

第一章

揭开基因的面纱

一、DNA 小史回顾

俗话说“龙生龙，凤生凤，老鼠生儿会打洞”。意思是说儿女总是很像父母。不过，亲代和子代并不完全相同。现在我们已经了解了其中的机制。但在 19 世纪之前，这可是一个天大的谜题。1865 年奥地利的修道士孟德尔从豌豆的交配的实验中发现了遗传具有法则。他认为某种形式的遗传物“因子”会从亲代传给子代，而这种因子相当于现在所说的“基因”。这一法则就是遗传学界著名的“孟德尔定律”。

令人遗憾的是，他的这一发现直到 1900 年才由荷兰的弗里斯、德国的科伦斯和奥地利的切尔马克分别重新发现遗传信息载体。1869 年，瑞士的米夏在医院收集绷带，用上面的脓作为材料。从脓细胞（淋巴细胞）核中发现了含有大量磷酸的前所未见的新物质——“核酸”，而这正是遗传信息的载体，不过当时还不很清楚细胞核中含有染色体。1902 年，美国的萨顿通过观察减数分裂内的染色体，推测基因位

于染色体上。美国的摩根从 1910 年之后开始遗传学的研究，他研究的果蝇的突变，发现基因位于染色体上的特定位置，并制作出了画有果蝇基因位置的染色体图。

英国吉利非斯在 1928 年利用肺炎双球菌进行的实验，否定了当时大多数人认为遗传信息的载体是蛋白质的看法。美国的艾布里于 1944 年提出遗传基因的载体是 DNA（脱氧核糖核酸）的看法。他的观点并未立刻得到认同，直到 1952 年，美国的赫尔西和蔡斯利用噬菌体做研究，才证明了他的观点。噬菌体是侵染细菌的病毒，赫尔西等人的研究采用的是侵染大肠菌的噬菌体。经过科学家研究发现，核酸是由称为核百酸的单位所组成，核苷酸则是由碱基、糖、磷酸等零件所构成，并且知道了碱基有 4 种。美国的沙尔加夫对核酸进行了分析，于 1949 年发现 DNA 的腺嘌呤（A）、胸腺嘧啶（T）、鸟嘌呤（G）、胞嘧啶（C）等四种碱基并非各占 1/4，而是依碱基的种类不同有所差异。此外还发现碱基之中，A 和 T、G 和 C 的含量比例相等。

DNA 的发现更是吸引了众多生物学家投入研究 DNA 是如何遗传的，其分子结构是破译遗传密码的关键。列文结构模型提出之后的 40 年间，生物学家们提出了五花八门的结构模型，但都不成功。直到 1953 年才由沃森和可里克给它们画上了一个圆满的句号。

1950 年夏天，对 DNA 一无所知的美国人沃森博士，在丹麦皇家学会听完劳伦斯布拉格关于 DNA 的演讲后，决定研究 DNA 的三维模型结构。沃森进入英国剑桥大学卡文迪许实验室后，认识了英国的可里克，他们很快发现彼此都对 DNA 的分子结构极感兴趣，便决定合作研究。他们提出生

命分子的三维结构是由线性密码中所蕴含的信息所决定。然而他们的研究招徕了实验室方面的非议。DNA 的 X 射线衍射图提供者威尔金斯对他们的研究也不热心。几起几落的遭遇使可里克心灰意冷。而沃森没有动摇，他坚信 DNA 是所有分子中最重要的王牌，是万木之本，是打开生命之门的钥匙。他和可里克一起，采用物理、化学的科学原理与方法，来揭示 DNA 结构的奥妙。最后，他们提出的 DNA 双螺旋模型认为，必须由两股核苷酸碱基的任意排列顺序，来决定高度有序的 DNA 三维结构。这是一个成功的模型，它由两条右旋但反向的链绕同一个轴盘旋而成，活像一个螺旋形的梯子，生命的遗传密码就刻在梯子的横档上。这个模型就是我们今天在挂图上和生物实验室看到的那个样子。

