Y-DNA检测在刑事侦查中的应用

尹国兴 著

上海交通大學出版社
SHANGHAI JIAO TONG UNIVERSITY PRESS

前　言

2016年8月26日，甘肃省白银市警方抓获了28年前强奸杀人系列案件的犯罪嫌疑人高某勇，指控其在14年内以残忍手段强奸并杀害了甘肃白银和内蒙古包头的11名女性。由于过去刑侦技术落后，疑犯在逍遥法外28年后方才落网。这次成功破案的关键，是使用了Y-DNA姓氏检测的新技术。该案的侦破，使得这项在民间一直默默无闻的医学检验技术一战成名，受到社会广泛关注，牢牢占据了媒体的重要版面。

通过检测人类Y染色体DNA侦破刑事案件是一项生物技术，起初来源于亲子鉴定，是亲子鉴定技术的精简改编。Y染色体DNA遵循严格的父系遗传规律，而姓氏的传递也与父系遗传有关，是与父系遗传相偶联的历史文化现象：父亲Y染色体传递给子代的同时，也将自己的姓氏传递给子代。Y-DNA破案法，即以Y染色体DNA侦破案件的方法，强化了对Y-DNA的检测，通过与数据库中已知姓氏家系的同源对比，找到待测样本的家族群体，首先确定嫌疑人的姓氏，然后顺藤摸瓜，筛排家系中的男性成员，从而锁定目标嫌疑人。

就甘肃白银这起案件而言，疑犯有作案11起的案底。一方面，警方在作案现场提取了案犯留下的含有遗传物质的身体遗留物，获取其Y-DNA信息，然后录入违法人员数据库中，以备进一步查找。另一方面，与他同姓的远房堂叔因行贿获罪被监视居住，警方因此采集此人的血样。此人遗传数据录入违法犯罪人员Y-DNA数据库中，经初步对比，发现与该案疑犯的Y-DNA信息相符合。这一结果表明，案犯与此人有相同的Y染色体遗传，是同一家系的男性成员，所以案犯姓氏较为明确。

为了进一步确认，样本又送到甘肃省公安厅做复核检验，结果符合前期检验，从而进一步确定案犯姓氏。随后，警方收集罗列了这一家系所有男性成员的Y-DNA信息，经筛排分析，确认高某勇具备作案条件，为该案的目标犯罪嫌疑人。高某勇归案后，对其本人指纹和DNA与案发现场指纹和DNA检验对比，

确认来自同一个体。经审讯，案犯对犯罪事实供认不讳，多年悬案由此得以侦破。

Y-DNA姓氏检测，无须针对整个Y染色体，只需选取其中一些特异性的片段作为标记物。DNA中有一种特异性的碱基序列称短串联重复顺序（short tandem repeat，STR），Y染色体上的STR称Y-STR。姓氏检测一般都是通过检测Y-STR实现。

目前中国约有15亿人，其中汉族数量约12.5亿，可谓人口众多。但中国人Y-DNA的多态性是有限的。若以Y-DNA区分，中国约只有100万个姓氏家系。换而言之，中国人口的绝对数量虽然不断增加，但Y-DNA的多态性并不会相应增加。Y-DNA的多态性分析，是通过电泳实现的。每个遗传家系都有固定的Y-STR电泳谱带，相当于家系的“条形码”，而这种“条形码”又与每个家系的姓氏相对应。通过扫描和检验这种“条形码”，可以确定男性个体的家族归属，从而做出有效的姓氏检测。

姓氏检测包括直接检测和间接检测。直接检测是通过与被测样本有明确亲缘关系的遗传样本的比较直接得出姓氏结论。例如，某男子在作案现场留下了毛发、血液、指甲、皮屑等遗传物质，但嫌疑人已逃之夭夭。此时，可寻找与嫌疑人有相同Y-DNA渊源的亲属，如祖、伯、父、叔、子、侄、孙等近亲或远亲，进行Y-STR的同源对比，直接确定（或排除）嫌疑人的姓氏。间接检测是只有被检测的遗传样本，没有直接可比的亲缘材料，而仅与数据库中的Y-DNA姓氏信息进行对比，此时获得的是姓氏的概率信息。例如，某男子的遗传样本与数据库中的信息对比，可知事主可能的姓氏。甘肃白银系列杀人案的侦破，是通过直接检测疑犯与其远房堂叔的Y-DNA，对比确定疑犯身份的。

甘肃白银系列杀人案的破获，是Y-DNA破案法在刑侦领域的一次具体应用，为未来破案提供了积极的借鉴和参考。既往亲子鉴定和DNA鉴定，都是作为司法审判的证据在法庭上使用，对公安机关侦破案件没有帮助。因为Y-DNA父系遗传的特性，通过Y-DNA检测，可以确定留在犯罪现场的人类遗传物质的主人性别；如果来自男性，可以进一步检测出姓氏，确定其家系来源。这样便缩小了排查范围，为破案提供便利条件。

