

DNA 测序技术概述

占爱瑶, 罗培高

四川农业大学 农学院, 四川 雅安 625014

[摘要] DNA 测序技术作为现代生命科学研究的核心技术之一, 自上世纪 70 年代中期 DNA 发明以来发展迅速。我们简要综述现有的几代 DNA 测序技术的原理及其发展历程, 并对未来可能出现的第三代测序进行预测。

[关键词] DNA 测序; 新一代测序; 单分子测序; 直接测序

[中图分类号] Q523 [文献标识码] A [文章编号] 1009-0002(2011)04-0584-05

The Overview of DNA Sequencing Technology

ZHAN Ai-Yao, LUO Pei-Gao

College of Agricultural, Sichuan Agricultural University, Ya'an 625014, China

[Abstract] DNA sequencing technology as one of the core technology of modern life science research has quickly developed since its establishing, the mid-1970s. In this article we mainly summarized the currently existing several generation DNA sequencing technologies, its evolution, and predicted the third generation sequencing may be appear in the future.

[Key words] DNA sequencing technology; next generation sequencing; single molecule sequencing; direct DNA sequencing

1944 年, Avery 通过肺炎双球菌转化实验证明 DNA 是遗传信息的载体^[1]。此后, 人们一直致力于 DNA 结构和序列研究, 使得 DNA 测序技术应运而生^[2-5]。该技术对探索生命奥秘、治疗疾病, 以及整个生物科学乃至医学的发展起到了巨大的推动作用, 并具有广阔的应用前景。

DNA 测序技术的发展经历了几个重要阶段: ①经典的手工测序, 主要依赖于一些生物化学方法和电泳分离技术, 包括 Sanger 测序和 DNA 化学降解测序; ②第一代测序技术, 主要基于 Sanger 法的测序原理, 结合荧光标记和毛细管阵列电泳技术来实现测序的自动化; ③第二代测序平台, 又称为新一代测序技术, 主要包括 Solexa 测序、Solid 平台、454 测序、HeliScope 遗传分析系统等, 它们的测序原理不完全相同, 但共同的特点是都不需要传统的克隆步骤, 实现了更高的通量。目前, 该技术仍在发生着迅速的变化, 它的发展使人们对遗传物质 DNA 的认识层次不断升华, 从对单一、局部的基因或基因片段的研究转变成了对整个基因组的研究^[6]。我们主要回顾了这几代 DNA 测序技术的发展历程, 并对未来可能出现的第三代测序进行了预测。

1 经典的 DNA 测序方法

20 世纪 70 年代中期, 2 种不同的 DNA 测序方法几乎同时发表: Sanger 等提出“DNA 双脱氧链末端终止测序”, 又称 Sanger 测序; Maxam 等提出“DNA 化学降解测序”^[4-5]。前者是利用 DNA 聚合酶的聚合反应, 在反应体系中引入一定比例的 2', 3'-双脱氧核苷三磷酸(ddNTP)作为终止剂, 由于 DNA 多聚酶不能区分 dNTP 和 ddNTP, 因此 ddNTP 可以掺入新生单链中, 而 ddNTP 的核糖基 3' 碳原子上连接的是氢原子而不是羟基, 因而不能与下一个核苷酸聚合延伸, 合成的新链在此终止, 终止点由反应中相应的双脱氧核苷酸三磷酸而定。后者的测序原理是在选定的核苷酸碱基中引入化学基团, 经过化合物处理使 DNA 分子在被修饰的核苷酸位置降解。虽然这 2 种测序方法的原理不同, 但它们都是根据核苷酸在某一固定的点起始, 随机在某一特定碱基处终止, 形成一系列以某一特定脱氧核糖核苷

[收稿日期] 2010-12-09

[基金项目] 国家自然科学基金(30971787); 霍英东高校教师基金(20070411148); 四川省杰出青年基金(2010JQ0042)

[作者简介] 占爱瑶(1984-), 女, 实验研究员, 硕士

[通信作者] 罗培高, (E-mail)lp052000@yahoo.com.cn

酸(A、T、C、G)为末端的长度各异的寡聚脱氧核糖核苷酸混合物,然后通过高分辨率的变性聚丙烯酰胺凝胶电泳,经放射自显影后,直接从胶片上读出DNA的顺序。

这2种测序方法都要依赖放射性同位素标记(^{32}P , ^{35}S)来实现,操作步骤繁琐,难以自动化,不能满足大规模测序的要求。因此,科研人员开始寻求新的非放射性标记的测序技术来克服这些缺点。