1962 年的诺贝尔医学生物学奖授予了沃森和可里克建立的 DNA 模型，引发了一门被称为“分子生物学”的新科学的诞生；它为破译生物的遗传密码提供了依据，导致遗传工程学的出现。用人工的方法将生物体内的 DNA 分离出来，重新组合搭配，再放回生物体内，创造新的品种，成为本世纪下半叶最活跃的领域。例如我国科学家实验成功的杂交水稻，抗棉铃虫棉花等等，都是 DNA 重组的产物；目前国际上正热门的克隆技术，也是 DNA 的绝妙之作。1954 年 4 月，DNA 双螺旋提出后，整个生物学界呈现出少有的五彩缤纷的景象，被认为是 20 世纪以来生物科学中最伟大的成果，是生物学史上一个新纪元，为生物科学、农业科学、医学的发展开辟了新天地。把 DNA 结构模型称为打开生命之门的钥匙，是一点也不夸张的。

进入 70 年代之后，利用限制酶切割和组合 DNA 的

DNA 重组技术问世，这项技术的出现使生物学家能够增殖基因或将基因另行组合。从此，有关 DNA 碱基排列组合和研究便突飞猛进。DNA 重组技术的进步也给医学带来了巨大影响。对于传统方法不易治疗的疾病，可考虑采用从基因层次来治疗的基因疗法。经过 80 年代的研究和实验，美国从 1990 年开始正式采用基因疗法，也有人期待这种疗法能对付癌症和艾滋病等疾病。某种生物具有的 DNA 的组合称为“基因组”，它含有该生物全部的遗传信息，而且也记录着该生物的演化过程等资料。1995 年，流行性感冒菌 DNA 的碱基排列组合全部揭晓后，大肠杆菌、酵母等也被拿来作基因组分析。

二、认识基因的奥秘

生物体有很多奇妙的本领，我们见惯了，也就不觉得奇怪，可细想想：为什么我们种下水稻的种子，长出的就是水稻，而不是小麦或者别的？遗传就是上一代和下一代基本上是一样的，不会从人变成别的什么东西，但遗传是什么因素控制的呢？是什么像接力棒一样一代传向另一代的呢？这些都牵涉到了基因，基因好比是计算机上的一串代码，区别是它不是由 0、1 组成，而是 A、T、G、C 四个密码按照一定的顺序排成，而且不是储存在电子芯片上，而是在染色体上，一条染色体上有成千上万的基因，它们经细胞解码、翻译，如同计算机程序的 0、1 码，被 CPU 一处理，哇，它显示出多种功能啦！

上面讲了基因的一个功能，那就是它控制生物体的性

状，那它又是怎么复制给下一代呢？这就是 A、T、G、C 这四个密码的怪癖了。这四个密码实际上是四种不同的物质，它们之间通过化学键连成一起，而基因中这样的信息链不是一条，而是两条。这条链上是 A，另一条链上对应就是 T，这条链上是 C，另一条对应就是 G，反之亦然。这是因为 A 和 T 之间、C 和 G 之间有特殊的吸附力（氢键），两条链通过这种吸附力缠在一起，构成所谓的 DNA 分子，形成染色体的物质基础。基因复制时，两条链在酶的作用下分开，根据 A 和 T、C 和 G 之间的互补原则，分别合成互补链，形成两条 DNA 分子，在细胞分裂时分别分配到两个子细胞中去，通过这样精确的配对关系，将遗传密码复制给下一代。认识了基因的两个特性，即基因控制性状和基因能稳定遗传后，我们就可以来看看什么是基因克隆了。

三、克隆引起大震撼

克隆这一词由克隆羊多莉的诞生而被媒体炒得沸沸扬扬，但我们这里所讲的基因克隆与克隆羊有一些区别。

克隆是指一个无形繁殖系，当用作动词时，克隆则是指利用不同方法产生无性繁殖系所进行的工作。简言之，克隆作动词用时是指研究或操作过程，作名词用时是指产生的结果。

克隆可根据其研究或操作的对象分为基因克隆、细胞克隆和个体克隆三大类。基因克隆是指在分子（DNA）水平上开展研究工作，以获得大量的相同基因及其表达产物。细胞克隆则是在细胞水平上开展研究工作以获得大量相同的细