Y-DNA破案法将姓氏检测应用在刑事侦查领域，大大降低了破案成本，节省了警力，提高了公安机关的工作效率。而且，姓氏检测操作简单、成本低廉，Y-DNA数据库很适合基层普及，是一项符合中国国情的实用技术体系。Y-DNA姓氏检测技术的普及，将改变中国刑警的工作状态，成为中国科技史上闪

光的一页。

Y-DNA 检测可以明确样本间的血缘关系，是一项成熟的生物技术。1998年，这项技术证明了一黑人男子是美国第三届总统托马斯·杰弗逊家族的直系后裔，显示这名男子的生父是杰弗逊家族中的某一成年男性。这是通过不同个体 Y 染色体间的两两比较得出的结论。

但是，这还不能满足姓氏检测的需要。因为在实际操作中，往往只有待测一方的生物检材或遗传标记信息，而可供比较的另一方信息难以寻获。因此，使用工程技术手段，不断积累和升级 Y-DNA 数据库必不可少。待测样本的数据信息与数据库中的已有信息逐一对比，寻找同源姓氏家系，以此确定姓氏和家系。姓氏检测乃至检索查询的新功能，是在数据的积累中得以形成并不断强化的，与原来的亲子鉴定已然完全不同。

一个姓氏家系的 Y-DNA 好比是一个战士，姓氏家系的 Y-DNA 组成数据库好比是一支军队，一个战士和一支军队的功能显然是不一样的。Y-DNA 数据库的功能比单个 Y-DNA 要强大得多，姓氏检测的顺利进行，基于 Y-DNA 数据库的建设和完善。

其实，姓氏检测技术的真正作用不局限于如何节约警力、提高工作效率，还在于成效明显高于其他刑侦方法（例如指纹鉴定、疑犯模拟画像），可用来侦破大案、要案和疑难悬案。甘肃白银系列杀人案动用了很大警力，采用了各种办法，均不能奏效，在采用该新技术后迅速侦破，该技术的优势和特色显得十分突出。

基层 Y-DNA 数据库的建设和姓氏检测技术，是破案的有力工具。如果 Y-DNA 数据库已经覆盖了足够多的家系，凡男性参与的各类犯罪，其破案率理论上可达到 100%。而且，Y-DNA 破案法的普及可有效预防犯罪，对犯罪行为形成有效威慑，使一些有计划、有预谋的犯罪不再发生。在普及了 Y-DNA 检测和 Y-DNA 数据库的某些地区，有证据显示，恶性案件的发生率已经呈现出“断崖式”下降的趋势，社会治安有了明显改善。

Y-DNA 破案法和 Y-DNA 数据库的普及，将为国家的长治久安提供有效的技术支撑，实现更快地破大案，更多地破小案，更准地办好案，更好地控发案。目前，中国各地的 Y-DNA 数据库建设已成流行趋势，基层干警和公众有了解和掌握这项技术的强烈需求。这是提升基层干警技术侦查水平的良好契机，也是对公众很好的科普教育。本书结合真实案例，详细描述了 Y-DNA 检测的原理和应用，消除读者的疑问和困惑，给读者启发并带去长远的福祉。

目　　录

第一章　Y-DNA 概述 …… 1
第二章　Y-DNA 与姓氏 …… 18
第三章　Y-DNA 破案法的科学原理 …… 27
第四章　Y-DNA 数据库 …… 38
第五章　Y-DNA 检测的技术体系 …… 69
第六章　常用的 Y-STR 基因座 …… 100
第七章　Y-DNA 破案的优势 …… 131
第八章　Y-DNA 破案法应用实例 …… 149
第九章　Y-DNA 检测的科学应用 …… 210
第十章　Y-DNA 问答 …… 225
参考文献 …… 241
索引 …… 250
后记 …… 254

第一章

Y-DNA 概述

1. 细胞、细胞核、蛋白质与基因

细胞是生命活动的基本单元，人体结构是由各种细胞组建起来的。功能各异的人体细胞构成了组织，分为上皮组织、肌肉组织、结缔组织、神经组织。例如，表皮属于上皮组织，心肌属于肌肉组织，骨骼属于结缔组织。各种组织结合在一起，形成了人体的各个器官，从而构成了复杂的人体。

人体细胞包括细胞膜、细胞质和细胞核。细胞质中有多种细胞器，包括线粒体、高尔基体、核糖体、内质网和中心粒等（见图 1－1）。作为遗传物质的核酸，主要存在于细胞核中。核酸是遗传信息的载体，包括核糖核酸（RNA）和脱氧核糖核酸（DNA）。其中，编码某一蛋白质的 DNA 链即是一个基因，是有遗传效应的 DNA。基因在染色体上有具体的位置和编号，称“基因座”或“位点”。