2 第一代 DNA 测序技术

2.1 第一代半自动 DNA 测序

20世纪80年代末,荧光标记的测序技术逐渐取代同位素标记测序,四色荧光标记的应用使测序反应物的分离能在一个泳道完成,从而降低了泳道间迁移率的差异对测序精度的影响,为一块平板胶做96个样品的高密度电泳的发展奠定了基础^[7]。此时,基于凝胶平板电泳的第一代半自动DNA测序仪应运而生^[8],如ABI公司首次推出的ABI 370半自动测序仪,其原理是采用4种具有不同发射波长的荧光染料标记由聚合酶链终止反应产生的一系列终止片段,经过聚丙烯酰胺凝胶电泳分离后,利用激光对分离的DNA片段携带的荧光信号进行激发,然后检测并记录不同发射波长的信号,经计算机处理,最后得到DNA序列信息^[9]。这种方法免除了同位素标记必须同时进行4组反应的麻烦,而且简化为由1条泳道同时读出4种碱基,大大提高了测序速度,为自动化加样及计算机阅读提供了技术基础,为基因组大规模测序提供了可能。但这种测序技术仍然使用平板凝胶电泳技术,费时费力,分析容量较低,提供的信息较少。

2.2 第一代全自动 DNA 测序

20世纪90年代,基因组学的出现和发展对测序的需求产生了数量级的跳跃。同时,毛细管阵列电泳DNA测序技术也开始产生^[10-13]。此时的测序仪用毛细管阵列电泳取代聚丙烯酰胺凝胶平板电泳,使样本分离可在一系列平行的石英毛细管中进行,可同时并行分析多个样本,加快了DNA的测序速度,如这一时期ABI公司开发的ABI 3730测序仪。与普通电泳相比,毛细管电泳具有许多优势,如灵敏度高、样品少、高效快速、可以在线检测等。此外,毛细管电泳设备的紧凑形式更易于实现并行化,可以获得较高的通量^[14]。这一代测序仪在人类基因组计划的后期起到了关键作用,使人类基因组计划比原计划提前2年完成^[15]。

这一代测序仪由于其原始数据的准确率高,读长长,目前仍在使用中。但是它仍然依赖于电泳分离技术,很难再进一步提升其分析速度和并行化程度,很难再降低它的测序成本^[17]。因此,需要寻求一些新的测序方法来突破这些局限。

3 第二代 DNA 测序技术

第二代测序技术又称新一代测序技术,属于循环阵列合成测序,采用大规模矩阵结构的微阵列分析技术,利用DNA聚合酶或连接酶及引物对模板进行一系列的延伸,通过显微技术观察记录连续测序循环中的光学信号来实现测序,可以同时并行分析阵列上的DNA样本^[17-19]。自2005年454焦磷酸测序被报道后,其他各类循环阵列合成测序相继出现^[20-21]。根据所分析样本在测序前是否需要扩增,大致可分为克隆扩增型和单分子测序^[22]。

3.1 克隆扩增型

克隆扩增型主要包括454焦磷酸测序、ABI公司的Solid系统、Illumina公司的Solexa测序等。尽管从模板文库制备、片段扩增到测序,这几种平台所采用的方法不一样,但它们都要经过模板文库制备、DNA片段扩增(加强测序过程中的光学检测灵敏)、并行测序、信号采集及序列拼接、组装等步骤(图1)。我们以454焦磷酸测序为例进行阐述。

454焦磷酸测序的技术核心是利用微乳液PCR技术(emulsion PCR)来实现DNA片段的扩增,利用焦磷酸法产生的光学信号进行显微观察检测,达到实时测定DNA序列的目的,是边合成边测序。

- ①模板文库制备:将待测DNA处理成<500 bp的片段并制备成单链DNA文库,加上接头^[23]。
- ②DNA片段扩增:将单链DNA文库模板及必要的PCR反应化合物与固化引物的微球(28 μm)混合,使每一个微球携带一个特定的单链DNA片段,微球结合的文库被扩增试剂乳化,这样就形成了只包含一个微球和一个特定片段的微乳滴,每个微乳滴都是一个进行后续PCR反应的微型化学反应器。微乳滴PCR最大的特点是可以形成数目庞大的独立反应空间以进行DNA扩增,同时避免了引物间及与模板DNA相互干扰的问题^[24]。整个片段文库平行扩增,多个热循环后,每个微球表面都结合了成百数千个相同的DNA拷贝,然后乳液混合物被打破,富集微球。
- ③并行测序:将富集的微球转移到刻有规则微孔阵列的微孔板上,微孔板一端用于测序反应的化合物通过,另一端与CCD光学检测系统的光纤部件