胞。个体克隆则是经过一系列的操作产生一个或多个与亲代完全相同的个体，这种克隆所用的生物材料可能是一个细胞，也可能是一个组织。很显然，基因克隆、细胞克隆和个体克隆是在三个不同的层次上所开展的工作。以原有的基因或细胞或生物个体作为模板，复制出多个与原来模板完全相同的基因或细胞或生物个体来。这就有点像大家利用复印机复印资料，或用胶片冲洗照片一样，从原有的资料或底片复制出许多完全一样的资料或照片来。

在长期的生产实践和科学研究中，人们已经掌握了许多克隆技术。“有心栽花花不开，无心插柳柳成荫”的诗句可以说是家喻户晓，诗句中的“插柳”实际上就是一个“克隆”过程，即把柳枝切成小段，插入土中。一段时间后，它就可以长出根来，并逐渐长成一棵小柳树。农民种红薯时，先将一根红薯藤剪成若干小段，然后将每一小段栽插在土里，几天以后，红薯藤就会长出根来，然后逐渐长成一根完整的红薯条。植物中还可以举出许多类似的例子。因此，人们对克隆不应该是陌生的，只是熟视无睹罢了。此外，科学家们利用组织培养技术也能使许多植物的不同组织再生成一棵完整的植株，而且已被广泛地应用于生产实践中，这种组织培养技术实际上也是一种克隆技术。英国罗斯林研究所试验成功的克隆羊“多莉”就是首次利用体细胞克隆成功的，它在生物工程史上揭开了新的一页。自然界有不少植物具有先天的克隆本能，如番薯、马铃薯、玫瑰等插枝繁殖的植物。而动物的克隆技术，则经历了由胚胎细胞到体细胞的发展过程。

早在 20 世纪 50 年代，美国的科学家以两栖动物和鱼类

做研究对象，首创了细胞核移植技术，他们研究了细胞发育分化的潜能问题，细胞质和细胞核的相互作用问题。

科学家选取了三只母羊，先将一只母羊的卵细胞中所有遗传物质吸出，然后将另一只 6 岁母羊的乳腺细胞与之融合，形成一个含有其遗传物质的卵细胞，并促使它分裂发育成胚胎。当这一胚胎生长到一定程度时，将它植入第三只母羊的子宫中，由它孕育并产出克隆羊多莉，多莉酷似提供乳腺细胞的 6 岁母羊。

小羊多莉是世界上第一个利用体细胞克隆成功的动物。克隆多莉的成功，从理论上说明了高度分化细胞，经过一定手段处理之后，也可回复到受精卵时期的合子功能；说明了在发育过程中细胞质对异源的细胞核的发育有调控作用。它对生物遗传疾病的治疗、优良品种的培育和扩群等提供了重要途径，对物种的优化、濒危动物的种质保存，对转基因动物的扩群均有一定作用。

自克隆小羊多莉成功以后，世界各国引起了强烈的反响，有的看做福音，有的视为祸水。我们以为对新技术应该采取支持态度，生物克隆获得突破，最大的好处是培养大量品质优良的家畜，丰富人们的物质生活，使畜牧业的成本降低、效率提高，还可以提供某些药物原料以提高人的免疫功能等等。在小多莉以前，罗斯林研究所曾培育出一只奶中含治疗血友病药物原料的转基因羊，一家公司以 50 万英镑的高价买去。如果利用体细胞大批“复制”这只羊，就可挽救更多患者的生命。另外，利用克隆技术可以大量复制珍稀动物，挽救濒危物种，调节大自然的生态平衡，为人类造福，何来忧患呢？当然，克隆技术也可能带来负面影响，一些克

隆动物在遗传上是全等的，一种特定病毒或其他疾病的感染，将会带来灾难。如果无计划地克隆动物，将会扰乱物种的进化规律，干扰性别比例，这种对生物界的控制会带来许多意想不到的危害。但只要采取相应的研究对策，制定科学的克隆计划，这种负面效应就可以最大程度地避免。

