

“人类基因组计划”自启动至“后基因组计划”的转折

杨焕明

(中国医学科学院 中国协和医科大学 教授)

(中国“人类基因组计划”重大项目 秘书长)

今年是“人类基因组计划”(Human Genome Project, HGP)正式提出十周年。对这一可与“曼哈顿原子弹计划”、“阿波罗登月计划”相媲美的创举作一历史性的回顾,纵览十年来这一计划的成果、瞻望这一计划的前景,对于我国抓住机遇,接受挑战是很有必要的。

Dulbecco“标书”

尽管对人类基因组的研究在七十年代便已开始,八十年代初在很多国家已形成一定规模。但一般还是认为全球性 HGP 的蓝图是由 Dulbecco 于 1986 年 3 月 7 日在“Science”上发表的短文中首先提出的。

在这篇题为“癌症研究的转折点——测定人类基因组序列”的短文中, Dulbecco 首先讨论了癌症研究的进展。癌症研究最重要的成果是使我们认识到,癌症与其它疾病的发生都与基因有关。因此,他提出:我们有两种选择,要么经过“零敲碎打”途径(piecemeal approach)各自寻找自己感兴趣的基因;或是测定人类的整个基因组序列。

Dulbecco 详细论述了这一工作的意义。他写道:“这样的工作是任何一个实验室难以承担的,至少应该成为国家级的项目。它的意义可以与征服宇宙的创举相媲美,我们应该以征服宇宙的气魄来进行这一宏伟的计划”;“更加吸引人的是使这一计划成为国际性的课题,因为人类的 DNA 顺序是人类的真谛,这个世界发生的一切,都与之息息相关。

Dulbecco“标书”的背景

八十年代初期,由于分子生物学技术,特别是 DNA 克隆、测序技术的进展,生物学、医学

的研究正酝酿着新的突破。例如:大批 onc 基因与 anti-onc 基因的发现,使 70 年代一度彷徨的肿瘤研究“柳暗花明又一村”;基因克隆技术的突破使遗传工程产业势在必发;基因表达研究技术趋于完善而使“讯号传导”研究初露曙光;人类高级神经活动研究取得的新的成果;大规模双相电泳、核磁共振、X 射线衍射等技术的建立与改进,蛋白质研究方兴未艾……

在世界范围内,就这些科研课题都提出过宏伟的计划:“肿瘤计划”、“遗传工程计划”、“讯号传导计划”、“脑的十年”、“蛋白质计划”、…。为什么只有 HGP 被大家接受为国际性重大计划,而又确实取得了比事先预计更好的成果呢!

自然科学有它本身的规律,有它的发展阶段和内在联系。只要想到所有这些计划的“掐脖子”因素都是需要基因来操作,就不能不钦佩 Dulbecco“标书”的高瞻远瞩。从某种意义上可以认为, HGP 是一个“补课计划”,只有得到人类整个基因组的所有或大部分基因,所有别的计划才有可能变为现实。

对中国来说, HGP 无疑是一把“双刃剑”。既是挑战,又是机遇。我们当时并没有认真接受挑战,我们也就有可能失去这一机遇。一位留学生当时就感叹:“如果中国不在这方面形成研究实力,只能等着扮演材料提供者的角色了。”

HGP 的成果

HGP 的成果主要表现在“四张图”上:

1. 遗传图:学术界长久以来一直认为世代漫长、婚配无序,难以进行遗传分析的人类居然

首先有了一张最详尽的遗传图。RFLP 继之 STR 标记的应用使经典遗传学真正获得了新生。现在我们已经有了 7000 多个信息丰富的遗传标记,分辨率已达 0.7 cM,已能满足单基因性状的定位克隆要求,也为多基因疾病的定位奠定了基础。这一工作比原定的 2000 年完成 2—5 cM 的计划提前了好几年。

2. 物理图:物理图有两方面,一是由一段已知序列为标记的 STS 图,作为基因组的“物理路标”,现分辨率已达 200 kb;二是以 DNA 克隆片段相互连接的“重叠群”,作为基因组研究的操作材料,现已完成 90% 以上。

3. 序列图:序列图即分子水平的物理图。由于测序技术的革命性改进,也可望在 5—6 年内完成。

4. 转录图:转录图是基因图的雏形,现已至少有 25 万个 cDNA 序列,并正在以每天 1000 多个的速率增长。

100 多年前, Henry Gray 的第一张人体解剖图奠定了近代医学的基础, HGP 所提供的这四张图组成了人类的第二张“解剖图”,也就奠定了现代生物学、医学的基础。并将成为生物学、医学研究中用之不竭的源泉。

我们要颂扬 HGP 精神,在人类自然科学史上,这是第一次全球科学家精诚合作,分担喜忧,共享结果。HGP 也是人类历史上第一次实现了以全球联网信息系统来传递实验材料。如最重要的试剂—DNA 探针,可以在自己的实验室里根据信息自行合成,再也不需索取、邮寄。

