2007年，新一代测序技术因其在众多领域中的成功应用而得到广泛承认。

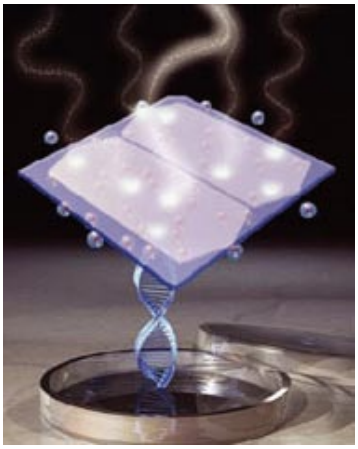
在北卡罗来纳州立大学（North Carolina State University）毒理学大楼里，小型基因组实验室主任Nigel Deighton和其他研究人员一道，依靠来自罗氏诊断公司（Roche Diagnostics）的仅能维持三个月研究的贷款，开发研制出实验室第一台新一代测序仪——454 GS FLX（见文后注释1）。他们花了几天时间，处理一个同事保存的细菌DNA，使其成为雾化状态，接着进行PCR扩增。随后他们将与磁珠连接的PCR产物点样在平板上（平板上有肉眼看不见的小孔），再将平板放入仪器，关上吊桥一样的舱门，仪器便开始自动运行，进行序列测定了，而不再需要人工操作，于是他们便决定一起出去喝上一杯。

第二天，Deighton浏览了该仪器生成的关于被测细菌基因组的最新测序图。在首次试运行中，GS FLX仪对待测序基因组测序的覆盖率达到原有仪器的6倍以上，平均阅读长度（read length）大约为250个碱基，“这个成绩对于第一次运行来说，算是不错的了。” Deighton说道。

在过去几年里，新一代测序技术平台在那些大型测序实验室中得到快速发展，各种新技术犹如雨后春笋般涌现。而在2007年，即使是在更小规模的实验室，例如Deighton的实验室，也得以更好的运用这种测序平台，George Church如是说，他在哈佛医学院（Harvard Medical School）研制出了相关的测序技术。

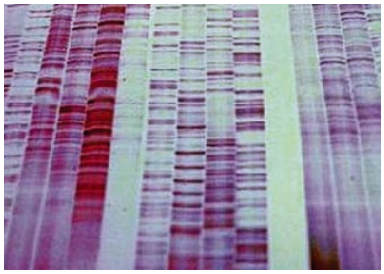
2005年推出的454 Life Sciences仪器，是进入市场的首个测序平台。该公司于2007年被罗氏诊断公司收购。而Solexa, Inc.在2006年6月也推出了自己的新一代测序仪器，该仪器将“合成测序”化学法（sequencing-by-synthesis chemistry）与“DNA簇技术”（DNA cluster technology）相结合，随后该公司被Illumina收购。2007年10月，Applied Biosystems宣布其研制出了ABI SOLiD系统，成为投入市场的唯一采用“基于序列连接的化学法”（ligation-based）的测序平台。在去年的全球基因组测序会议上，Church就曾指出上述测序平台在多方面得以应用。Broad研究院基因组测序及分析小组的副组长Chad Nusbaum说，研究人员们目前越来越多地将测序技术作为一种多功能手段去解决各种疑难问题。

2007年，研究人员用新旧两代测序平台对人类基因组进行了测序。Baylor医学院（Baylor College of Medicine）和454 Life Sciences公司的研究人员，用两个月时间，花费一百万美元，对James Watson的基因组进行了测序。除此之外，还完成了其他两个个体基因组的测序，一个是Craig Venter，对Venter基因组的测序，就是以其建立的研究为基础而进行；另一个测序对象是一位中国人，测序工作在北京基因组研究所进行。J. Craig Venter



454生命科学公司所研发的新一代测序平台以光纤微流体技术和包裹了待测DNA片段的乳化液滴技术为基础

图片来源：*454 Life Science*

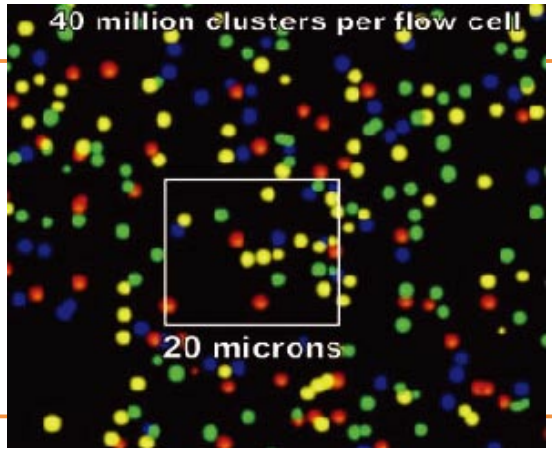


经典的桑格测序法已经成为老一代的测序技术

图片来源：*Nature*

研究所对Venter的DNA的测序采用的是Sanger法，耗时数年，共花费七千万美元左右。目前，该研究所致力于对更多个体的全基因组进行测序，测序工作将以新一代测序技术为平台，同时以Sanger法所得到的参考序列为基础，J. Craig Venter研究所的联合技术中心（J. Craig Venter Institute's Joint Technology Center）科学主任Yu-Hui Rogers这样对我们说。