围绕多莉的最明显的令人不安的问题是，她是否是人克隆的先兆。从一开始，威尔穆特就一直强调，没有合理的理由为克隆人作辩护。他说，此外，已知目前缺乏有关克隆如何奏效的知识，所以这类试验太危险。“过去，我们曾培育出有明显畸形的羔羊。”至于克隆人，这是一个颇有争议的研究课题。当代生物史证明，克隆技术只能复制出外貌特征相同的生物，不能够克隆出被复制者原有的才能。人的思想才能受后天的制约。所以，即使有人能够克隆出酷似历史上的伟大领袖、伟大科学家那样的人物，却缺乏伟大领袖、伟大科学家那样的思想、气质、才能。试问，这样的克隆具有什么意义？至于有人主张克隆人以取得人体器官，用于医学上人体器官的移植，这也是不可行的。因为克隆出来的人首先是一个人，他享有人权。如果克隆人不肯捐赠器官，你发明者也不能侵犯人权。至于克隆无头的人，那也是不现实的，因为克隆人要生存，首先要吃饭，要思维，没有头颅是不可能的，我们总不能培植一个无头的植物人吧？而且，最重要的是克隆人不符合世情国情，当今世界人口急剧膨胀，不少国家已实行计划生育，控制人口增长，在这种情况下怎能斥巨资做违背社会发展规律的事呢？正如德国研究技术部长吕特格斯所说：“复制人类将不被允许，也一定不会发生。”

20 世纪末，克隆技术在英国又有了新的进展，他们把这一技术应用于人类造血事业。英国的 PPL 公司是克隆技术的经济后台，它的主管罗思·詹姆斯博士说：“从研究多莉中我们知道，我们可以用一个细胞制造出一只转基因动物。我们现在正利用这一技术生产人类血液中最重要的组成部分，也就是血浆。”他们与罗斯林研究所合作研究一种带有 人类基因的牛和羊。他们先把动物体内的血浆取出，再取代人类的血浆，这种改变了基因的牛和羊体内就含有 人类血浆的重要成分，通过对这些动物的饲养、再克隆或繁殖，就可以得到稳定可靠而且相对便宜的血资源。

四、基因工程面面观

基因工程就是人们对生物基因进行改造，利用生物技术生产人们想要的特殊产品。随着 DNA 的内部结构和遗传机制的秘密一点一点呈现在人们眼前，特别是当人们了解到遗传密码是由信使 RNA 转录表达的以后，生物学家不再仅仅满足于探索、揭示生物遗传的秘密，而是开始跃跃欲试，设想在分子的水平上去干预生物的遗传特性。

如果将一种生物的 DNA 中的某个遗传密码片段连接到另外一种生物的 DNA 链上去，将 DNA 重新组织一下，不就可以按照人类的愿望，设计出新的遗传物质并创造出新的生物类型吗？这与过去培育生物繁殖后代的传统做法完全不同，它很像技术科学的工程设计，即按照人类的需要把这种生物的这个“基因”与那种生物的那个“基因”重新“施工”，“组装”成新的基因组合，创造出新的生物。这种完全

按照人的意愿，由重新组装基因到新生物产生的生物科学技术，就被称为“基因工程，或者称之为“遗传工程”。

基因工程的核心技术是 DNA 的重组技术，也就是利用供体生物的遗传物质，或人工合成的基因，经过体外或离体的限制酶切割后与适当的载体连接起来形成重组 DNA 分子，然后在将重组 DNA 分子导入到受体细胞或受体生物构建转基因生物，该种生物就可以按人类事先设计好的蓝图表现出另外一种生物的某种性状。

基因工程技术几乎涉及到人类的生存所必需的各个行业。比如将一个具有杀虫效果的基因转移至棉花、水稻等农作物种中，这些转基因作物就有了抗虫能力，因此基因工程被应用到农业领域；要是把抗虫基因转移到杨树、松树等树木中，基因工程就被应用到林业领域；要是把生物激素基因转移到动物中去，这就与渔业和畜牧业有关了：如果利用微生物或动物细胞来生产多肽药物，那么基因工程就可以利用到医学领域。

五、实施 **HGP** 计划

1. 人类基因组计划及其对生物医学研究的推动

基因是控制生物遗传性状的基本单位。每个基因都是由 4 种简单的脱氧核糖核苷酸（又称碱基）按一定的顺序排列所组成的。人类发育、生殖、生长、疾病、衰老、死亡及精神、行为等都不同程度地与基因有关。决定一个人生、老、病、死等一切生命现象的遗传信息总和称为人类基因组。人

