

DNA测序技术发展及其展望

孙海汐^{1,2} 王秀杰¹

1. 中国科学院遗传与发育生物学研究所植物基因组学国家重点实验室, 北京100101
2. 中国科学院研究生院, 北京100049

摘要 DNA测序技术是现代生物科学研究中重要的手段之一。自从1977年第一代测序技术问世以来, 经过三十几年的努力, DNA测序技术已经取得了很大的发展, 在第一代和第二代测序技术的基础上, 以单分子测序为特点的第三代测序技术已经诞生。本文回顾了每一代测序技术的原理和特点, 并对测序技术的发展趋势和它在生物科学研究中的应用做了展望, 以期更好地帮助人们理解测序技术在生命科学研究中的重要作用。

关键词: 第一代测序技术; 第二代测序技术; 第三代测序技术; 基因芯片技术

The Development and Future Perspectives of DNA Sequencing Technology

Sun Haixi^{1,2}, Wang Xiujie¹

1 State Key Laboratory of Plant Genomics, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China

2 Graduate University of Chinese Academy of Sciences, Beijing 100049, China

Abstract: DNA sequencing technology has played important roles in modern biological research. The first generation of sequencing technology appeared in 1977. During the past decade, the so-called next generation sequencing technologies had emerged and caused a revolution in biological research. Recently, the third generation sequencing technology, which is able to determine the base composition of single DNA molecules, has joined the race. The mechanisms and features of each generation of sequencing technology and their future perspectives will be discussed here.

Keywords: first generation sequencing technology; second generation sequencing technology; third generation sequencing technology; microarray

1. 引言

快速和准确地获取生物体的遗传信息对于生命科学研究一直具有十分重要的意义。对于每个生物体来说，基因组包含了整个生物体的遗传信息。测序技术能够真实地反映基因组DNA上的遗传信息，进而比较全面地揭示基因组的复杂性和多样性，因而在生命科学研究中扮演了十分重要的角色。

测序技术最早可以追溯到20世纪50年代，早在1954年就已经出现了关于早期测序技术的报导，即Whitfeld等用化学降解的方法测定多聚核糖核苷酸序列^[1]。1977年Sanger等发明的双脱氧核苷酸末端终止法和Gilbert等发明的化学降解法，标志着第一代测序技术的诞生^[2, 3]。此后在三十几年的发展过程中陆续产生了第二代测序技术，包括Roche公司的454技术^[4, 5]、Illumina公司的Solexa技术^[6-8]和ABI公司的SOLiD技术^[9, 10]。最近，Helicos公司的单分子测序技术^[11-13]、Pacific Biosciences公司的单分子实时(Single Molecule Real Time, SMRT)测序技术^[14, 15]和Oxford Nanopore Technologies公司正在研究的纳米孔单分子测序技术^[16]被认为是第三代测序技术^[17]。测序技术正在向着高通量、低成本、长读取长度的方向发展^[18]。

2. 第一代测序技术

早在1954年，Whitfeld等就提出了测定多聚核糖核苷酸链的降解法^[1]，该方法利用磷酸单酯酶的脱磷酸作用和高碘酸盐的氧化作用从链末端逐一分离寡核糖核苷酸并测定其种类。但由于操作复杂等原因，该方法并没有被广泛应用。

在1977年，Sanger等提出了经典的双脱氧核苷酸末端终止测序法^[2]。该方法的原理是：由于ddNTP的2'和3'都不含羟基，在DNA合成反应中不能形成磷酸二酯键，因此可以被用来中断DNA合成反应。在4个DNA合成反应体系中分别加入一定比例的带有放射性同位素标记的某种ddNTP，通过凝胶电泳和放射自显影后，可以根据电泳带的位置确定待测分子的DNA序列。同一年，Gilbert等提出了化学降解法^[3]。该方法与Sanger法类似，都是先得到随机长度的DNA链，再通过电泳方法读出序列。二者的不同之处在于，Gilbert法是用特定的化学试剂标记碱基再用化学方法打断待测序列，而Sanger法是通过ddNTP随机中断合成待测序列。

此后，在Sanger法的基础上，80年代中期出现了以荧光标记代替放射性同位素标记、以荧光信号接收器和计算机信号分析系统代替放射性自显影的自动测序仪^[19]。另外，90年代中期出现的毛细管电泳技术使得测序的通量大为提高^[20]。

除此之外，这一时期还出现了一些其他的测序方法，如焦磷酸

测序法(pyrosequencing)^[21]、连接酶测序法(sequencing by ligation, SBL)^[22]、杂交测序法(sequencing by hybridization, SBH)^[23]等。其中焦磷酸测序法即为后来Roche公司454技术使用的测序方法，连接酶测序法即为后来ABI公司SOLiD技术使用的测序方法。

3. 第二代测序技术

尽管第一代测序技术已经帮助人们完成了从噬菌体基因组到人类基因组草图等大量的测序工作，但由于其存在成本高、速度慢等方面的不足，并不是最理想的测序方法。经过不断的开发和测试，进入21世纪后，以Roche公司的454技术^[4, 5]、Illumina公司的Solexa技术^[6-8]和ABI公司的SOLiD技术^[9, 10]为标志的第二代测序技术诞生了。与第一代技术相比，第二代测序技术不仅保持了高准确度，而且大大降低了测序成本并极大地提高了测序速度。使用第一代Sanger的测序技术完成的人类基因组计划，花费了30亿美元巨资，用了三年的时间；然而，使用第二代SOLiD的测序技术，完成一个人的基因组测序现在只需要一周左右的时间。由于第二代测序技术产生的测序结果长度较短，因此比较适合于对已知序列的基因组进行重新测序，而在对全新的基因组进行测序时还需要结合第一代测序技术。下面对三种第二代测序技术的原理和特