接触,用于信号检测^[25]。每个微孔只能容纳一个微球,将微孔板放置在 GS FLX 中,测序反应开始。引物与模板 DNA 退火后,在 DNA 聚合酶、三磷酸腺苷硫酸化酶、萤光素酶和三磷酸腺苷双磷酸酶等 4 种酶的协同作用下完成循环测序反应。放置在 4 个单独的试剂瓶里的 4 种碱基,依照 T、A、C、G 的顺序依次循环进入微孔板,每次只进一个碱基,当加入的 dNTP 与模板互补时,DNA 模板与互补的 dNTP 聚合时可以产生等摩尔 PPi,在三磷酸腺苷硫酸化酶的催化下,PPi 与 5'-磷酸化硫酸腺苷(APS)反应生成等量的 ATP;在萤光素酶的催化下,生成的 ATP 又可以和萤光素结合形成氧化萤光素,同时释放光信号。最后为信号采集及序列组装、拼接。在 454 焦磷酸测序中需要注意 2 点:①用 dATP 的类似物 dATP α S 代替 dATP,因为 dATP 的结构与 ATP 相似,能与萤光素反应产生背景荧光,而 dATP α S 几乎不产生背景荧光;②用三磷酸腺苷双磷酸酶使测序能循环进行并得到信号峰,三磷酸腺苷双磷酸酶能在聚合反应完成后降解剩余的 dNTP 和产生的 ATP,因此不须洗涤或分离步骤来除去剩余的 dNTP,同时也解决了测序反应中产生的 ATP 信号

会使背景累积而溢出而使测序无法进行的问题^[26]。

与 454 的情况相同,Solid 系统也采用了微乳滴 PCR 与微球相结合的策略来扩增 DNA 模板,不过 Solid 系统的微球要小得多,只有 1 μm ^[27-29]。它独特之处在于其边合成边测序采用的是连接反应而不是聚合反应,同时,它运用双碱基编码策略来协助检测错误^[30]。Solexa 测序系统则是利用其专利核心技术“DNA 簇”和“可逆性末端终结”,实现自动化样本制备及基因组数百万个碱基大规模的平行测序,其单链文库片段的扩增是通过所谓“桥式扩增”过程实现的^[31-33]。

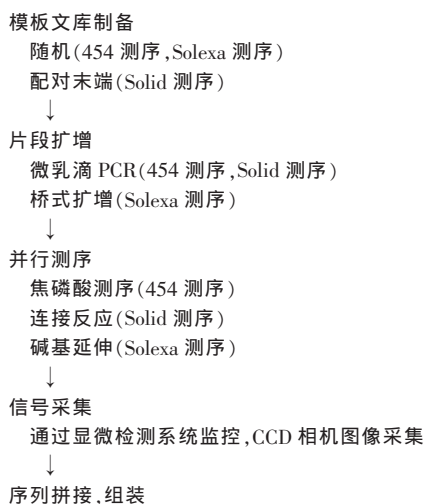
如果说第一代测序技术的特点是“线性的”,仪器每次运行只能读取有限的几条“通道”上的数据;那么第二代克隆扩增型测序则是“平面的”,仪器每运行一次可以读取整个平面上成千上万个测定点的信号,即它具有高度的并行性、高通量、操作简单、成本低等优势。但克隆扩增型测序也存在致命的缺陷,即读长很短,甚至还不如传统的 Sanger 法,这就给后期的序列组装带来了巨大的压力。

3.2 单分子测序

为了克服克隆扩增型读长短的问题,单分子测序应运而生。它也是一种合成测序技术,能够直接观察和测定 DNA 或 RNA 分子的数量和序列的结构^[34-35],主要包括 HeliScope 遗传分析系统、VisiGen 公司的单分子合成测序仪及 Pacific Biosciences 公司的单分子实时测序技术等。与克隆扩增型所不同的是,它不需要模板的预先扩增,通过在单一 DNA 分子组成的阵列上进行合成测序,通过物理等各种技术手段来观察和记录 DNA 聚合酶合成 DNA 的过程,如 Helicos Biosciences 公司的全反射显微镜技术、VisiGen 生物技术公司的蛋白质纳米装置、Pacific Biosciences 公司的零级波导纳米结构。

就测序技术本身而言,单分子测序和克隆扩增型的区别并不大,但测序结果却有不小的差别^[22]:①单分子测序不需要 PCR 扩增,更能反映细胞或组织内分子的真实情况,特别是在需要定量分析的情况下^[36];②该技术具有更高的通量,因为在一个特定的表面,单个分子可以增加独立分析的 DNA 片段数量^[21];③该项技术有望综合连续的长的读长和无以伦比的准确性。尽管如此,单分子测序也遇到了新的问题:如何降低非检测特异性背景的干扰,如何更准确快速地记录测序反应的结果,如何自动快速处理巨量的 DNA 序列信息等^[37-38]。这些都是亟待人们进一步研究解决的问题。