类基因组所含的全部遗传信息由 DNA（即脱氧核糖核苷酸链）携带，分布于 1~22 号染色体、X 染色体、Y 染色体以及细胞质中的线粒体 DNA 上。

目前，人类尚不能完全破译和解读贮存于基因组的信息或遗传语言，基因组 DNA 碱基顺序的排列仍是一部“天书”。因此，如何破译和解读这部“天书”已成为各国科学界极为关注的前沿科学领域。

“人类基因组计划”（HGP）最初是由美国科学界于 80 年代中提出，并于 1990 年开始实施的。其最终目标是阐明记录人类基因组全部遗传信息的 3×10^9 碱基对的 DNA 序列，识别其中的约 10 万个基因，建立储存这些信息的数据库，开发数据分析工具，并研究由于 HGP 的实施所带来的伦理、法律和社会问题。全部研究计划在 15 年内即于 2005 年完成。

基因组计划的实施将对整个生物医学研究产生重要影响。这一计划的最终完成将阐明人类全部基因与基因调控元件的结构和位置，了解染色体结构对基因活性的控制与调节，以及包括遗传性疾病、恶性肿瘤、心血管病、内分泌病、老年性疾病、放射性疾病甚至感染性疾病等的分子改变及其发病机理，为疾病的预后以及分子水平上的诊断、风险预测、预防和治疗提供依据。通过对基因功能的了解，促进新药开发，形成以基因及其表达产物为基础的生物技术产业。对染色体和个体之间的 DNA 序列多样性的研究将促进基因诊断、个体识别、亲子鉴定、药物过敏、耐受或抵抗、器官移植的组织配型、发育进化等诸多医疗、司法和人类学研究。

2. 人类基因组研究对其他科学领域和社会、经济发展的影响

“人类基因组计划”自 1990 年启动以来，在世界各国科学家的共同努力下取得了迅速的发展。其研究成果不仅带动了生物学和医学研究的不断创新，深化了对人类遗传语言、基因结构与功能的关系、生命的起源与进化、细胞发育、生长、分化、及疾病发生、发展的分子机理等问题的理解。同时也使各种先进的工业化技术（自动化技术、信息技术、大规模集成技术等）运用到生物医学领域。³该计划的实施促进了生命科学与信息科学、材料科学和高新技术产业的结合，推动了基础性研究和不同科学的交叉融汇，刺激相关学科技术领域的发展，并带动起一批新兴的高技术产业。基因组研究还影响到农业、畜牧业（转基因动物、植物）；微生物学（微生物药物、保健品；抗病原微生物的药物；土壤改良等）；计算机科学、自动化技术、纳米科学、微电子技术、激光、信息科学，以及社会、伦理、道德、法律和经济问题（医疗、人寿保险等）等诸多基础与应用科学的研究以及人文科学研究。基因组研究的成果既可以直接指导和转化为实际应用，并可能改变人类社会生产、生活和环境的面貌，具有不可估量的社会和经济效益。因此称 HGP 为人类科学史上与曼哈顿原子弹计划和阿波罗登月计划齐名的三大科学工程之一，其科学价值和人类社会发展的影响将远远超过前两个计划。