点分别加以具体介绍。

1) Roche公司的454测序技术

454测序系统是第二代测序技术中第一个商业化运营的测序平台^[4]。它的原理是^[4, 24, 25] (图1^[5])：

① 待测DNA文库的构建

把待测序列用喷雾法 (nebulization) 打断成300–800 bp 的小片段并在小片段两端加上不同的接头，或将待测序列变性后用杂交引物进行PCR扩增，连接载体，构建单链DNA (ssDNA) 文库。

② Emulsion PCR

将这些ssDNA与水油包被的直径大约28 μm的磁珠在一起孵育、退火，由于磁珠表面含有与接头互补的寡聚核苷酸序列，因此

ssDNA会特异地连接到磁珠上。同时孵育体系中含有PCR反应试剂，因此可以保证每一个与磁珠结合的小片段都会在各异的孵育体系内独立扩增，扩增产物仍可以结合到磁珠上。反应完成后，破坏孵育体系并富集带有DNA的磁珠。经过扩增反应，每一个小片段都将被扩增大约100万倍，从而达到下一步测序反应所需的模板量。

③ 测序

预先用 *Bacillus stearothermophilus* 聚合酶和单链结合蛋白处理带有DNA的磁珠，然后将磁珠放置在一种叫做PicoTiterPlate (PTP) 的平板上。PTP板上含有很多直径约为44 μm 的小孔，每个小孔仅能容纳一个磁珠，通过这种方法固定每个磁

珠的位置以监测接下来的测序反应。测序反应采用焦磷酸测序法，将一种含有比PTP板上小孔直径更小的磁珠放入小孔，启动测序反应。测序反应以磁珠上大量扩增的ssDNA为模板，每次反应加入一种dNTP进行合成反应。如果这种dNTP能与待测序列配对，则会在合成后释放焦磷酸基团。释放的焦磷酸基团会与反应体系中的ATP硫酸化酶反应形成ATP。生成的ATP和荧光素酶共同氧化反应体系中的荧光素分子并发出荧光。测序反应产生的荧光信号由放置在PTP板另一侧的CCD照相机记录，再经过计算机分析转换为测序结果。由于每种dNTP在反应中产生的荧光颜色不同，因此可以根据荧光的颜色来确定被测分子的序列。反应结束后，游离的dNTP会在双磷酸酶的作用下降解ATP，导致荧光淬灭，从而使反应体系再生。作为一个反应器，由于PTP板上每个小孔之间相互独立，因而大大降低了反应的干扰和误差。

