

# 人类基因组计划的进展及其影响<sup>\*</sup>

Progress of The Human Genome Project and Its Impact on Economy

吴 ③

(中国医学科学院肿瘤研究所, 中国科学院院士 北京 100021)

## 一、缘 起

人类基因组计划(Human Genome Project, 简称HGP)最初由诺贝尔奖获得者 Renato Dulbecco 于1986 年建议, 主要目标是测出人的遗传物质共约  $3 \times 10^9$  个核苷酸的全序列, 使人类对自身的认识达到一个新的高度。建议提出后在科学界引发了一场辩论。许多科学家认为这是一个大胆的、富有想象力的建议, 完成这一计划就像基督教徒找到了圣杯(Holy Grail)(意指梦寐以求的东西或目标); 反对者则认为实行这个计划按当时的技术水平估计(主要指测序能力)生物学界的大量人力和基金将被这个计划吸去而影响整个生命科学的的研究水平; 还有一种意见是认为 30 亿个碱基序列即使全部弄清了, 记载于电话簿那样的书本, 也会堆满一间屋子, 打开一看, 乃是密密麻麻只有 4 个字母排列组合的天书, 何况其中只有大约百分之几是编码蛋白质或 RNA 的序列, 百分之九十几可能是“垃圾”, 因而主张把大约 5~10 万个基因的编码序列(即 cDNA 序列)弄清楚就行了, 不必去测全序列。经过大约 3 年的讨论后, 美国政府决定于 1990 年 10 月正式启动这项将耗资 30 亿美元的 15 年计划, 预期到 2005 年弄清人基因组大约 30 亿个碱基的全序列。最初(1990)计划的目标是: 描绘全部人类遗传物质(即基因组, genome)的特征, 即改进现有的遗传图和构建所有染色体的物理图, 并最终确定全部碱基序列, 发现总数超过 50 000 的人类基因, 使之适用于深入的生物学研究。这个计划还包括对一系列模式生物体基因组的全测序, 最初提出的有大肠杆菌、酵母、拟南芥、线虫、果蝇和小鼠。对这些处于生物演化不同阶段的生物体的研究是认识人基因组结构和功能绝对不可缺少的。1993 年对原先的计划进行了修订, 提出了 1993~1998 年五年计划各项指标。随着 HGP 的实施, 又有科学家提出一些其他的

模式生物体, 如河豚鱼、斑马鱼等。此外, 一些具有重要生产价值的农作物, 如水稻基因组等的研究也被提上议事日程。1998 年 10 月又提出了 1998~2003 年的五年计划。美国政府资助的 HGP 由国家卫生研究院(NIH)和能源部承担。NIH 为此专门成立了一个新的机构, 现称人基因组研究所(NHGR)。

## 二、HGP 的实施和发展

美国关于 HGP 的辩论和启动引起了全世界科学家的兴趣, 并带动了欧共体、日本、加拿大、苏联、巴西、印度等国提出各自的计划。通过科学家和政府的共同努力, HGP 成为国际合作的大项目。吸引了世界各国的许多科学家, 不断取得令人惊叹的成就。有鉴于此, 各国政府在这方面的投入不断加强。如日本, 即使在目前的经济危机中仍计划把 1998 年的基因组研究经费增加 1 倍; 韩国要在生物工程 2000 计划中投入数十亿美元。从基因组研究获得的 DNA 和蛋白质序列数据与日俱增, 名副其实地体现了信息的“爆炸”。美国最初的 HGP 已包含了生物信息学(Bioinformatics)的内容, 近年来用在这个方面的资助不断增长。日本 1998 年用于基因组生物信息学的经费为 1997 年的 533%! 生物信息学在寻找和定位新基因、探索基因和蛋白质功能、拼接已测序的 DNA 片段、修饰和设计具有特定功能的新分子等方面的作用与日俱增, 已成为基因组研究中绝对不可缺少的新学科。HGP 吸引了数学、物理、化学、材料、计算机等各方面的人才, 综合利用纳米工程、组合化学、波谱学、光刻学、高效筛选、转基因技术、机器人、DNA 芯片和生物信息学等的一切新进展。对人基因组的研究已经不再满足于原先计划的目标。1997 年底有人提出 HGP 的最终目标应该是提供生物学周期表(Bibby's Periodic Table), 这个周期表的“元素”就是决定人类一切性状的大约 5~10 万个基因。完成 30 亿个碱基的测序只不过为这