20 世纪 90 年代以来，国际人类基因组研究持续升温，其主要原因在于人类所有疾病和健康状况都与基因直接或间

接相关。每种疾病都有其相应的致病基因或易感基因存在。许多有战略发展眼光的企业家已意识到基因组研究必将带动和促进治疗性蛋白、医疗诊断、基因治疗、组合化学分子治疗剂、新药筛选以及生物信息、数据库管理、微电子技术、自动化技术设备等领域的开发生产研究，带来巨大的商业利益。如某些基因编码的蛋白有可能作为药物矫正遗传缺陷或调节代谢功能（如凝血因子、干扰素、细胞因子等）；通过检测疾病或重要生物学功能基因的 DNA 或蛋白质诊断疾病或监控代谢；将基因送入细胞中，矫正遗传缺陷、破坏异常细胞或调节细胞功能，进行基因治疗；以编码细胞受体或其他特异蛋白的基因为治疗应用的目标分子；开发小分子治疗剂；从组合化学库中筛选开发治疗性药物；DNA 序列自动化分析技术与数据分析能力集成的商业开发；生物信息学、基因组信息数据库和分析软件的商业开发；纳米技术、生物芯片、微电子等技术的商业开发等。众多国际大型制药企业将巨额资金投入疾病相关基因研究和“基因组工业”中。并导致了愈演愈烈的基因专利之争和由此诱发的疾病家系、特殊人群以及染色体异常材料等遗传资源之争。美国和欧洲一些研究单位或公司已开始通过不同途径和方式进入或试图进入发展中国家的人类基因组研究领域。显然，新基因的克隆分离、功能鉴定和开发应用研究已经与国家、民族和人民的利益联系在一起了。随着公司、药厂的投入，功能明确的基因序列，特别是与严重危害人民健康的遗传易感性疾病（如恶性肿瘤、心血管疾病、神经、精神性疾病及免疫性疾病）的基因、一些编码重要的生长因子、细胞因子及其受体的基因，或与重要生物学过程如发育、分化、衰老相关的基因专

利化趋向已下可逆转。而一旦某个功能基因序列被专利化，其下游产品的开发也就必须引进此种专利。某些基因的专利转让费已达上千万美元。从这个意义上讲，基因序列本身就是产品，而且是一种不能再生的资源。可以说，谁在基因组研究中处于领先地位，谁就占领了 21 世纪生命科学的制高点；谁拥有了基因，谁就拥有了 21 世纪的生物技术产业。

3. 各国政府对基因组研究的态度

美国是人类基因组研究计划的创始国之一，美国国会批准对 HGP 投入 30 亿美元。英国政府近年来每年在 DNA 序列测定和欧洲生物信息中心的建设等方面的投入达到 1000 万英镑以上。法国的 Genethon 实验室和 CEPH 中心，仅民间集资即达 6500 万法郎，1997 年又新建了国家基因研究中心。欧共体每年投入 500 万美元以上的基因组研究专项基金。日本 1996 年生物技术研究预算达 2050 亿日元，1997 年研究开发预算要求额为 2508 亿日元，主要用于克隆技术、核苷酸序列测定、基因治疗、农业畜牧业研究、环保生物技术等。瑞典政府 1997 年 ~ 2000 年每年对基因组研究的投入为 1000 万到 1500 万法郎。新加坡、印度、巴西等政府纷纷建立国家级生物技术研究开发机构，保存、保护遗传资源，积极实施开发研究。

中国 HGP 在 90 年代初的发展尚处于萌芽时期，科技部和国家自然科学基金委员会在生物技术领域内组织了一批与人类基因组研究相关的研究课题。1993 年启动的《中华民族基因组中若干位点基因结构的研究》重大项目，标志着我国人类基因组研究正式启动。国家“863”高技术计划生物

领域 1996 年起设专题项目支持人类重大疾病基因的研究，该项目 1997 年升级为重大项目。目前，中国人类基因组研究在我国多民族基因组的保存和分析、重要遗传家系资源的收集、建立和发展基因组研究的新技术以及疾病相关基因研究等方面已取得了多项重要进展。1999 年科技部和中国科学院联合支持了“国际人类基因组计划 1% 基因组测序”项目，郑重向世界宣告中国科学界加入国际人类基因组 DNA 序列测定计划，承担其中 1% 的测序任务。

4. 国际人类基因组研究的进展和发展趋势

人类基因组计划包括构建遗传连锁图、物理图、测序和基因识别 4 项工作。前两项任务已提前完成。1995 年完成了由 225 个酵母人工染色体（YCA）连续克隆系组成的第一代人类基因组物理图谱。目前已建立了 15086 个序列标志位点，平均每 199kb（千碱基对）就有 1 个序列标签位点。其中 X 染色体和 7 号染色体达到 75 ~ 80 kb 一个序列标签位点。1996 年完成了由 5264 个微卫星 DNA 标记组成的人类第二代遗传连锁图，覆盖 4000cM（分摩，遗传距离单位），分辨率 < 0.7cM。第三代遗传标志 &0; &0; 单核苷酸多态性（SNP）标记研究的迅速发展，不仅可以进一步提高遗传图与物理图的分辨率，还有望用于疾病候选基因的筛选鉴定与药物基因组学研究。1998 年 10 月发表的第二张人类基因图将人类 30181 个基因定位在染色体的不同区带上。截至 2000 年 3 月 17 日，在公共数据库中可以查到的表达序列标签（EST）已超过 378 万条，其中人类基因的表达序列标签为 1,810,391 条。在基因组序列测定方面，根据美国国家生