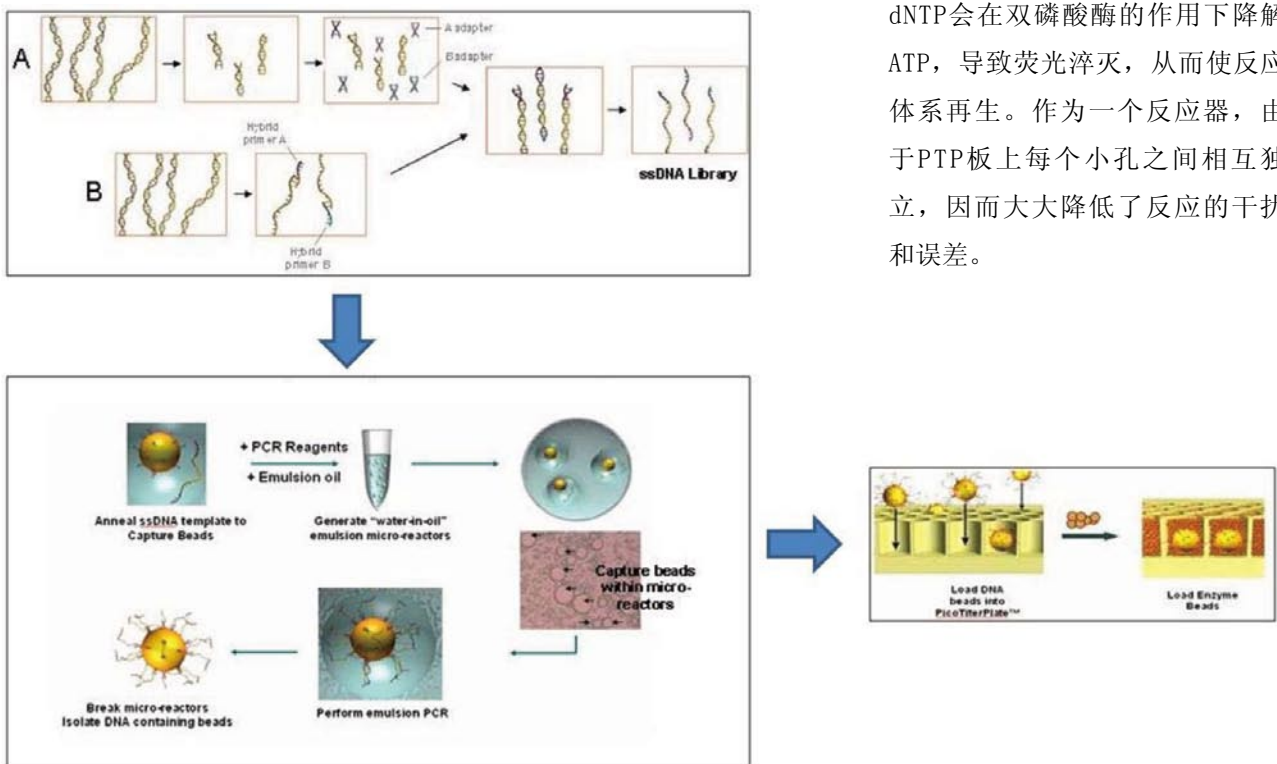


图1 454测序技术流程

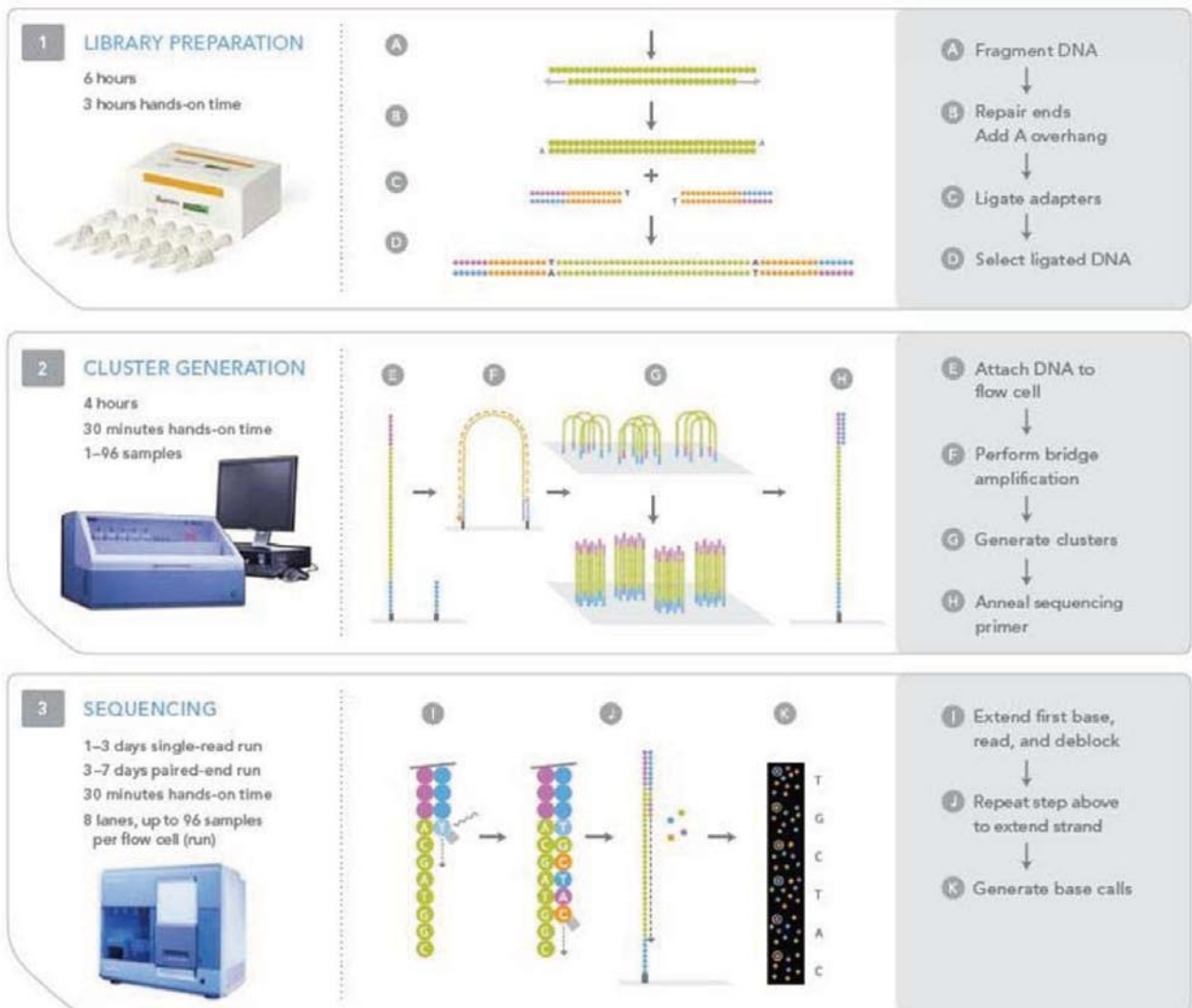


图2 Solexa测序技术流程

在2008年，Roche公司推出了454技术最新的GS FLX Titanium系列试剂和软件，提升了读取长度与测序通量，使454技术的平均读取长度达到400 bp，每个循环能产生总量为400-600 Mb的序列，耗时约10小时^[5]。454技术的主要缺点是无法准确测量同聚物(homopolymer)的长度。例如当待测序列中出现Poly(A)的情况下，测序反应中会一次加上多个T，而加入T的数目只能从荧光信号的强度来推测，有可能造成结果不准确。也正是因为这个原因，454技

术主要的错误不是来自核苷酸的替换，而是来自插入或缺失。454技术最大的优势在于较长的读取长度，使得后继的序列拼接工作更加高效、准确。

2) Illumina公司的Solexa技术
Illumina公司的Genome Analyzer于2006年问世，它是一种基于Solexa技术的测序系统。该技术利用边合成边测序的方法(sequence by synthesis, SBS)，它的原理是^[6, 7, 25] (图2^[8])：

① 待测DNA文库的构建
把待测序列打断成200-500 bp的小片段，并在小片段两端加上不同的接头，连接载体，构建ssDNA文库。

② DNA与流动槽(flow cell)的附着将这些ssDNA随机地附着在流动槽表面的channel上。流动槽是一种含有8个channel的微纤维板，它的表面固定有很多接头，能支持ssDNA在其表面进行桥式扩增。