克隆扩增型:



单分子测序型:

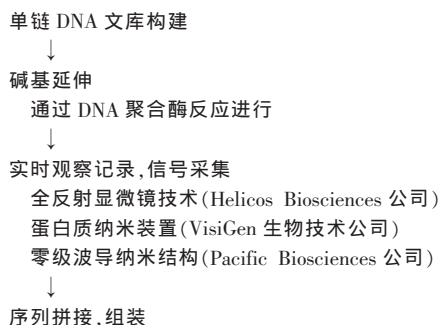


图 1 新一代测序流程图

4 第三代测序技术——直接测序技术

近年来人们一直致力于直接测序技术的相关研究。直接测序技术,即通过现代光学、高分子、纳米技术等手段来区分碱基信号差异的原理,以达到直接读取序列信息的目的,而不需要使用生物或化学试剂,这对于进一步降低测序成本是非常可取的。它有以下几种类型:①非光学显微镜成像:就是将核苷酸(主要是碱基)的空间线性排列方式可视化,如果一个 DNA 链的图片具有足够高的分辨率,可以把 DNA 链上的 4 种碱基区分开来,那么 DNA 序列将非常容易被读出,这一设想是想借助具有原子水平分辨率的非光学显微镜来实现,如扫描隧道显微镜(STM)、原子力显微镜(AFM)等^[39-40]。日本大阪大学与 ZS Genetics 公司都在致力于这方面的研究,并且有一定的进展^[41]。②石墨烯和碳纳米管:石墨烯是由碳原子组成的二维晶体,碳原子排列与石墨的单原子层一样,非常稳定并且具有良好的导电性,适合制作核酸测序用电极。这种设想是在石墨烯上打出约 1 nm 宽的缝隙,引导 DNA 分子垂直通过缝隙,当 DNA 通过时,缝隙两边的石墨烯边缘可以作为电极来直接确定 DNA 的序列^[42]。目前,IBM 公司的 Almaden Research Center 及哈佛大学正在致力于这项研究^[43-45]。③纳米孔:即直径在纳米尺度(1~2 nm)的小孔,通常利用生物分子或固态物质制成。设想在电场驱动下,当线性 DNA 分子通过小孔时,利用电流的变化来直接读取 DNA 的碱基序列。目前,许多研究机构正在开发纳米孔测序技术,如美国波斯顿大学、阿肯色大学、加利福尼亚大学、IBM 公司等^[46-48]。以上几种设想都还停留在理论验证阶段,还存在很多技术瓶颈需要解决,还无法达到对单个碱基的准确辨别。相信经过科学家们的不懈努力,直接测序技术将作为第三代 DNA 测序技术应用分子生物学的各类研究。

5 前景与展望

当今的科技革命不再是传统意义上的以某项重大科技突破为标志和某个领域的突起为代表,而是以众多学科、领域全面发展,交叉融合为特征。纵观 DNA 测序技术的发展历程,会发现 DNA 测序技术不再单纯使用生物化学或化学方法,而是慢慢融合了物理手段,并结合计算机应用,大大提高了测序速度。随着现代光学、纳米技术、计算机科学等多学科的全面发展,人们将克服困难,提高显微镜的

分辨率、增强纳米孔的灵敏度、实现石墨烯上 1 nm 缝隙,最终达到利用显微镜技术、纳米孔或石墨烯来直接读取 DNA 序列的目的。届时,DNA 测序技术将不会存在测序缺口的问题,测序速度和精度将得到更大的提升,测序成本将被人们所接受。随着测序技术的不断完善,生命科学将进入一个空前繁荣的新时期,并进而极大地影响人类文明的进程。