物信息中心 (NCBI) 的基因数据库 (GenBank) 2000 年 2 月公布的数据 , 已测定的基因组序列超过 5,691,000 条 , 碱基总数达到 5,800,050,000。线虫以及酵母等 10 余种生物基因组的全序列已测定完成。

1999 年 10 月 2 日出版的《自然》杂志发表了人类 22 号染色体 DNA 序列测定结果。该工作由英、美、日等国科学家合作 , 测定了人类 22 号染色体 3350 万碱基对的序列 , 占 22 号染色体长臂 DNA 总序列的 97%。这些序列分布在 12 个 DNA 片段上 , 尚有 11 个空隙或区域未被覆盖。 12 个 DNA 片段中最长的一个达到 2300 万碱基 , 这也是人类基因组测序研究迄今获得的最长的 DNA 片段。在 22 号染色体上发现了至少 545 个基因和 134 个假基因 (在进化过程中曾经发挥过功能 , 但现在已不再发挥功能的基因)。此外尚有 200 到 300 个尚待确认的基因。基因长度从 1000 碱基到 583,000 碱基 , 平均长度 19 万碱基。若以这一基因密度估算 , 人类基因的总数将不超过 8 万个。人类 22 号染色体上的基因与免疫功能、先天性心脏病、精神分裂症、智力低下、出生缺陷及若干肿瘤如白血病等有关。 22 号染色体 DNA 全序列的测定无疑将对相应疾病的研究起到极大的促进作用。22 号染色体 DNA 序列的测定还使科学家们第一次从整条染色体的水平上去观察、分析和理解 DNA 碱基排列的科学意义。或者说从 DNA 碱基排列的水平上去理解染色体的结构 , 包括基因或基因家族的染色体定位、基因表达如何受染色体的调节和控制、染色体缺失或重排如何导致疾病以及染色体如何复制与遗传等。 22 号染色体 DNA 序列的测定使科学家们在着丝粒附近发现了以前未曾想到的复杂的重

复顺序结构。这一信息也许有助于解释为什么 22 号染色体在这一区域的重组导致 Di George 综合症。

截至 2000 年 2 月 24 日，人类基因组测序计划已完成了 64.1% 的任务。其中 16.7% (535,133,000 碱基) 为完成图 (100% 完成该区域的 DNA 序列测定)，47.4% (1,524,139,000 碱基) 为工作框架图 (初步测定该区域 90% 的 DNA 序列)。以这样的规模和速度发展，到 2000 年夏天以前，完全可以完成预定的工作框架图测定任务，于 2003 年提前完成人类基因组全序列的测定工作。

未来 3 至 5 年国际人类基因组计划的研究内容将集中在：

(1) 克隆和测定人类及模式生物 (特别是小鼠) 的全长 cDNA 序列；

(2) 开发和完善 DNA 序列整合、分析，以及表达基因识别的软件；

(3) 开展第三代遗传标记，单碱基多态性 (SNP 遗传标记研究，用于人类基因组多样性、药物基因组学研究和识别、定位疾病相关基因等) 研究；

(4) 继续研究由于基因组计划带来的社会、伦理和法律问题等内容上。

在大规模进行基因组 DNA 和 cDNA 序对测定的同时，一个以破译、解读和开发利用基因组功能信息、特别是与人类重要疾病和重要生物学功能相关的遗传信息为主要研究内容的功能基因组学研究将伴随结构基因组研究共同发展。功能基因组学主要研究基因的识别及其功能信息的提取、鉴定和开发利用，涉及生物信息学、计算机生物学和生物医学研