③ Bridge PCR
向反应体系中添加未标记的

核苷酸和酶,进行Bridge PCR扩增反应。Bridge PCR以流动槽表面固定的接头为模板,经扩增将桥型ssDNA扩增成桥型dsDNA。

④ dsDNA的变性

将桥型dsDNA变性成ssDNA,继续扩增。经过不断的扩增变性循环,每一种ssDNA都在各自的位置集中成束(cluster),每一个束含有单个模板分子的500-1000个克隆拷贝,从而达到能支持下一步测序反应所需信号强度的模板量。

⑤ 测序

测序方法采用边合成边测序的方法(SBS)。向反应体系中同时添加DNA聚合酶、接头引物和带有碱基特异荧光标记的4种dNTP。由于这些dNTP的3'羟基被化学方法保护,因而每轮合成反应都只能添加一个dNTP。在dNTP被添加到合成链上后,所有未使用的游离dNTP和DNA聚合酶会被洗脱。加入激发荧光所需的缓冲液,用激光激发荧光信号,用光学设备完成荧光信号的记录,再通过计算机分析转化为测序结果。当荧光信号的记录完成后,加入化学试剂淬灭荧光信号并去除dNTP的3'羟基保护基团,以便进行下一轮测序反应。

在2009年,Solexa推出了对读测序(paired-end)方法,即在构建待测DNA文库时在两端的接头上都加上测序引物结合位点,在第一轮测序完成后,去除第一轮测序的模板链,用对读测序模块(Paired-End Module)引导互

补链在原位置再生和扩增,以达到第二轮测序所用的模板量,进行第二轮互补链的合成测序。

使用对读测序方法,Solexa技术的读取长度可以达到 2×75 bp,相比454技术,其后续的序列拼接工作的计算量和难度均大大增加。Solexa技术每个循环能获得20.5-25 Gb的测序结果,耗时约9.5天^[8]。由于Solexa技术在合成中每次只能添加一个dNTP,因此很好地解决了同聚物长度的准确测量问题。Solexa技术主要的错误来源是核苷酸的替换,而不是插入或缺失,目前它的错误率大约在1-1.5%之间。

3) ABI公司的SOLiD技术

ABI公司的SOLiD测序系统于2007年10月投入商业使用,它基于连接酶测序法,即利用DNA连接酶在连接过程中进行测序。它的原理是^[9, 10, 26, 27](图3^[10]):

① 待测DNA文库的构建

把待测序列打断成很小的片段,并在小片段两端加上不同的接头,连接载体,构建ssDNA文库。

② Emulsion PCR

与454技术的Emulsion PCR类似,将带接头的ssDNA固定在磁珠表面,进行PCR扩增,并对扩增产物进行3'端修饰。

③ 连接酶测序

将3'端修饰的磁珠沉积于上样玻片(slide),进行连接酶测序。在沉积过程中可对磁珠密度进行调节,以达到最大通量。向

体系中加入DNA连接酶、通用测序引物n和具有3'-XXnnnzzz-5'结构的八聚核苷酸。在这个八聚核苷酸中,第1和第2位(XX)上的碱基是确定的,并根据种类的不同在第6-8位(zzz)上加了不同的荧光标记。这种由两个碱基决定的测序方法被称为两碱基测序(two base encoding)。当八聚核苷酸由于第1和第2位配对而被连接酶连接上时,会发出荧光。在记录下荧光信息后,通过化学方法在第5和第6位之间进行切割,淬灭荧光信号,以进行下个位置的测序。通过这种方法,每次测序的位置都相差五位,即第一次测第1和第2位,第二次测第6和第7位……在测到末尾后,将新合成的链变性、洗脱。而后用通用测序引物n-1进行第二轮测序。通用测序引物n-1与通用测序引物n的差别是,二者在与接头配对的位置上相差一个碱基,即通用测序引物n-1在通用测序引物n配对位置上向3'端移动了一个碱基。因此在加入DNA连接酶和八聚核苷酸后,可以测定第0和第1位、第5和第6位……第二轮测序完成后,接下来再分别加入通用测序引物n-2、通用测序引物n-3、通用测序引物n-4进行第三轮、第四轮、第五轮测序,最终可以完成全部位置的测定,并且每个位置均被测定了两次。