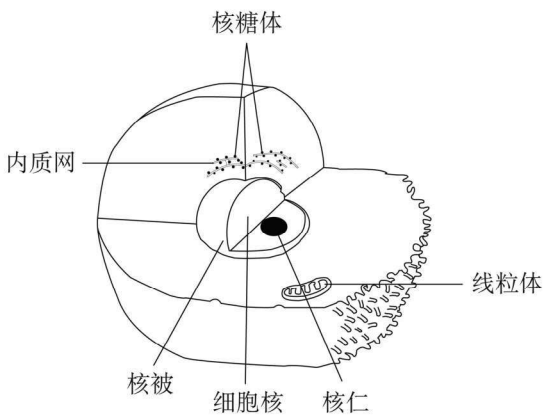


图 1－1　人体细胞的基本结构

DNA 是生物体遗传信息的载体，由 C、H、O、N、P 五种元素组成。DNA 分子的基本单位是脱氧核苷酸，由一分子磷酸、一分子脱氧核糖和一分子含氮碱基构成。DNA 分子分为三级结构。DNA 链以磷酸糖为骨架，由 A(腺嘌呤)、G(鸟嘌呤)、C(胞嘧啶)、T(胸腺嘧啶)四种碱基连贯而成，这四种碱基连贯而成的单链，即 DNA 的一级结构。1953 年，美国科学家 J. Watson 和 F. Crick 发现了 DNA 双螺旋结构，是按照碱基的互补配对原则形成的双链，即 DNA 的二级结构。其中，G 与 C 配对，中间以 3 个氢键相连；A 与 T 配对，中间以 2 个氢键相连。以双螺旋结构为基础，进一步盘曲折叠，形成更加紧密的超螺旋结构，即 DNA 的三级结构(见图 1-2)。

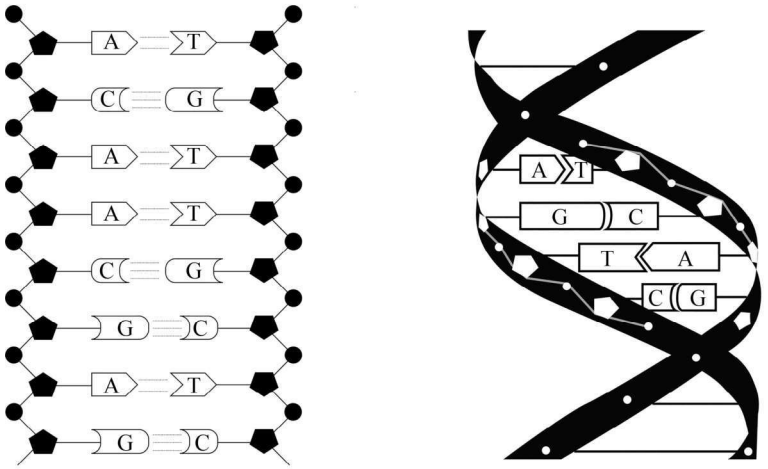


图 1-2　DNA 链平面结构和立体的双螺旋结构

2. 蛋白质——生命的本质

蛋白质是细胞的功能成分，是细胞结构和代谢的主体。人体和细胞内的各种生命活动，都与蛋白质的功能息息相关。

东汉医学家张道陵提出了元白之论，是世界上最早的蛋白质理论。他把蛋白质置于突出的地位，用于阐释生命的本质。在其传世著作《老子想尔注》中提到："魄，白也，故精白，与元同色"。意思是说，人体是由蛋白质构成的，所以精液也是蛋白质，与乳汁有一样的性质。魄，体魄，机体；元，乳汁，蛋白质；色，颜色，性质。张道陵的元白之论开启了人类认识生命现象和本质的历程。

迄今为止，蛋白质的研究已经十分细致和深入。氨基酸是蛋白质的组成单

位，分为必需氨基酸和非必需氨基酸。必需氨基酸在人体内部不能合成，只能依靠对外摄取；非必需氨基酸在体内可以合成。氨基酸与氨基酸脱水缩合形成二肽，进而形成多肽。多肽链是蛋白质的基本结构（见图 1－3）。

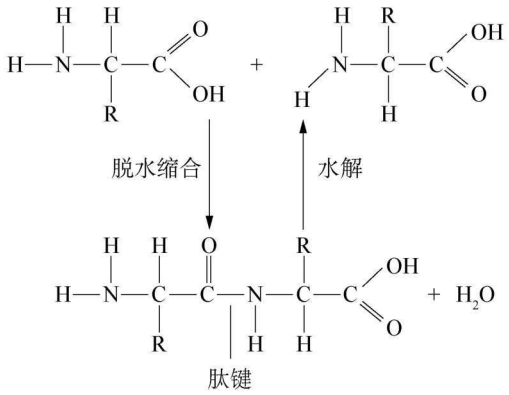


图 1－3　氨基酸与肽键

蛋白质分为四级结构。一级结构是由氨基酸组成的多肽链，其肽键顺序由基因上遗传密码的排列顺序所决定。各种氨基酸按遗传密码的顺序，通过肽键连接起来，成为多肽链。蛋白质的一级结构决定了蛋白质的二级、三级和四级结构。二级结构是指多肽链中主链原子的局部空间排布即构象，包括 α－螺旋、β－片层、β－转角和无规则卷曲。在二级结构的基础上，蛋白质再进一步盘曲或折叠形成具有一定规律的三级结构。具有 2 条或 2 条以上独立三级结构的多肽链结合在一起，此为蛋白质的四级结构。

