氨酸	AGU } 精氨酸	UAA } 终点	UGA } 终点
CUC } 亮氨酸	AGC } 精氨酸	UAG } 终点	UGG } 色氨酸
CUA } 亮氨酸	CCU } 脯氨酸	CAU } 组氨酸	CGU } 精氨酸
CUG } 亮氨酸	CCC } 脯氨酸	CAC } 组氨酸	CGC } 精氨酸
AUU } 异亮氨酸	CCA } 脯氨酸	CAA } 谷氨酰胺	CGA } 精氨酸
AUC } 异亮氨酸	CCG } 脯氨酸	CAG } 谷氨酰胺	CGG } 精氨酸
AUA } 甲硫氨酸	ACU } 苏氨酸	AAU } 天冬酰胺	AGA } 精氨酸
AUG } + 合成起步信号	ACC } 苏氨酸	AAC } 天冬酰胺	AGG } 精氨酸
GUU } 缬氨酸	ACA } 苏氨酸	AAA } 赖氨酸	GGU } 甘氨酸
GUC } 缬氨酸	ACG } 苏氨酸	AAG } 赖氨酸	GGC } 甘氨酸
GUA } 缬氨酸	UAU } 酪氨酸	GAU } 天冬氨酸	GGA } 甘氨酸
GUG } 缬氨酸	UAC } 酪氨酸	GAC } 天冬氨酸	GGG } 甘氨酸

1) 如在起始位点上则是甲硫氨酸的符号, 否则照旧。

由此可知,人体内蛋白质或酶是在基因所携带的遗传信息的控制下合成的。通俗地说,有什么样的基因,就决定产生什么样的蛋白质或酶,就导致出现什么样的遗传性状。一旦DNA发生了改变,即遗传学上所说的基因突变,蛋白质或酶亦随之发生改变,遗传性状也就会跟着发生某种相应的变化,许多分子病和遗传代谢病就是如此。例如,人类有一种较严重的遗传疾病叫做镰形细胞贫血症,正常人的红细胞好似一个双凹的圆碟子,而这种病人的细胞在缺氧时像一把镰刀。患者严重贫血,发育不良,容易感染各种疾病,一般都在成年之前死亡。本病主要分布在非洲,也散布于地中海区,由于杂合子对疟疾有较大的抗性,故与疟疾的地理分布相一致。早在40年代末,人们就知道这种疾病是一种发生在血红蛋白分子上的遗传病。血红蛋白分子(HbA)是由4条多肽链组成:两条 α 链,每条链含有141个氨基酸,决定该链的基因位于16号染色体;两条 β 链,每条链含有146个氨基酸,决定该链的基因位于11号染色体上。镰形细胞的血红蛋白(HbS)和正常人的血红蛋白在组成上有什么不同呢?原来正常人的 β 链从氨基开始的一端数起,第6位上的氨基酸是带有负电荷的极性侧链的亲水性的谷氨酸,而镰形细胞贫血症患者在这个位置上的氨基酸变成了不带电荷且具有非极性侧链的疏水性的缬氨酸。也就是说,在血红蛋白分子的4条多肽链的574个氨基酸中,只有2个氨基酸的差别,即2个疏水的缬氨酸替代了2个亲水的谷氨酸。为什么会出现这2个氨基酸的替代呢?原来是位于11号染色体上的决定 β 链mRNA形成的DNA分子(基因链)发生了一个由核苷酸T被A替换的突变,即CTT变成了CAT。基因的突变导致mRNA中这个有关的遗传密码子中的第2个核苷酸由A变成了U,即GAA变成了GUA,从而使血红蛋白的 β 链氨基端第6位上的谷氨酸变成缬氨酸(表2-3)。

表 2-3 正常血红蛋白与镰形血红蛋白的比较

β 链氨基酸序列	...6—7.....26.....63.....
HbA 的氨基酸	...谷——谷.....谷.....组.....
HbA 的密码子	...GAA...GAA.....GAA.....GAU.....
HbA 的基因	...CTT...CTT.....CTT.....CTA.....
HbS 的氨基酸	...缬.....谷.....谷.....组.....
HbS 的密码子	...GUA...GAA.....GAA.....GAU.....
HbS 的基因	...CAT...CTT.....CTT.....CTA.....

由于被2个缬氨酸替代了谷氨酸的这种异常血红蛋白,在稀氧状态下,比正常血红蛋白的溶解度低5倍,于是在氧分压低的毛细血管区,HbS的溶解大大降低而形成管状结构和凝胶化,导致红细胞变形成镰刀状,即“镰变”。镰变是可逆的,氧张力增加时,红细胞能够恢复。但如果镰变时间延长,细胞膜便遭受到不可逆的损害,这种镰形细胞就被网状内皮细胞从循环系统中清除掉,造成严重的贫血、虚弱和疲乏无力。由于异常的红细胞变得很僵硬,不能通过微循环,又加上凝胶化的结构使血液粘滞性增加,致使微循环阻塞,局部组织缺血、缺氧甚至坏死,从而出现急性骨关节痛和腹部疼痛;根据血管闭塞的部位不同,临床上可出现不同器官的病变。如果涉及肾血管,可引起肾功能衰竭;如果涉及心脏的血管,可引起心力衰竭;如果累及肝、脾、神经,则可造成肝、脾、神经损伤,等等。这就告诉我们,虽然基因(DNA分子)并没有携带红细胞的形态及贫血症等性状,而只带有合成有关蛋白质的遗传信息,信息变了,翻译出来的蛋白质结构也就改变了,由此引起的蛋白质构型和功能发生改变,从而破坏了人体代谢调节的正常活动,造成了一系列的临床症状——腹痛、四肢疼痛、血尿、肾衰、心衰、脾大、梗死,乃至脑血管意外,等等(图2-4)。