SOLiD技术每个循环可以测两个上样玻片,读取长度可达 2×50 bp,与Solexa技术类似,后续的序列拼接工作也比较复杂。SOLiD

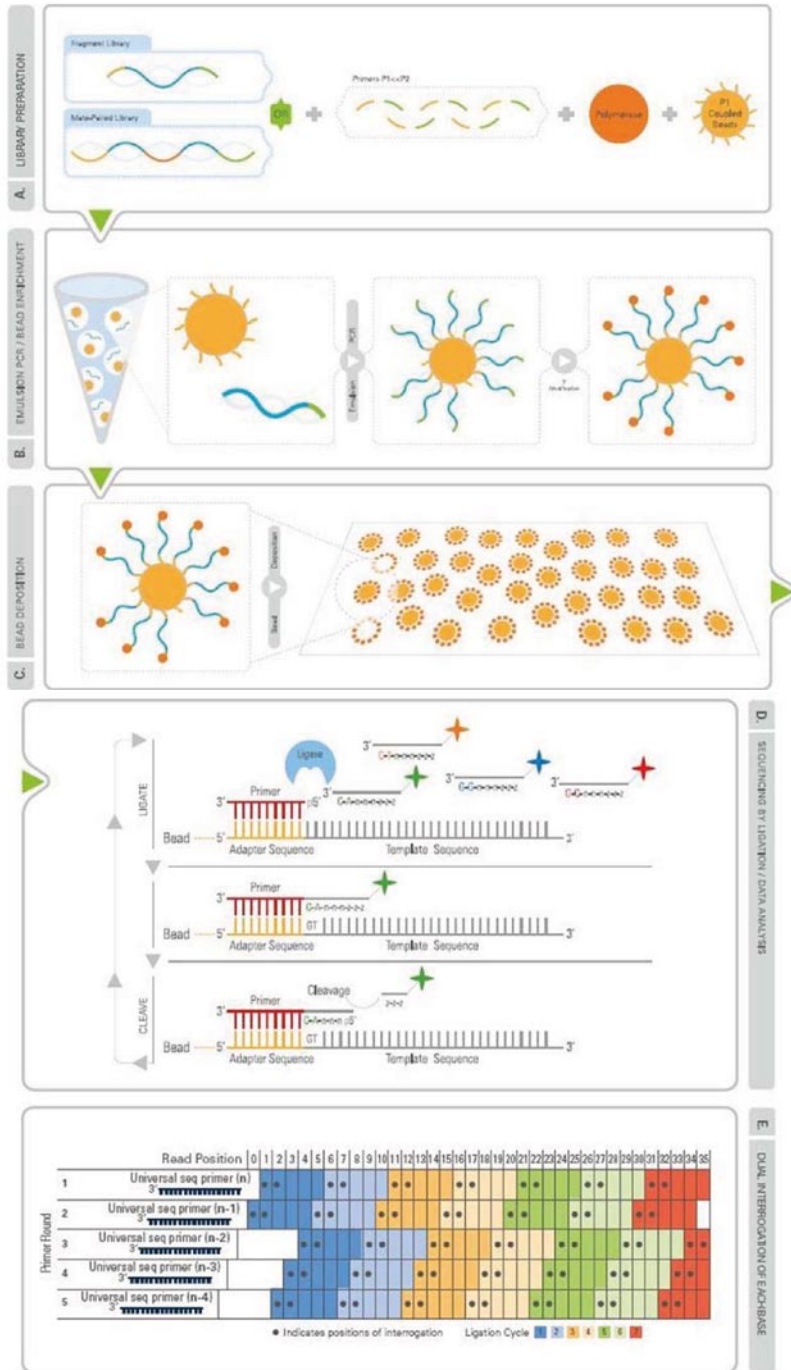


图3 SOLiD测序技术流程

► 技术每个循环的数据产出量为10-15 Gb, 耗时约为6-7天。由于采用两碱基测序, 该技术的准确率能达到99.94 %以上^[10]。

4. 第三代测序技术

近期出现的Helicos公司的

Helicoscope单分子测序仪、Pacific Biosciences公司的SMRT技术和Oxford Nanopore Technologies公司正在研究的纳米孔单分子技术, 被认为是第三代测序技术。与前两代技术相比, 他们最大的特点是单分子测序。其中, Helicoscope技术和SMRT技术利用荧

光信号进行测序, 而纳米孔单分子测序技术利用不同碱基产生的电信号进行测序。

Helicos公司的Helicoscope单分子测序仪基于边合成边测序的思想, 将待测序列随机打断成小片段并在3'末端加上Poly(A), 用末端转移酶在接头末端加上Cy3荧光标记。用小片段与表面带有寡聚Poly(T)的平板杂交。然后, 加入DNA聚合酶和Cy5荧光标记的dNTP进行DNA合成反应, 每一轮反应加一种dNTP。将未参与合成的dNTP和DNA聚合酶洗脱, 检测上一步记录的杂交位置上是否有荧光信号, 如果有则说明该位置上结合了所加入的这种dNTP。用化学试剂去掉荧光标记, 以便进行下一轮反应。经过不断地重复合成、洗脱、成像、淬灭过程完成测序。Helicoscope的读取长度约为30-35 bp, 每个循环的数据产出量为21-28 Gb。值得注意的是, 在测序完成前, 各小片段的测序进度不同。另外, 类似于454技术, Helicoscope在面对同聚物时也会遇到一些困难。但这个问题并不会十分严重, 因为同聚物的合成会导致荧光信号的减弱, 可以根据这一点来推测同聚物的长度。

此外, 可以通过二次测序来提高Helicoscope的准确度, 即在第一次测序完成后, 通过变性和洗脱移除3'末端带有Poly(A)的模板链, 而第一次合成的链由于5'末端上有固定在平板上的寡聚Poly(T), 因而不会被洗脱掉。第二次测序以第一次合成的链为模板, 对其

反义链进行测序。对Heliscope来说，由于在合成中可能掺有未标记的碱基，因此其最主要的错误来源是缺失。一次测序的缺失错误率约为2-7%，二次测序的缺失错误率约为0.2-1%。相比之下替换错误率很低，一次测序的替换错误率仅为0.01-1%。总体来说，采用二次测序方法，Heliscope可以实现目前测序技术中最低的替换错误率，即0.001%^[12, 13, 28]。

Pacific Biosciences公司的SMRT技术基于边合成边测序的思想，以SMRT芯片为测序载体进行测序反应。SMRT芯片是一种带有许多ZMW(zero-mode waveguides)孔的厚度为100 nm的金属片。将DNA聚合酶、待测序列和不同荧光标记的dNTP放入ZMW孔的底部，进