3. 染色质、染色体与 Y 染色体

细胞在分裂期，DNA 聚集在一起，与蛋白质一起组成了复合体，即染色体。DNA 一般呈酸性，可以与碱性染料结合被染色，所以称染色体。染色体一般为分散状态，称染色质。在细胞的有丝分裂期，染色质发生卷曲、折叠，成为在光学显微镜下可见的染色体。

染色质是遗传物质的载体。1879 年，W. Flemming 提出“染色质”一词，用以描述细胞核中能被碱性染料强烈着色的物质。1888 年，Waldeyer 正式提出染色体的命名。经过多年研究，人们逐渐认识到，染色质和染色体是在细胞周期不同阶段可以相互转变的形态结构。

染色体的中心区域是着丝粒，控制着细胞的有丝分裂。以着丝粒为中心，从

中心到两边叫“臂”，短臂记为“p”，长臂记为“q”，终端叫端粒。p(Yp)、q(Yq)又细分为区、带等。染色质中的蛋白质为碱性蛋白，包括组蛋白和非组蛋白。组蛋白富含精氨酸和赖氨酸，能和带有负电荷的 DNA 结合。非组蛋白种类很多，包括 DNA 聚合酶、RNA 聚合酶等。

人类体细胞一般含有 46 条染色体，其中 X 染色体和 Y 染色体为性染色体，决定了男女性别的差异。其中，1 号染色体最长，22 号染色体最短。Y 染色体为男性所特有，是第 3 小的近端着丝粒染色体，仅大于 21 号和 22 号染色体(见图 1-4)。Y 染色体有 2 个遗传功能不同的区域：拟常染色体区和 Y 特异区。5%的区域属于拟常染色体区，位于 Y 染色体的两端，在减数分裂的时候与 X 染色体交换重组；其余 95%的区域为 Y 特异区，不发生重组，遵循严格的父系遗传规律。

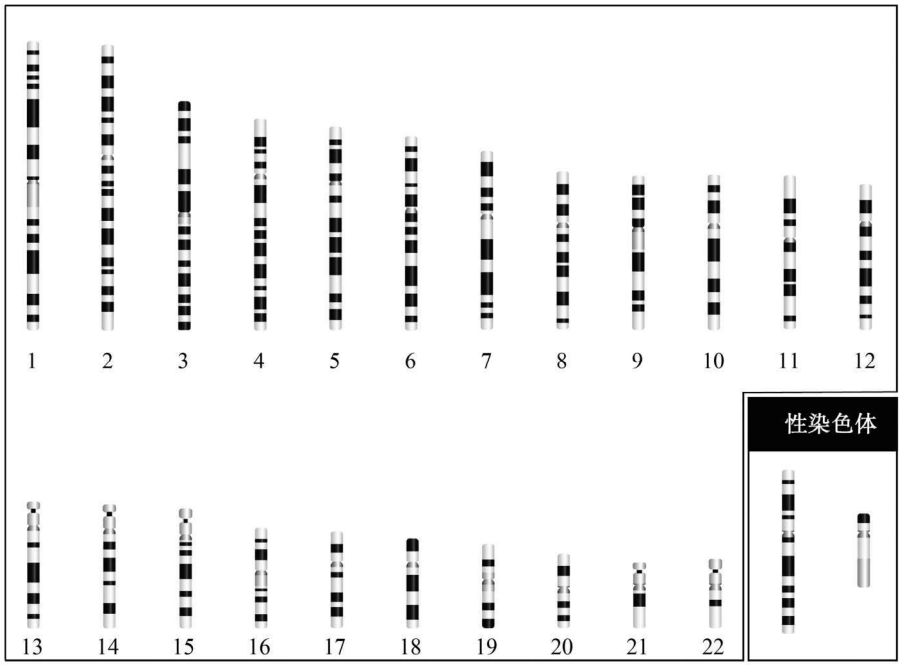


图 1-4　人类染色体

4. Y 染色体的遗传特性

X 染色体和 Y 染色体为同源染色体，是人类的性染色体。有学者认为，两个性染色体 3 亿年前起源于一对常染色体。人类的 Y 染色体非重组区呈单倍体父系遗传，除了拟常染色体区外，Y 染色体在减数分裂期间不与 X 染色体进

行重组交换，只能从父亲遗传给儿子。Y 染色体传男不传女，因此，Y 染色体控制的性状只在男性后代中表现。

Y 染色体有以下区、带：短臂和长臂各只有 1 个区。短臂只有 1 个带(Yp11)，分为 3 个亚带(Yp11.1，Yp11.2，Yp11.3)，Yp11.3 又分为 2 个亚亚带(Yp11.31，Yp11.32)。长臂分为 2 个带(Yq11，Yq12)，Yq11 分为 2 个亚带(Yq11.1，Yq11.2)，Yq11.2 分为 3 个亚亚带(Yq11.21，Yq11.22，Yq11.23)，还有学者将 Yq11.21 再分为 2 个带(见图 1-5)。