耶鲁大学（Yale University）研究人员与454 Life Sciences公司研究人员合作，将454测序技术与一种称为“paired-end mapping”的技术（见文后扩展阅读）相结合，检测人类基因组内的基因结构变异，研究结果发表于2007年10月的 《科学》杂志上。“该技术使我们有了新的发现。”该研究的第一作者Michael Snyder说，他所指的是基因序列中的某些基因转位与断裂点，如果采用比较基因组杂交及fosmid法，这些位点是无法检测出来的。Illumina DNA测序部门总经理John West和Applied Biosystem's SOLiD测序系统研发人员之一Kevin McKernan都表示，他们的公司都采用了“paired-end mapping”技术法对不同大小的基因结构变异进行研究。



图中彩色圆点表示在Solexa测序仪上采用可逆性终止反应（reversible terminators）读序的待测序列

图片来源：*Illumina*

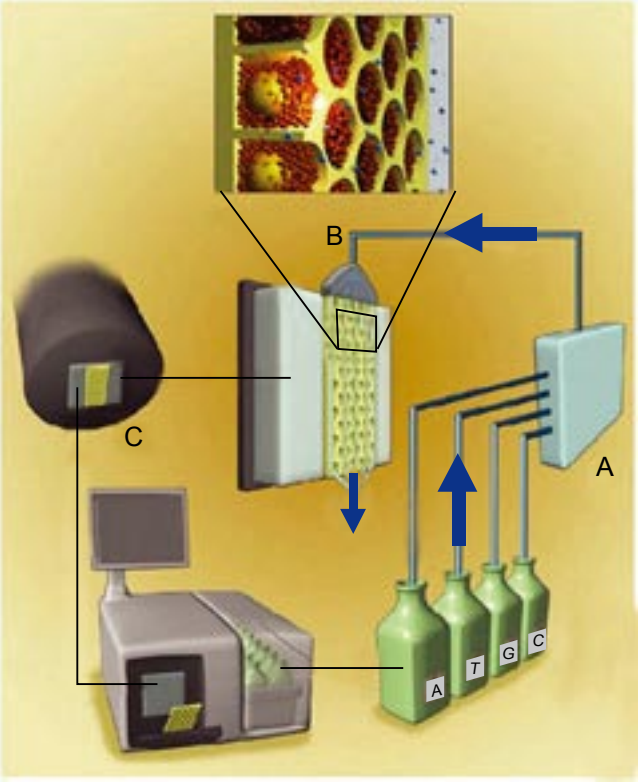
新一代测序技术对于较短的遗传序列，如与组蛋白结合的DNA片断、转录因子等的测定来说，亦是一大福音。来自美国国家心肺血液研究所（US National Heart, Lung and Blood Institute）的Keji Zhao所带领的研究小组，采用了染色质免疫沉淀分析（ChIP）技术对组蛋白修饰进行分析研究，研究结果发表于2007年的《细胞》杂志上。在其研究中，正是利用了Solexa平台较短的阅读长度这一特点，该研究也是最早一批运用Solexa技术的项目。

Zhao说，以前我们运用传统的ChIP与基因表达系列分析（SAGE）技术相结合的方法进行此类研究，其中SAGE采用Sanger法对串联DNA标签进行测序。每完成一个组胺酸修饰的序列测序，需要花费将近10万美元；而如今采用Solexa仪器进行测序，每一个标签测序只需要2,000美元，大大降低了费用，节省下来的经费可用于对更多标签的测定。“没有这新一代的测序平台，我们不可能进行目前规模的任何研究。”Zhao说。另一位科学家Elaine Mardis说，新的测序平台对于Zhao的研究领域来讲是非常理想的，因为该技术可以直接对基因片段进行测序及鉴定，无需再用杂交法对序列进行鉴定。