行合成反应^[29]。与其他技术不同的是，荧光标记的位置是磷酸基团而不是碱基。当一个dNTP被添加到合成链上的同时，它会进入ZMW孔的荧光信号检测区并在激光束的激发下发出荧光，根据荧光的种类就可以判定dNTP的种类。此外由于dNTP在荧光信号检测区停留的时间（毫秒级）与它进入和离开的时间（微秒级）相比会很长，所以信号强度会很大。其它未参与合成的dNTP由于没进入荧光型号检测区而不会发出荧光。在下一个dNTP被添加到合成链之前，这个dNTP的磷酸基团会被氟聚合物（fluoropolymer）切割并释放，荧光分子离开荧光信号检测区（图4^[15]）。SMRT技术的测序速度很快，利用这种技术测序

速度可以达到每秒10个dNTP^[15, 30]。

Oxford Nanopore Technologies公司正在研究的纳米孔单分子技术是一种基于电信号测序的技术。他们设计了一种以 α -溶血素为材料制作的纳米孔，在孔内共价结合有分子接头环糊精（图5^[31]）。用核酸外切酶切割ssDNA时，被切下来的单个碱基会落入纳米孔，并和纳米孔内的环糊精相互作用，短暂地影响流过纳米孔的电流强度，这种电流强度的变化幅度就成为每种碱基的特征。碱基在纳米孔内的平均停留时间是毫秒级的，它的解离速率常数与电压有关，180 mV的电压就能够保证在电信号记录后将碱基从纳米孔中清除。纳米孔单分子技术的另一大特点是能够

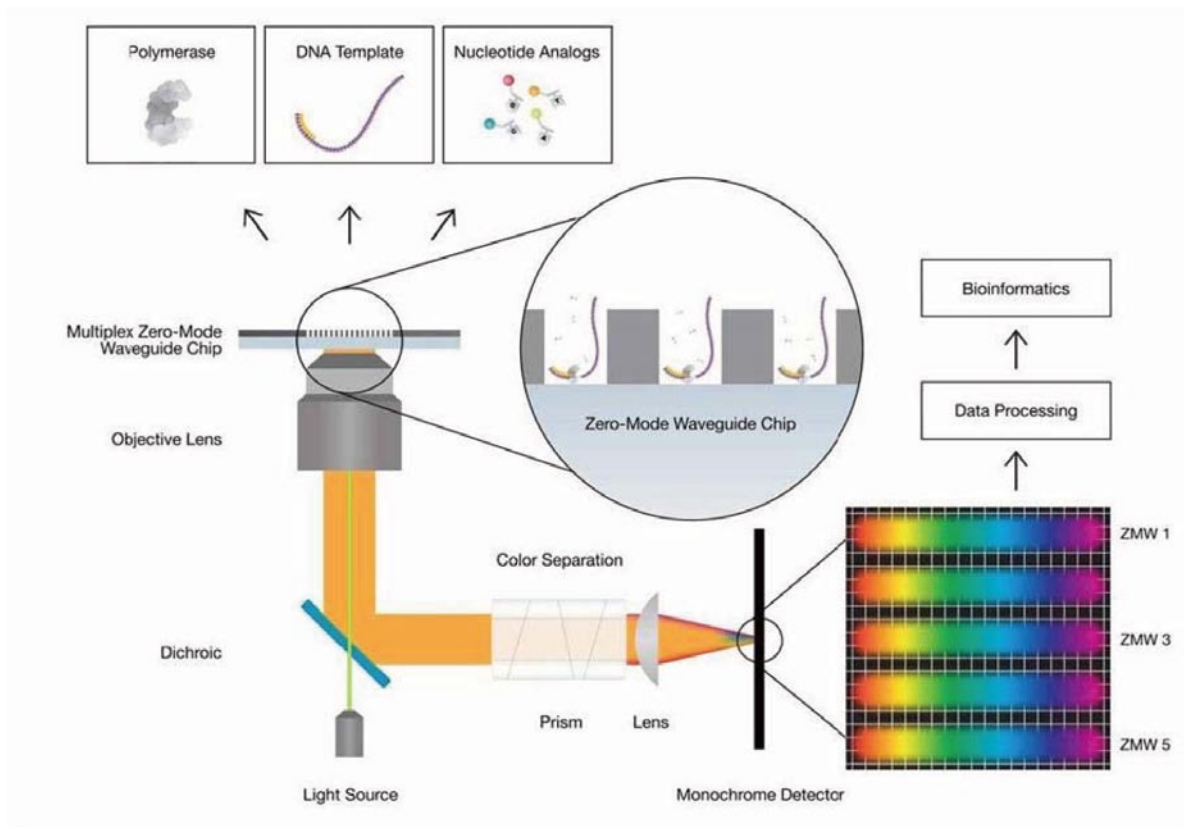


图4 SMRT测序技术流程

►直接读取甲基化的胞嘧啶，而不像传统方法那样必须要用重亚硫酸盐 (bisulfite) 处理，这对于在基因组水平研究表观遗传相关现象提供了巨大的帮助。纳米孔单分子技术的准确率能达到99.8%，而且一旦发现替换错误也能较容易地更改，因为4种碱基中的2种与另外2种的电信号差异很明显，因此只需在与检测到的信号相符的2种碱基中做出判断，就可修正错误。另外由于每次只测定一个核苷酸，因此该方法可以很容易地解决同聚物长度的测量问题。该技术尚处于研发阶段，目前面临的两大问题是寻找合适的外切酶载体以及承载纳米孔平台材料^[17]。