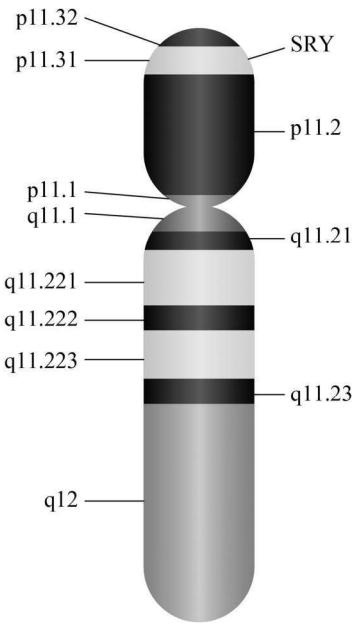


图 1-5　Y 染色体的分区和分带

Y 染色体上有编码睾丸决定因子的基因，决定了男性特征。胚胎之初，生殖腺的发育男女是一致的，18 周后性别决定区 Y 基因(SRY)控制性别分化。SRY 的作用是调节性腺的发育。如果不存在 SRY 基因，性腺发育为卵巢，产生雌性激素，促进发育形成女性生殖系统；如果存在 SRY 基因，性腺发育为睾丸，产生雄性激素，促进发育形成男性生殖系统。

Y 染色体上没有生命必需的基因，女性没有 Y 染色体，也可以正常生活。从功能的角度来讲，所有正常男性的 Y 染色体都是一样的，Y 染色体各异性应与功能无关。

Y 染色体并不是一成不变的，存在一定的变异性。父系遗传中世代积累的突变，经过漫长的年代，形成了 Y 染色体特异区的 DNA 多态性，包括长度多态性和序列多态性。随着越来越多的 Y 染色体上多态性标记被发现和应用，Y 染色体在父系鉴定和群体遗传研究方面显示了突出的价值。

5. DNA 编码区与非编码区

DNA 上有编码区和非编码区。基因即编码区，是蛋白质的遗传密码所在，可以表达为蛋白质。一个基因的长度从几千碱基到几万碱基数量不等。人类只有不到 3 万个基因用于编码蛋白质。DNA 非编码区不能指导蛋白质合成，其中有大量的重复顺序，表现出长度多态性和序列多态性。序列中的单核苷酸多态性(SNPs)，在人类基因组中分布广泛，Y-SNPs 是鉴定父系遗传重要的标志点。

非编码 DNA 不能转录为相应的信使 RNA(mRNA),也不能进行蛋白质的翻译,但非编码区是不可缺少的。因为非编码区上有与 RNA 聚合酶的结合位点,具有调控作用,启动子、终止子位于非编码区。这些是遗传信息转录过程中不可或缺的,对蛋白质的功能状态起着调控作用。

Y 染色体与父系遗传有关,有与性别有关的基因序列。Y 染色体上约含 5.9 万 kb 的 DNA,其中 60%位于其长臂的异染色体区 Yq12,其短臂约含 1.3 kb DNA。预测 Y 染色体上共有 7 117 个基因,已克隆出的基因有 200 多个。

6. 查伽夫法则

查伽夫法则,即 DNA 中 A∶T=G∶C=1∶1 的法则。20 世纪 40 年代末,Erwin Chargaff 等通过人、猪、牛、羊、细菌和酵母等多种 DNA 样品水解物的分析,发现 DNA 样品中的碱基组成有一定的规律性和定量关系。

Chargaff 等的结果表明,在不同生物的 DNA 之间,4 种脱氧核苷酸 A、T、C、G 的数量和比例不相同,但无论哪种物质的 DNA,嘌呤碱基总数等于嘧啶碱基总数,都有 A=T 和 G=C,(A+G)/(T+C)=1,(A+C)/(T+G)=1。这被称为 DNA 化学组成的查伽夫法则,也就是后来确定的双链 DNA 的碱基互补配对原则。1952 年,Chargaff 向 J. Watson 和 F. Crick 讲述了这一规则,从而确定了 DNA 双螺旋结构的基础理论(见图 1-6)。

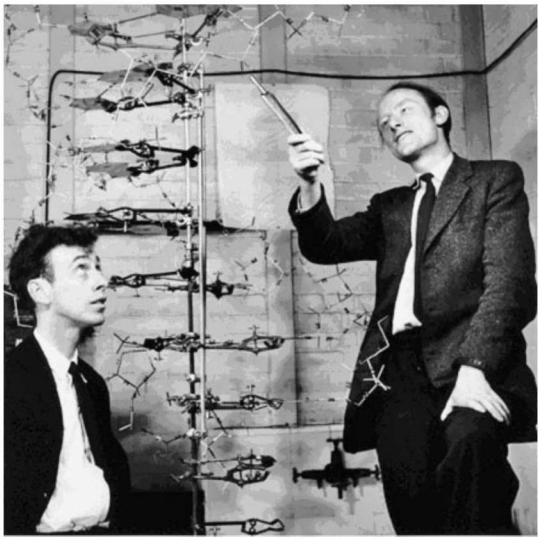


图 1-6 J. Watson 和 F. Crick 发现 DNA 双螺旋结构

7. “废物 DNA”