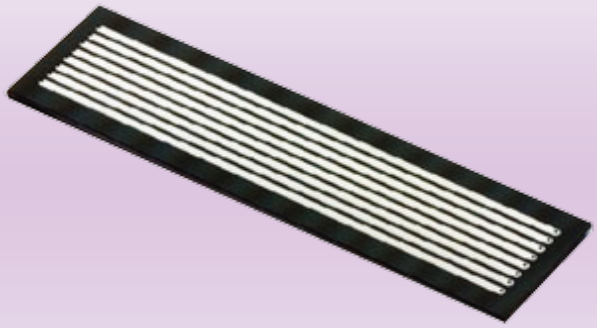
Mardis接着说道，环境基因组学（见注释2）这一新兴研究领域也从新一代测序技术的发展中获益。为揭示人体健康疾病状态与人体微生物群系的关系，美国国立卫生研究院（NIH）于2007年12月19日宣布正式启动的一项新的基因工程——人体微生物群系项目（HMP）就是一个例子。2007年9月，来自哥伦比亚大学，以W. Ian Lipkin为首的研究小组与454 Life Sciences联合在《科学》杂志上发表了一篇研究报告。报告中指出，他们在实验中鉴定出了一种与数十亿蜜蜂死亡有关的病毒。“环境基因组学的研究在传统测序时代是非常昂贵且繁琐的，” Mardis说：“现在，随着新一代测序平台的出现，这些实验变得简单易行了，而且也更容易从中获得新的信息。”

454 Life Sciences测序仪的内部工作原理图

图片来源：*454 Life Sciences*



Church预言说，在2008年，人们将能够借助新的技术平台，更加透彻地认识基因序列与功能之间的关系。事实上，到目前为止，遗传序列与功能之间的关系的研究在2007年已经取得了很大的进步，速度由之前的一年分析一个序列显著提高到2007年的一年能分析研究60个序列。“那是个相当了不起的进步，但是对于绝大部分序列，并没有继续研究以找到形成相应功能的等位基因或等位片段。” Church说。想要了解序列与功能之间的确切关系，需要更多个体的基因组序列及相关表型数据。来自美国能源部联合基因组研究所（US Department of Energy's Joint Genome Institute）的遗传学家Len Pennacchio说：“从这一点上看，所测得的Watson的DNA还只是相当于一份组成基因组各项“零部件”的清单，尚不能够告诉我们基因组的各个因子是如何协调工作，使我们成为独一无二的个体的。”



在芯片上进行的测序：Illumina测序平台

图片来源：*Illumina*

“早在2004年美国国立卫生研究院就提出研究目标——1000美元的成本进行人类基因组测序。尽管新一代测序技术还不能达到这个目标，但相关的研究人员正在向这个目标迈进。”Richard Mayers是斯坦福人类基因组中心主任，他这样告诉我们。

目前看来，所要进行的下一步应该是建立起更为先进的计算机算法，以用于基因组测序过程中的拼接和比较。基因组测序中成千上万的序列片段需要研究人员深入研究以便鉴别出冗余序列，从而保证阅读框的正确。“这是个棘手的问题，但是我们已经开始将新一代技术应用于此。” Richard说。

同时，Deighton还将把454 Life Sciences技术应用于北卡罗来纳州立大学目前的研究中。他们已经将细菌基因组进行了拼接，并且希望能够达到20倍的基因组覆盖率，这样就可以填补基因片段间的空隙。大约15名研究人员已经签订协议，将使用该测序仪器对各种生物物种进行测序。到3个月的贷款期期满之际，Deighton会进行经费分析，让同事们认识到，尽管该仪器售价高达50万美元，但其带来的低成本测序效益却是无法想象的。

原文检索：http://www.nature.com/nmeth/journal/v5/n1/full/nmeth1154.html

编者注：

1. GS FLX 罗氏诊断公司和454公司不断加大创新投入，又于2007年推出了通量更大，读长更长，准确性更高的第二代基因组测序系统——Genome Sequencer FLX System （GS FLX）。GS FLX是世界上第一个可以提供超高通量基因组和长片段DNA测序的商业化系统设备，只需要一步简单的全基因组样品制备过程，就可以在7.5小时内获得超过1亿个碱基信息。

2. 环境基因组（metagenomics）是指同时被研究的整个微生物群落的DNA。从历史观点上来看，微生物学主要集中于研究能够在实验室培养并在显微镜下检测的单个种类。但大多数微生物的复杂群落无法在实验室里培养，环境基因组学研究将使研究人员能够研究整个微生物群落，进而改变现代微生物学。这些被研究的微生物大多数是此前未知，且不能在实验里培养的物种。研究目的在于弄清微生物间的相互作用如何平衡大气组成、对抗疾病及支持植物生长。环境基因组研究是从生活于一个特定环境样本中的所有微生物中提取出DNA。提取出的遗传物质由上百万个DNA随机碎片组成。新的DNA序列测定技术和功能更强大的计算机，能够帮助研究人员让这些环境基因组图谱变得有意义。

扩展阅读：http://www.sciencemag.org/cgi/content/abstract/1149504v1

原文检索：Nature Methods - 5, 11 - 14 （2008） doi:10.1038/nmeth1154

筱玥 编译