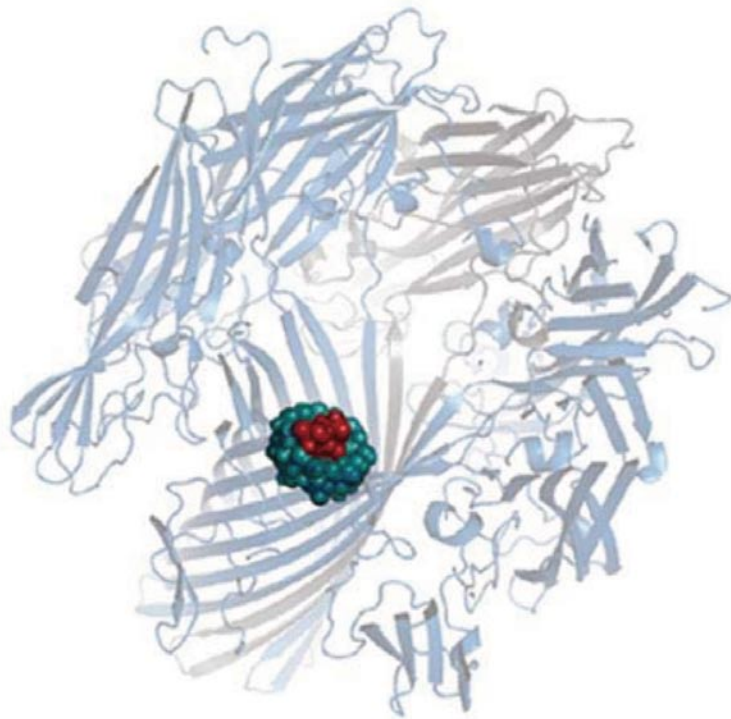


图5 纳米孔单分子结构

5. 测序技术的发展及应用趋势

三代测序技术特点的比较见表1。测序成本、读取长度和测序通量是评价测序技术先进与否的重要标准。测序成本是一个很重要的因素，它在一定程度上决定了基因组测序应用的普及性。在1995年，随着自动测序仪的出现，检测一个碱基的成本大约是1美元。随后到1998年，使用ABI Prism® 3700 DNA Analyzer检测一个碱基的成本已经降到了0.1美元。而目前广泛使用的第二代测序技术的测序成本更低。National Human Genome Research Institute (NHGRI) 预测，在未

来5年内，测序成本还将下降100倍。读取长度是指一个反应能测得数据的长度，它对测序成本和数据质量有很大的影响。更长的读取长度可以减少测序后的拼接工作量，但可能也会降低测序结果的准确性。测序通量是指在一定时间内获得的数据输出量。仪器的测序通量是指在样品准备充分的前提下，测序仪每24小时产生的数据量。更高的测序通量也能够一定程度上降低测序成本、提高科研工作的效率。

第二代测序技术与第一代Sanger测序法的原理都是基于边合成边测序的思想，在测序原理上没有本质的飞跃，那么测序的时间和费用都大大降低的原因是什么呢？其关键在于第二代测序技术采用了高通量测序技术，使测序通量大大提高，从Sanger测序法一次读取一条序列到毛细管测序的一次读取96条序列再到现在的一次读取几百万条序列的实现，不得不说这是对第一代测序技术的一次革命性的变革。

然而第二代测序技术并不完美，由于其在测序前要通过PCR手段对待测片段进行扩增，因此增加了测序的错误率。并且由于Illumina和SOLiD的测序结果都较短，比较适合重测序，而不太适用于没有基因组序列的全新测序。第三代测序技术解决了错误率的问题，通过增加荧光的信号强度及提高仪器的灵敏度等方法，使测序不再需要PCR扩增这个环节，实现了单分子测序并继承


	测序方法/平台	公司/公司网站	方法/酶	测序长度	每个循环的数据产出量	每个循环耗时	主要错误来源
第一代测序技术	Sanger/ABI3730 DNA Analyzer	Applied Biosystems www.appliedbiosystems.com	Sanger法/DNA聚合酶	1000 bp	56 Kb		
第二代测序技术	454/GS FLX Titanium Series	Roche www.roche-applied-science.com	焦磷酸测序法/DNA聚合酶	400 bp	400-600 Mb	10 h	插入、缺失
	Solexa/Illumina Genome Analyzer	Illumina www.illumina.com	边合成边测序/DNA聚合酶	2*75 bp	20.5-25 Gb	9.5 d	替换
	SOLiD/SOLiD 3 system	Applied Biosystems www.appliedbiosystems.com	连接酶测序/DNA连接酶	2*50 bp	10-15 Gb	6-7 d	替换
第三代测序技术	Heliscope/Helicos Genetic Analysis System	Helicos www.helicosbio.com	边合成边测序/DNA聚合酶	30-35 bp	21-28 Gb	8 d	替换
	SMRT	Pacific Biosciences www.pacificbiosciences.com	边合成边测序/DNA聚合酶	100000 bp			
	纳米孔单分子	Oxford Nanopore Technologies www.nanoporetech.com	电信号测序/核酸外切酶	无限长			

表1 三代测序技术特点的比较

了高通量测序的优点。目前正在研发的纳米孔单分子技术则更是在原理上做出本质变革，不再基于目前所用测序技术广泛使用的边合成边测序的思想，而是使用外切酶从ssDNA的末端逐个切割形成单碱基，并采用新技术对切落下来的单碱基进行检测，这样可以更好地提高读取长度，减少测序后的拼接工作量，实现对未知基因组进行重新测序。

高通量是第二代和第三代测序技术的共同特点。由于这些测

序技术一次可获得上百万条、甚至几百万条序列信息，因此可对某一组织、某一时间表达的所有mRNA进行序列测定，故又被称为深度测序^[32]。由于使用高通量测序技术可以实现对不同组织mRNA表达差异进行检测，通过与基因组进行比较可以确定组织特异表达的基因，因此具有取代“基因芯片”技术的态势。由于基因芯片的检测范围取决于芯片上探针的信息，因此只能检测人们已知序列的特征，缺乏发现寻找新