“废物 DNA”，或称垃圾 DNA，是指细胞内不合成蛋白质的那部分 DNA。人类基因组中包含 30 亿个碱基序列，编码蛋白质的 DNA（外显子）只占全序列的 3%左右，与编码基因有关的序列（内含子）加在一起只占 5%，剩余部分占了 95%，与蛋白质的合成没有直接、明显的关系。

人类基因组 DNA 中大量的重复序列，因为不表达蛋白质或 RNA 产物，一度被称为“废物 DNA”。这些 DNA 其实有重要的用途，除了维护染色体正常的形态结构，还有重要的辅助功能。所以，废物 DNA 是生命活动不可或缺的细胞成分，并不是真正的“废物”。

Y-DNA 中绝大部分为非编码区，表现为父系单倍体遗传。其中有 Y-STR、Y-SNPs、Alu 元件、InDel 元件、DNA 甲基化表观遗传等，这些是用于家系识别和个体识别的重要遗传标志点。

8. 转录、逆转录与翻译

DNA 是双螺旋结构，是遗传信息的载体。mRNA 转录遗传信息，将双链 DNA 信息转录成单链 mRNA 信息，经转运 RNA（tRNA）翻译成多肽链，进而完成蛋白质的合成。遗传信息从核酸到蛋白质的过程，F. Crick 称之为“中心法则”（见图 1-7）。

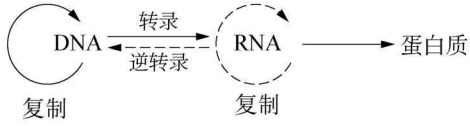


图 1-7　遗传信息传递的中心法则

DNA 与 mRNA 除了双链和单链的区别以外，在结构上也有差异。在核糖的第 2 个碳原子处前者是 H，后者是 OH，碱基也变换了 1 个，即 DNA 中的胸腺嘧啶 T 在 RNA 中被尿嘧啶 U 所取代。两种碱基的差别很小，尿嘧啶 U 比胸腺嘧啶 T 仅多了 1 个甲基。

转录是在由 RNA 聚合酶和辅助因子组成的转录复合物的催化下，从双链 DNA 分子中复制生物信息生成 1 条 RNA 链的过程。在低级生命体（如病毒中），还有逆转录存在。以 RNA 为模板合成 DNA，通常这与转录过程中遗传信息从 DNA 到 RNA 的方向相反，故称为逆转录。

mRNA 由起始密码子 AMG 开始，每 3 个核苷酸组成的三联体构成 1 个密码子，64 个密码子编码 20 个氨基酸。tRNA 为三叶草形，在蛋白质合成中处于关键地位，不但为每个三联子密码翻译成氨基酸提供了接合体，还为准确无误地将所需氨基酸运送到核糖体上提供了载体。

在蛋白质合成期间，存在于 mRNA 上代表一个多肽的核苷酸序列翻译为多肽链序列。tRNA 能够与核糖体的 P 位点、A 位点结合，其顶端突起部位通过密码子与反密码子的配对与 mRNA 相结合，而其 3'末端恰好将所转运的氨基酸送到正在延伸的多肽上。

Y-DNA 编码区有特异性的序列，可表达为功能蛋白质，如 Y-GATA-H4、Y-GATA-C4 等，是家系鉴定中经常使用的重要标志点。

9. 环境、基因调节与基因突变

基因调节是指蛋白质对基因状态的调控，是生物适应环境变化的一种内在调节机制。例如外源性胰岛素介入，内源性胰岛素的表达就会被关闭；外源性吗啡进入体内，会对内啡肽的表达产生抑制。但这是一种功能调节，不会遗传给子代。

表观遗传是指 DNA 序列不发生变化，但基因表达发生了可遗传的改变。表观遗传受环境因素的调节，是生物体为了适应环境而发生的分子改变。表观遗传可以遗传给子代。在 Y-DNA 上，同样有表观遗传相关位点，受环境因素的宏观调控。

环境是生命物质存在的空间，这种空间的动态变化往往会导致细胞核内部基因物质的变化。基因突变就是其中最简单的形式之一。基因的功能状态受到以物理、化学、生物因素为内容的环境因素的调控，只有适应环境，个体的生存才能够延续。在代谢过程中，生物体不断与外界环境进行物质、信息、能量的交换，形成了平衡和谐的相互关系。如果以物理、化学、生物因素为内容的环境发生变化，生物与环境之间已达成的动态平衡就会遭到破坏，从而威胁个体的生存。

无论是常染色体还是性染色体，基因突变是重要的现象。既往观念认为，基因突变是 DNA 复制的过程出现了某种“错误”。基因突变发生于基因组内部，在编码区和非编码区都有突变出现，只是非编码区的基因突变不与蛋白质的表达发生关系。研究表明，基因突变，包括移码突变和点突变，同样是蛋白质介导的生物学事件，是一种基因编辑现象，也是由基因参与调控这一过程。