<sup>\*</sup> 本文系作者 1998 年 1 月在生物物理所, 6 月在中国科学院第九次院士大会生物学部 and 同月在第 100 次香山科学会议上的报告, 同年 11 月正式校定完的。

个目标打下结构上的基础(即结构基因组学)。对各个基因(包括其编码的蛋白质)和整个基因组的功能研究(功能基因组学)已经开始,比较基因组学和演化基因组学在认识基因组功能方面将发挥愈来愈大的作用。这方面的研究估计将延续到整个 21 世纪。

对基因组计划感兴趣的不仅有科学家和政府,还有许多大药厂、大公司以及他们投资组建的形形色色生物工程公司。其中较早的有史克公司,斥资数千万美元把原在美国国家卫生研究院(NH)工作的科学家 J. Craig Venter 请去组建了一个私营的基因组研究所(The Institute for Genomic Research, 简称 TIGR)。这个研究所在基因组学领域已占有一定的地位,在人基因组和病原微生物基因组的测序方面已作出不少重要贡献。1996 年史克公司更向基因组研究投入 1.25 亿美元巨资,用以获取更多的人基因组序列,并将生物信息学家的编制从 2 名增加到 70 名,使其计划中开发的新药有 25% 以上立足于基因组学。还有一些公司资助人类常见疾病基因的克隆和重要生物体基因组的测序和修饰。据最新的估计,世界各公司投向基因组学研究的经费已达每年数十亿美元,大大超过了政府的投入。

### 三、知识产权之争

这种形势一方面极大地加速了基因组研究的进展,另一方面也加剧了知识产权的争夺:参加政府资助的 HGP 的科学家有义务将新获得的 DNA 和蛋白质数据随时存入大家都能参考的公共数据库。公司资助和私营研究机构的数据库往往不对外公开,凡有潜在经济效益的基因、EST (expressed sequence tags, 表达序列标签)和新近成为争夺热点的 SNPS (single nucleotide polymorphisms, 单核苷酸多态)等都保存于自己的数据库中。近年来这方面的争论日趋激烈。今年 5 月 1 日的《科学》杂志上刊登了两篇不同观点的评论:密执安大学法学院的两位学者著文认为生物医学研究成果的私有化(指专利)既刺激了私人投资,也带来一种风险:由于支离破碎和互相重叠的知识产权迅速增长,使上游的知识向为改善人类健康的下游产品的转化遇到许多麻烦;结果是上游的知识产权越多,有用的产品反而越少;因此,决策者应对上游的专利进行合理的界定,最大限度地减少影响下游产品开发的限制性特许证。在同一期上还刊登了美国专利和商标局负责生物工程检查的官员写的一篇题为:“DNA 的专利”的文章,认为没有必要对现行的专利制度进行修改。杂志编辑部号召对两篇文章的观点展开讨论。

### 四、主要成就,第三次技术革命和对世界经济的影响

HGP 实施以来的将近 8 年中,由于投入的不断研究和研究队伍的扩大,进展是十分惊人的。测序速度由于采用了大规模自动化手段,一个测序中心可以安装数十台至上百台自动测序仪,结合机器人操作,电脑储存和分析,日夜不停地进行。过去一位博士生花数年功夫才能完成的数十万碱基大小的序列,现在在一个测序中心已成为常规工作。许多涉及重要疾病的基因如肥胖基因、支气管哮喘基因等常常成为传媒大肆炒作的新闻,或以数千万美元的高价转让给公司,或在新闻发布会上宣称已经克隆到某某基因而使公司的股票飙升。

更重要的进展在于模式生物体和病原微生物基因组的测序:1996 年完成了酿酒酵母 1 206.8 万个碱基的全序列,从中大致确定了 5 885 个编码蛋白质的基因、140 个 rRNA 基因、40 个 snRNA 和 275 个 tRNA 基因。从而第一次知晓一种最简单的真核生物所需的全部基因的数目和大体上的功能分类。在此前后,病原微生物基因组方面也取得了一系列重要的进展,如流感嗜血杆菌、生殖器支原体、幽门螺杆菌、包柔氏螺旋体、结核杆菌和梅毒螺旋体等影响千百万人健康和生命的病原体的全基因组序