基因的能力，而高通量测序则能够很好地弥补基因芯片这方面的不足。但就目前而言，高通量测序技术建立的时间还很短，技术不是很成熟，其信息储备量也很有限；而基因芯片技术已经发展了近16年，其实验技术及后期数据分析理论已经很成熟很完备，也积累了庞大的公共数据库，因此短时间内基因芯片技术还会占据主导优势。但相信在不久的将来，高通量测序技术将会越来越成熟并得到更加广泛的应用。 

参考文献:



- [1]Whitfeld, P.R. A method for the determination of nucleotide sequence in polyribonucleotides. *Biochem J*, 1954, 58(3): p. 390-6.
- [2]Sanger, F., S. Nicklen, and A.R. Coulson, DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*, 1977, 74(12): p. 5463-7.
- [3]Maxam, A.M. and W. Gilbert. A new method for sequencing DNA. *Proc Natl Acad Sci U S A*, 1977, 74(2): p. 560-4.
- [4]Margulies, M., et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 2005, 437(7057): p. 376-80.
- [5]<http://www.454.com>.
- [6]Fedurco, M., et al. BTA, a novel reagent for DNA attachment on glass and efficient generation of solid-phase amplified DNA colonies. *Nucleic Acids Res*, 2006, 34(3): p. e22.
- [7]Turcatti, G., et al. A new class of cleavable fluorescent nucleotides: synthesis and optimization as reversible terminators for DNA sequencing by synthesis. *Nucleic Acids Res*, 2008, 36(4): p. e25.
- [8]<http://www.solexa.com>.
- [9]Shendure, J., et al. Accurate multiplex polony sequencing of an evolved bacterial genome. *Science*, 2005, 309(5741): p. 1728-32.
- [10]<http://www.appliedbiosystems.com.cn/>
- [11]Braslavsky, I., et al. Sequence information can be obtained from single DNA molecules. *Proc Natl Acad Sci U S A*, 2003, 100(7): p. 3960-4.
- [12]Harris, T.D., et al. Single-molecule DNA sequencing of a viral genome. *Science*, 2008, 320(5872): p. 106-9.
- [13]<http://www.helicosbio.com/>
- [14]Turner, S. Harnessing nature's powerful DNA sequencing engine: single molecule real time sequencing-by-synthesis in *Advances in Genome Biology and Technology meeting*. 2008, Marco Island, FL.
- [15]<http://www.pacificbiosciences.com>.
- [16]Clarke, J., et al. Continuous base identification for single-molecule nanopore DNA sequencing. *Nat Nanotechnol*, 2009, 4(4): p. 265-70.
- [17]Rusk, N. Cheap third-generation sequencing. *Nature*, 2009, 461(7262): p. 244-245.
- [18]Chan, E.Y. Advances in sequencing technology. *Mutat Res*, 2005, 573(1-2): p. 13-40.
- [19]Melamede, R.J. Automatable process for sequencing nucleotide. 1985, US patent.
- [20]Heiger, D.N., A.S. Cohen, and B.L. Karger. Separation of DNA restriction fragments by high performance capillary electrophoresis with low and zero crosslinked polyacrylamide using continuous and pulsed electric fields. *J Chromatogr*, 1990, 516(1): p. 33-48.
- [21]Hyman, E.D. A new method of sequencing DNA. *Anal Biochem*, 1988, 174(2): p. 423-36.
- [22]Pfeifer, G.P., et al. Genomic sequencing and methylation analysis by ligation mediated PCR. *Science*, 1989, 246(4931): p. 810-3.
- [23]Drmanac, R., et al. Sequencing of megabase plus DNA by hybridization: theory of the method. *Genomics*, 1989, 4(2): p. 114-28.
- [24]Ronaghi, M. Pyrosequencing sheds light on DNA sequencing. *Genome Res*, 2001, 11(1): p. 3-11.
- [25]Mardis, E.R. The impact of next-generation sequencing technology on genetics. *Trends Genet*, 2008, 24(3): p. 133-41.

[26]Ng, P., et al. Gene identification signature (GIS) analysis for transcriptome characterization and genome annotation. Nat Methods, 2005, 2(2): p. 105-11.

[27]Dressman, D., et al. Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. Proc Natl Acad Sci U S A, 2003, 100(15): p. 8817-22.

[28]Shendure, J. and H. Ji. Next-generation DNA sequencing. Nat Biotechnol, 2008, 26(10): p. 1135-45.

[29]evene, M. J., et al. Zero-mode waveguides for single-

molecule analysis at high concentrations. Science, 2003, 299(5607): p. 682-6.

[30]Korlach, J., et al. Selective aluminum passivation for targeted immobilization of single DNA polymerase molecules in zero-mode waveguide nanostructures. Proc Natl Acad Sci U S A, 2008, 105(4): p. 1176-81.

[31]<http://www.nanoporetech.com/>

[32]Sultan, M., et al. A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. Science, 2008, 321(5891): p. 956-60.

收稿时间：2009年5月8日

基金项目：科技部973项目课题“菜籽油积累潜势的基因调控及生物学效应”（2006CB101605）。

作者信息



孙海汐

中国科学院遗传与发育生物学研究所，硕士研究生，研究方向为生物信息学。



王秀杰

中国科学院遗传与发育生物学研究所植物基因组学国家重点实验室，博士，研究员，研究方向为生物信息学。