放射线是诱发基因突变的重要环境因子。放射线作为一种环境因素作用于

生物体，改变了生命物质分子的能量状态，扰乱了正常的生命秩序，导致伤害发生。放射线先作用于细胞质，然后作用于细胞核。但核酸构成的细胞核对放射线的作用比蛋白质构成的细胞质敏感，因而细胞质对细胞核的作用很大程度被放射线对细胞核的直接作用掩盖了。为了探讨环境、细胞质与细胞核的相互作用关系，就需要引进细胞核自体移植的技术方案。

10. DNA 的修饰和甲基化

除了 DNA 中的遗传密码外，细胞中还存在着其他指令，负责在不改变基因序列下决定基因的表达或不表达，此即表观遗传。表观遗传是指 DNA 序列不发生变化，但基因表达却发生了可遗传的改变。这种改变一般是通过 DNA 的修饰作用完成的。

表观遗传标志会一代一代传下去，老一辈的 DNA 修饰会在子代中继续存在，一个基因一旦开启或关闭则意味着在后代中永远开启或关闭。例如，狼和狗有亲缘关系，但狗会摇尾巴而狼不会。狗在与人类接触的过程中，进化出了摇尾巴的功能，这是与人类交流的一种有效方式。这个功能首先是一种表观遗传，然后被固化，稳定遗传下去，形成了狗与狼一个标志性的行为区别。

人是有知识、有道德的文化群体，与动物有异。进化论表明，人是由猿进化而来，但人与猿有本质的不同，不单是外观形体和基因物质的不同，还有精神和文化气质的差异。这是有遗传学基础的社会现象，是由表观遗传支配的。

对于人而言，一个人的生活方式或受到的压力有可能影响到儿女和孙辈，有长远的效力和记忆，由此形成了一个家族的传统和家风，所谓“老子英雄儿好汉”。基因没有好坏之分，但表观遗传有好坏之分。一个好的传统，好的家风，可以代代相传。这并不是说这个家族有了一个好基因，但一定是有一个好的表观遗传在起作用。俗语说“龙生龙，凤生凤，老鼠的儿子会打洞”，这种差异，除了基因的差异，DNA 修饰的表观遗传也是一个不可或缺的重要因素。

DNA 的修饰并非与生俱来，而是后天形成。随着年龄增长，为应对不同的生活方式和环境变化，DNA 修饰会发生很大的变化。同卵双胞胎的 DNA 序列完全一样，但在行为和生活细节等方面会表现出差异，这在很大程度上就是 DNA 修饰起了作用。

DNA 甲基化在基因表达和调控中起重要作用，是一种重要的表观遗传修饰。一个基因发生了甲基化，通常这个基因的表达会被抑制。例如，冷藏过的番茄，其芳香物质显著减少，变得不好吃。因为低温诱导了 DNA 甲基化的瞬时增

加，芳香物质的产生减少了，所以番茄味道变差了。细胞内部，DNA 甲基化还与衰老及个体的年龄推断有关，DNA 修饰作为衰老指标直接预测个体年龄和寿命，已经显示了很好的前景。

基因组 DNA 中的 5-甲基胞嘧啶（5-^{m}C），是一种稳定存在的表观遗传修饰，这种非基因序列的改变会使基因表达水平改变。启动子区域的甲基化与去甲基化调控着基因的沉默与表达。DNA 甲基化犹如一把锁，当甲基化出现，基因表达被锁死，基因处于沉默状态；当去甲基化发生，锁被打开，启动子开始工作，基因开始表达，产生各种蛋白质的性状。

由于外界环境的作用，DNA 甲基化后基因表达被关闭，这种抑制作用可以传递给子代。相比 DNA，表观遗传标志更容易发生变化，然后影响基因的活性。DNA 的甲基化在群体中广泛存在，而且具有个体特异性，即便是同卵双生子，DNA 甲基化的水平也是不同的，因此，可以使用 DNA 甲基化修饰进行同卵双胞胎的个体识别。

11. 获得性遗传与表观遗传标记

《晏子春秋·杂下之十》有“橘生淮南则为橘，生于淮北则为枳，叶徒相似，其实味不同。所以然者何？水土异也”之说，形成了“南橘北枳”的成语。意思是说环境变了，事物的性质也变了。这里面涉及一个遗传学中十分重要的观念，即获得性遗传。

获得性遗传是后天发生的遗传，指生物在个体生活过程中，受外界环境的影响，产生带有适应意义和一定方向的性状变化，并能够遗传给后代的现象。在不同的自然环境中，生活着不同形态功能的生物，这是人们可以看到的最简单的事实。

获得性遗传是 20 世纪最具争议的经典遗传学话题，提出和支持这个学说的代表人物有拉马克、米丘林、李森科等。摩尔根的基因理论和弗朗西斯·克里克中心法则的提出，一度否定了获得性遗传的存在，将之列在了“伪科学”之列，但基因修饰的发现和表观遗传在理论上的完善，使获得性遗传的经典理论卷土重来。越来越多的证据证明，获得性是可以遗传的，尤其在生物的群体进化中发挥着重要的作用。获得性遗传的关键是看是否有表观遗传标记出现，也就是 DNA 的修饰达到了怎样的水平。之所以称为“表观遗传修饰”，是因为这种表面的修饰可以稳定遗传给子代。表观遗传使生物体的性状遗传与环境变化之间联系得更加紧密，获得性遗传学说也因此具有了坚固的分子生物学的支撑。