表 全基因组测序举例

|                                  |                                     |                                                   |
|----------------------------------|-------------------------------------|---------------------------------------------------|
| 梅毒螺旋体                            | 1 138 000bp                         | Science 281, 375, 1998                            |
| 结核杆菌                             | 4 410 000bp<br>富含 GC                | Wellcome Trust 资助英,法科学家于 Sanger 中心完成, 1997. 12 宣布 |
| 幽门螺杆菌                            | 1 667 867bp                         | Nature 388: 539, 1997                             |
| 枯草杆菌                             | 4 214 810bp                         | Nature 390: 249, 1997                             |
| 包柔氏螺旋体                           | 901 725bp (染色体)                     | Nature 390: 580, 1997                             |
| Borrelia burgdorferi Lyme 病      | 533 000 (质粒)                        |                                                   |
| 流感嗜血杆菌                           | ~ 1 800 000bp                       | Science 269: 496, 1995                            |
| 生殖器支原体                           | ~ 600 000bp                         | Science 270: 397, 1995                            |
| 酿酒酵母<br>Saccharomyces cerevisiae | 12 068 000bp<br>5 885+ 140+ 40+ 275 | Science 274: 546, 1996                            |
| 人                                | 完成 3%                               |                                                   |

列均在最近3年内完成(见表),为认识这些病原体的致病分子机制,设计诊断、预防和治疗的新方法新途径开辟了无限的可能性,也为拥有专利的药厂、公司提供了滚滚财源的前景。

比如,幽门螺杆菌可能同胃、十二指肠溃疡,甚至同胃癌的发生都有密切关系;结核病长期以来是人类健康的大敌,近年来又趋猖獗,不但危害发展中国家的人民,在发达国家也卷土重来,每年造成300万人死亡,结核杆菌全基因组序列为Wellcome药厂设计准确和快速的诊断方法及有效的治疗药物提供了新途径,1998年10月法国科学家已经找到结核杆菌的致病基因 $ep$ 。梅毒这种可怕的性传播疾病自16世纪危害人类以来,始终未能在体外培养繁殖其病原体,因此至今无法制备疫苗。TIGR和德克萨斯的一组科学家经过长期的努力,终于完成梅毒螺旋体全基因组的测序。而且从全序列的分析中初步了解到这种病原体诱发可怕临床症状的秘密以及为什么不能在体外培养繁殖的原因。在弄清基因组的全序列后,科学家们十分乐观地预言,配制出一种培养基来繁殖这种螺旋体,深入遗传学研究和制备疫苗的前景已经在望。

对人类生活产生更深刻和广泛影响的是对重要农作物基因组的研究。从80年代已经开始的对植物遗传物质的修饰已经和正在对人类社会产生愈益巨大的影响。全世界绝大部分的食物、燃料、纤维、化学原料和部分药物将来都可能取自遗传修饰的植物。如通过遗传修饰使之对某种除草剂(如Roundup)产生抗性的大豆、棉花、油菜、土豆、玉米等的种子已经育成,而且通过了美国FDA严格的审查,开始在大田播种。孟山都公司的抗Roundup大豆在1996年种植了100万公顷,1997年上升到900万公顷,今年将达到2000万公顷。该公司育成的抗棉铃虫棉花种籽已经在我国北方大面积长期推广。今年全世界种植的基因组修饰的各种作物大约会超过2500万公顷。

预防重于治疗的思想逐渐深入人心,保护和提高健康水平、延长寿命已经成为一股势不可挡的潮流。随着对人基因组的深入了解和检测方法的改进(如用于诊断的DNA芯片技术等),医生将逐步根据对象的基因型开出个体化的处方,处方的内容将不限于传统的药物。营养品、保健品的市场将不断扩大。可采用基因工程手段改善食用油成份、蛋白质和碳水化合物含量和氨基酸组成,生产更有益于人体健康,能预防动脉硬化或骨质疏松症等常见病的农作物;把某些病原体的抗原基因或毒素基因转入普通的食物或水果(如土豆、香蕉等)中,以服食这种食品替代传统的疫苗达到预防疾病的目的,在实验中已初露端倪;利用细胞衰老的分子机制研究成果生产保持皮肤光泽、有弹性的新一代化妆品也

