

“国际人类基因组计划”完成人类基因组 “工作框架图”

2000年6月26日,6国合作、公众支持的国际人类基因组计划协作组今天在全球同一时间宣布已完成人类生命的蓝图——人类基因组的“工作框架图”。这是“人类基因组计划”进展的一个重要的里程碑。

“工作框架图”的构建包括两大任务:首先是确定几十万个大片段DNA克隆在人类染色体上的位置,然后测定这些DNA大片段的化学组成单位——核苷酸(共四种,分别以A、T、C、G来表示)的序列,即将每一个位置上的核苷酸一一确定并连接成线,最后将测定多次的这些序列依其彼此重叠的相互关系及其在人类染色体上的位置,组装成代表人类整个基因组的DNA序列图。

已组装好的“工作框架图”已经能够覆盖人类基因组的97%,其中至少85%的序列已组装得准确无误。国际合作、公共支持的各中心的测序能力在去年出乎意料地翻了几番,每秒钟合计的测序量差不多达到1000多个核苷酸,并且夜以继日、分秒必争。60%以上的序列都是在短短的前6个月里完成的。已测序列的总长度已超过180亿核苷酸,差不多已把整个人类基因组测定了6次。组装好的、没有“空洞”的连续片段的平均长度为20万核苷酸左右。

“工作框架图”的总体质量远远超过了原先的估计,50%以上的序列已非常接近最终的“完成图”的质量要求,20%的序列已经达到“完成图”的标准。

根据所有参与者与资助者都认同的人类基因组计划的原则——所有的序列只要组装后达到了一定长度(一至二千个核苷酸)都必须在24小时内递交至公众基因数据库。这一原则得到了美国、英国、法国和其他国家的广泛赞同。1999年,中国参与了人类基因组计划,并对这一原则表示支持,由此推动了国际社会与组织对这些原则的一系列支持,并促使联合国教科文组织于今年5月发表了支持声明。正因为这样,人类基因组计划所产生的所有最新的序列信息都源源不断地向全世界公布,完全免费地供全球的科学家、产业界、特别是专门为生物产业提供信息服务的企业使用。对所有序列数据的下载、转递、再分析都没有任何方式的限制。应用“工作框架图”及其序列,已有数万个基因被鉴定,多达数十种与疾病有关的基因被定位、克隆。

尽管“工作框架图”已可满足生物、医学大多数研究的需要,但要全面地分析、彻底解读人类基因组还需要完美无缺、高度准确并经得起时间考验的“完成图”。“完成图”的要求是没有“工作框架图”尚能允许的“空洞”,每个核苷酸序列的准确率都要在99.99%以上。人类的21和22号染色体就已经达到了这样的要求。因此,国际人类基因组计划的重点已转入“完成图”的最终完成。这一任务原定在2003年完成,由于“工作框架图”的如期完成与比预期要高的质量,“完成图”的绘制成功将比原计划大大提前。

下一阶段的另一任务是尽快鉴定人类DNA序列的差异,单核苷酸多态性(SNP)。人类基因组计划原先计划在2003年前要发现10万SNPs。据今天的已组装的序列以SNP协作组积累的其他数据,科学家已发现了30万SNP,至年底将有100万个。现在人类基因组计划已完成的只是一个(或数个)无名个体的参照基因组序列,人类DNA序列的差异才能为人类疾病与人类历史的研究提供有力的工具。

公众支持的国际人类基因组计划协作组是由6个国家16个中心的1000多位科学家组成的。这6个成员国是美国、英国、日本、法国、德国和中国。参与这个计划的16个中心紧密合作,全面协调,使人类基因组每一区域都已“包干”到位,减少了重复劳动。所有中心都根据统一的质量标准,按照“百慕大原则”暨“人类基因组计划精神”,每天将数据递交至公共基因数据库。