获得性遗传在个体识别中也有重要作用。同卵双胞胎有相同的基因，有相同或相近的外观，但是因为他们所处的环境和经历不同，DNA 的修饰水平是不一样的，因此可以通过科学的方法检测和鉴别。在 Y 染色体上，不同个体 DNA 修饰的水平不同，上面有很多的甲基化位点，这也是家系识别和个体识别的重要参照物。

12. 细胞核自体移植方案

细胞核移植是研究核质关系的有力工具，也是体细胞动物克隆的核心技术。方法是将静息期的哺乳动物体细胞核移植入去核卵母细胞中，经过电融合后，开始启动发育，并在培养基中形成胚胎。这是体细胞核移植的一般过程，需要体细胞和生殖细胞的参与。

细胞核自体移植技术是一项细胞工程技术，从核移植发展而来。与核移植技术相比，细胞核自体移植操作上的最大特点是可只使用一个细胞或一种细胞，而此前所有的核移植都是在两种细胞间进行的。该技术的核心是利用环境与细胞质作用，通过改变细胞质以调控细胞核内基因的状态，从而获得适应新环境的细胞株、生物品系，是研究环境因子、细胞质、细胞核相互作用的基础工具（见图 1－8）。

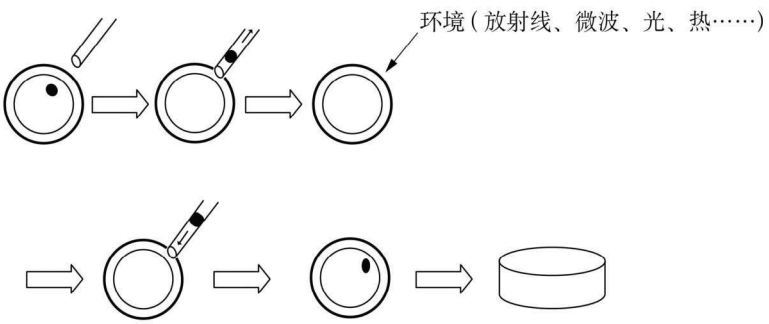


图 1－8　细胞核自体移植方案

环境因子与核质关系的建立和协调有重要作用。环境因子作用于去核细胞，省去了环境因子对细胞核的直接破坏作用，在高能量下，保全了细胞核的结构和功能。此时环境因子对细胞质内蛋白质的直接作用成为环境损伤的主要对象，此后将细胞核重新移回，蛋白质的损伤就会对细胞核内的基因状态造成影响，从而构成蛋白质对基因表达的调控以适应环境因子的作用。这种基因调节

可以稳定遗传给子代，是生物适应环境变化在分子水平的反映。

细胞核自体移植通过改变细胞质的功能状态以对细胞核内的遗传信息进行调控，借助显微操作系统，产生环境作用因素，可以对去核细胞进行处理，或对移出的细胞质进行处理，使细胞质具有与先前不同的功能状态，并对细胞核的表达产生影响。高能环境因子的作用，如绿色激光、X 线、高能电流等，会造成细胞内线粒体的不可逆损伤，引起细胞的凋亡。所以在移植的过程中要注意适度补充新鲜的线粒体，以保证细胞的存活率。

DNA 双键断裂是放射引起 DNA 损伤的主要机制之一，它的修复途径包括同源重组和非同源末端连接两种方式。Y 染色体有大量的移码突变及甲基化修饰位点，这些现象与环境因子有密切关系。因而，细胞核自体移植也是研究 Y-DNA 遗传的有力工具。

13. 适应的分子机制

适应是一种生命现象，即生命体对环境有适应性。生物不断适应环境的过程，也就是进化的过程。整体而言，在环境因素的作用下，基因决定蛋白质的结构和功能，蛋白质决定基因的表达和调控。基因表达和基因调控，是表观遗传发生的重要环节。基因与蛋白协调作用形成的系统构成了生命体适应环境的动力机制（见图 1-9）。

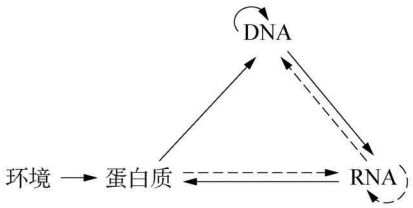


图 1-9　适应的分子机制

适应的分子机制强调生命重要分子间的整体关系。适应的分子机制，揭示了环境、蛋白质、DNA、RNA 间的整体关系，可以对生命系统的起源及其动态变化作出解释。

表观遗传受到环境因素的调控，此过程中 DNA 序列不发生变化，但基因表达发生了可遗传的改变，这是生命体适应环境变化的一种机制。DNA 甲基化在常染色体是很重要的表观遗传标志，在 Y 染色体上，也有大量的甲基化位点。