将进入市场。这些遗传修饰的食品、化妆品等等都将成为医生处方的内容。我国传统医学中的预防原则(“上医治未病”)和食疗方法将在基因组水平上得到进一步发扬光大和提高。

能源、矿产和环境3方面的企业也正在生物分子水平上汇合。能源公司开始寻找其他形式的能源产品。利用古细菌(Archaea)在无氧条件下产生甲烷的甲基辅酶M还原酶有可能使今后的能源主要来自植物而不再依赖煤、石油和天然气等矿产品。有的公司已经在其附属工厂生产乙醇以代替石油,利用基因组学的成果则可使生产成本大大降低。美国能源部正在研究如何利用一种细菌(*Deinococcus radiodurans*)去清除重金属和放射性废弃物对环境的污染。改变拟南芥的基因组,使之能生产可生物降解的塑料的工作已取得成就。上述的例子仅仅代表一种大趋势的开端。今后10年随着HGP的进展,人们对各种动植物基因组的测序和改造能力将更加强大,各种各样具有新的遗传特性的植物和动物将为我们提高食物的产量和质量,提供有特殊性能的纤维和其他化工原料、廉价能源、易于自然降解避免污染环境的塑料。去年3月有人提出,现在已出现继工业革命、电脑引发的信息革命(computer-based revolution)后的第三次技术革命,即基因组学革命(genomics revolution)。不管这个观点是否会被普遍接受,基因组学的发展对人类社会方方面面(保健、食物、能源、环境乃至观念和行为)产生的巨大影响是无法否定的。

由于基因组学的发展,设计和创造了许多重要的生物分子,这些分子可以应用到制药、农业、食品、化工、化妆品、环境、能源各领域,并能产生惊人的经济效益,自然而然地形成了一个以基因组学为基础的生命科学工业。这个工业将成为世界经济中发展最快的部门。美国专利和商标局(PTO)在1991年收到核苷酸序列的专利申请有4000件,1996年上升到50万件。面对这种态势世界上最大的制药、化工和农业公司和企业正在进行改组、合并和结成新的联盟,以加强自己的竞争实力。如老牌的汽巴药厂(Ciba-Geigy)和Sandoz组成了Novartis这个世界上最大的联合体,以其雄厚的资本和R&D不仅参加到保健业,而且也进入营养和农业领域的竞争;他们的资产超过1000亿美元。最近宣称,将任命伯克莱加州大学的化学家P. Schultz为其新建的Novartis功能基因组研究所所长,该所将招聘100位科研人员,10年中给予2.5亿美元的资助,以新发现基因的功能为基础,开发新药物。葛兰素(Glaxo)和史克的近似联盟成为世界第三大公司,占领了全球超过7.5%的医药市场,新公司1年的研究经费即超过人基因组测序的全部估计费用,它的市场资本总量超过世界上143个国家的年总产

值之和! 孟山都(Monsanto)这个传统的化学公司自1985年至今舍弃了许多原有的主要企业,1997年以来,向生物工程和基因组学投资66亿美元。现已彻底改造成生命科学公司。孟山都的股票自1995年到1998年中上涨了5倍,还要把美国的大制药厂American Home 制药公司合并进去。著名的杜邦公司起初想在保留传统的化学和能源企业的同时扩展到生命科学,但未能成功。1998年4月杜邦公司宣布改组成由生命科学领衔的3个企业单位。一周之内它的股票就上涨了12%。这还不算,到了5月份,又宣布(尽管在先前曾一再否认),将把能源公司Conoco卖掉,把收入用于将杜邦公司改建成一个生命科学公司。它还用26亿美元购进了一项联合制药风险投资的Merck公司的股份,用于研究基因组、生物学、化学和生物工程。Dow Chemical公司用9亿美元购进一项旨在修饰谷物和食品的联合风险投资的Eli Lilly厂的40%股份。今年夏天起Dow化学公司宣布它将成为一家生命科学公司。Hoechst公司卖掉了它的基础化学部门,并通过与另一家公司合作将资本投入到制药业和农业。