后 HGP 计划

在目标具体、任务明确、“定时定量”的 HGP 完成之后,还有什么事情要做呢? 这便是在 HGP 之后我们所要考虑的内容:

已有很多提法:1) 基因克隆计划:基因鉴定作为科研项目有其双重性,它在 1993 年被列为美国 HGP 的重要内容之一。而它本身又是基于 HGP 的“spin-off”。HGP 的硬任务(遗传图、物理图、序列图)完成后,数目达 10 万之巨的基因鉴定还将是长期的艰巨的任务。而就表现型而言,基因鉴定又面临从单基因经典性状到多

基因复杂性状的转折。

2) “基因组多样性计划”:人类以基因的相似性而成为人类,又以多样性而成为具体的人。公孙龙在《白马篇》中的“白马非马”之说,充分体现了群体中个体多样性的含义。多样性计划有两个含义:一是多个族群基因组即群体多样化的研究;二是由来自一个人的“代表性基因组”,深入到每一个个体的基因组。

3) cDNA 计划:这一计划的目标是建立人体不同组织在不同时期、不同基因的不同表达“目录”,即人体基因表达的时空图。这一计划意味着从基因组研究到表达系列研究的转变。

上述两条计划都与基因鉴定一样,已作为 HGP 的一部分开始实施,又因超越 HGP 的原定硬任务而将在 HGP 后继续执行。

除此之外,还有:

4) Proteome (“蛋白质组”计划):这一计划的两个基点都来源于 HGP。①在技术上 HGP 的基因序列可以马上转化为信息的(电脑推定)和物质的(体外翻译)的蛋白质一级结构。②在策略上仿照 HGP 即从单一的蛋白质转向大规模的种、结构、功能的研究。这一计划意味着从研究基因的结构转向研究基因产物的结构与功能。

5) 细胞计划:细胞是自然界中生命活动的基本结构与功能单位。在阐明了以基因力核心的生命物质的本质后,以讯号传导为主线,阐明发生在细胞中一切生命现象的奥秘,是“细胞计划”的基本设想。五十年代生物学从细胞水平深入到分子水平曾经是一次科学的飞跃,今天生物学从分子水平上升到细胞水平仍将是一次飞跃。

当然,还有很多其它提法。

我们又回到了十年前 Dulbecco 提出的 HGP 时,面临的所有质疑。

我们怎么办?

在 70 年代,当遗传工程列为我国的重大科技项目时,我国的一位遗传学前辈说过:“我们需要补经典遗传学的课”。科学的发展阶段是很难超越的,我们嘲笑爬行,梦想飞跃,我们今

天还在为此付出代价。

今天 we 认识到,“分享数据”的真实含义是只有在积极参与的过程中,同步形成研究能力,才能真正分享这四张图即 HGP 的所有宝贵数据。我们需要补课,即形成我们自己的研究能力。遗憾的是至今我们也没有补好这一课。这个代价现在已慢慢展现:一些国外实验室特别是一些大公司,纷纷到我国来寻找遗传家系的材料,我国一些珍贵的家系材料正以多种合法或非法的途径流往国外。因为伴随 HGP 而来的“基因专利化”,正在改变生物学、医学基础研究的格局,迫使从事生物学、医学研究的各国实验室不得不重视克隆自己的基因。

二十一世纪,我们的基因诊断需要基因、基因治疗需要基因,我们遗传工程产业、我们的蛋白质计划需要基因,我们的“细胞计划”也需要基因。我们不能坐等人家来“掐脖子。”

有待克隆的未知基因肯定要比已知的多得多,重要得多。

我国克隆基因得天独厚。我国地广人多,

历史上形成了很多的族群(ethnic groups)与隔离群(isolates),保留了不少多世代、多个体的大家系。如果发现了典型的遗传性状,不管是正常还是疾病,是经典的遗传病还是癌症,或是别的常见病,都将是可克隆相关基因的珍贵材料。会有更多的国外实验室、公司把眼光投向我国的这一资源。

我国当前紧要做的事情是,一方面尽快挖掘我国宝贵的遗传病资源,发挥我国的资源优势;另一方面充分利用国际基因数据库中的基因信息,建立我国的基因组分析与基因克隆能力。

中国的 HGP,指导思想是参与、分享,重点是以我们自己的资源,依靠我们自己的力量,为我们子孙后代克隆我们自己的基因。(with ourselves, by ourselves, of ourselves, for ourselves)。

基因克隆计划,应该是我国的 HGP 与后 HGP 的连接点。

Transitions from the initiation of Human Genome Project(HGP)to post-HGP

--The 10th Anniversary of HGP

Abstract

It is worthwhile to have a retrospective review of HGP when its achievements have been highly acknowledged and listed as one of the greatest projects in the history of natural sciences 10 years after its initiation. HGP is both a challenge and an opportunity to China. China should reformulate its own HGP and focus on the gene identification by taking advantage of its rich genetic resources and by sharing the human genome data achieved by the world-wide HGP.