由于保健策略的重点从治病为主转向个体化的预防,农业、食品和营养业也同生物工程和制药业融合。Genzyme Transgenics公司正在培育用遗传工程改造的山羊。有一群羊已可生产足够的抗凝血酶III(Antithrombin III),可取代价值1.15亿美元的工厂。用转基因动物生产药物的成本只及用细胞培养的1/10。Novartis正在修饰玉米,使之能预防骨质疏松症。1998年5月孟山都用42亿美元购进了两家公司,同世界上最大的私人公司之一(Cargill)联合投资生产和组装遗传修饰的食品。其他大的食品生产商也在投资和销售遗传修饰的谷物。他们正在创造一个新的工业部门,叫做农作物制造业(agriceuticals)。

## 五、HGP 面临的挑战

尽管在人类疾病基因、微生物、植物基因组测序和功能研究方面取得了重大进展,人基因组全测序这个基本任务却是个硬核桃。人基因组内大量存在的重复序列给拼接工作带来极大困难。至今才完成了3%(大约9000万个碱基)。在今后的7年中是否能完成剩下的97%是个无法回避的问题。今年第一季度参加HGP的许多科学家已经发现,他们对人基因组全测序的估计过于乐观了。从1996年开始进行大规模测序的几个著名中心的成绩都不够理想。今年NHGRI给6个著名的测序中心加强了资助(共6000多万美元),要求在未来的1年中完成11700万个碱基。各中心都正在为能否完成这个指

标而绞尽脑汁。

去年5月9日,前面提到的J. Craig Venter和Perkin-Elmer公司(世界上最大的DNA自动测序仪厂家)突然宣布,他们要组建1家以盈利为目的的私家公司,投资3亿美元在3年内大体上完成人基因组的全测序。从今年4月开始,装备230台新式测序仪(AB PRISM 3700型),每天可测1亿个碱基,第一年即可完成30亿碱基的99%,第2、3年进行拼接和填补重要的空隙;将采用强大的计算机。在5月11日的一次新闻发布会上,Venter甚至建议政府资助的HGP从事模式生物(如小鼠)基因组的测序,把人基因组的测序工作让给他来做。6月5日的《科学》上发表了Venter等关于人基因组测序新策略的文章。8月中,又有一家在美国加州的遗传数据公司(hcyte 药物公司)宣布,他们计划在今后两年内投资2亿美元把人类基因组中编码蛋白质的序列全部测出,还要找出SNPs,把它们全部定位于一张电脑图上,这些数据只对愿向hcyte付高额费用的人开放。面对如此强大的挑战,HGP界的头头们已从最初的震惊中冷静下来,研究对策。从今年10月公布的美国HGP1998~2003五年计划指标中可以看出政府资助的科学家们正在全力以赴,迎接挑战。不管结局如何,政府和私人企业之间在人才、知识产权、专利等方面的竞争将更趋剧烈。人基因组计划在进展速度、技术路线乃至组织形式上,正在经历一场变革。

## 六、伦理、法律和社会问题

遗传学的进展在西方世界一直引起宗教界的不安,认为遗传学家要扮演上帝的角色。70年代重组体DNA技术(即遗传工程)出现后担心其对环境、社会和个人产生不利影响的人已不限于宗教界。美国的HGP一开始就充分注意到这种态势,每年都拨出一定的款项资助伦理、法律和社会问题的研究(ethic, legal, and social issues, 简称ELSI),1998年的拨款为690万美元。比如,照西方的观点,一个人的基因组应属于个人隐私,其中含有什么致病基因或对某种疾病的易感基因,用现代方法很容易查出,查出结果泄漏出去,被检对象就可能在就业、婚姻、保险等方面受到歧视;如果没有相应的法律加以保护,DNA测试就会受到拒绝。至于试管婴儿、人工受精乃至克隆人等引发的伦理、法律和社会问题则更是公众十分关注的问题。如果处理不好,不仅会影响遗传学、基因组学的进展,还会引发种种意想不到的社会问题

(14篇摘自Science和Nature的英文参考文献省略)  
(责任编辑 蔡德诚)