参考文献

- [1] Avery O T, MacLeod C M, McCarty M. Studies on the chemical nature of the substance inducing transformation of pneumococcal types[J]. J Exp Med, 1944,98:451-460.
- [2] Meselson M, Stahl F W. The replication of DNA in E.coli[J]. Proc Natl Acad Sci USA, 1958,44:671-682.
- [3] Sanger F, Coulson A R. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase[J]. J Mol Biol, 1975,94(3):441-446.
- [4] Sanger F, Nicklen S, Coulson A R. DNA sequencing with chain-terminating inhibitors[J]. Proc Natl Acad Sci USA, 1977, 74(12):5463-5467.
- [5] Maxam A M, Gilbert W. A new method for sequencing DNA [J]. Proc Natl Acad Sci USA, 1977,74(2):560-564.
- [6] 周晓光, 任鲁凤, 李运涛, 等. 下一代测序技术: 技术回顾与展望[J]. 中国科学: 生命科学, 2010,40(1):23-37.
- [7] 甄志成, 姚志建. 毛细管阵列电泳与规模化 DNA 测序[J]. 色谱, 2001,19(4):361-364.
- [8] Prober J M, Trainor G L, Dam R J, et al. A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides[J]. Science, 1987,238(4825):336-341.
- [9] 马洪明, 路新枝, 王勇. 四色荧光标记 DNA 序中一例典型错读码的校正[J]. 海洋科学, 2002,26(3):10-12.
- [10] Swerdlow H, Wu S, Harke H, et al. Capillary gel electrophoresis for DNA sequencing: Laser-induced fluorescence detection with the sheath flow cuvette[J]. J Chromatogr, 1990,516(1):61-67.
- [11] Chen D Y, Swerdlow H P, Harke H R, et al. Low-cost, high-sensitivity laser-induced fluorescence detection for DNA sequencing by capillary gel electrophoresis [J]. J Chromatogr, 1991,551(1-2):237-246.
- [12] Ruiz-Martinez M C, Berka J, Belenkii A, et al. DNA sequencing by capillary electrophoresis with replaceable linear polyacrylamide and laser-induced fluorescence detection [J]. Anal Chem, 1993,65(20):2851-2858.
- [13] Dovichi N J. DNA sequencing by capillary electrophoresis[J]. Electrophoresis, 1997,18(12-13):2393-2399.
- [14] Hunkapiller T, Kaiser R J, Koop B F, et al. Large-scale and automated DNA sequence determination[J]. Science, 1991,254 (5028):59-67.
- [15] Zubrisky E. How analytical chemists saved the human genome project[J]. Anal Chem, 2002,74(1):23-26.
- [16] Metzker M L. Sequencing technologies--the next generation[J].

- Nat Rev Genet, 2010,11(1):31–46.
- [17] Mitra R D, Church G M. In situ localized amplification and contact replication of many individual DNA molecules[J]. Nucleic Acids Res, 1999,27(24):34–39.
- [18] Mitra R D, Shendure J, Olejnik J, et al. Fluorescent in situ sequencing on polymerase colonies[J]. Anal Biochem, 2003,320(1):55–65.
- [19] Shendure J, Porreca G J, Reppas N B, et al. Accurate multiplex polony sequencing of an evolved bacterial genome[J]. Science, 2005,309(5471):1728–1732.
- [20] Bentley D R. Whole-genome re-sequencing[J]. Curr Opin Genet Dev, 2006,16(6):545–552.
- [21] Harris T D, Buzby P R, Babcock H, et al. Single-molecule DNA sequencing of a viral genome[J]. Science, 2008,320(5872):106–109.
- [22] 沈树泉. 单分子测序与个体医学[J]. 生理科学进展, 2009,40(3):283–288.
- [23] 王传超, 李辉. 古 DNA 分析技术发展的三次革命[J]. 现代人类学通讯, 2010,16(4):35–42.
- [24] 晏菱, 葛芹玉, 蒋小青, 等. 微乳液多重 PCR 基因芯片法在早孕期无创性胎儿性别诊断中的应用[J]. 现代医学, 2007,35(2):83–88.
- [25] 张晓丹, 武海萍, 周国华. 焦测序技术及其在遗传分析中的应用[J]. 分析化学, 2006,34(6):582–586.
- [26] Dressman D, Yan H, Traverso G, et al. Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations [J]. Proc Natl Acad Sci USA, 2003,100(15):8817–8822.
- [27] Mardis E R. Next-Generation DNA sequencing methods[J]. Annu Rev Genomics Hum Genet, 2008,9:387–402.
- [28] Barbee K D, Huang X H. Magnetic assembly of high-density DNA arrays for genomic analyses[J]. Anal Chem, 2008,80(6):2149–2154.
- [29] Housby J N, Southern E M. Fidelity of DNA ligation: a novel experimental approach based on the polymerisation of libraries of oligonucleotides[J]. Nucleic Acids Res, 1998,26(18):4259–4266.
- [30] Fedurco M, Romieu A, Williams S, et al. BTA, a novel reagent for DNA attachment on glass and efficient generation of solid-phase amplified DNA colonies[J]. Nucleic Acids Res, 2006,34(3):e22.
- [31] Turcatti G, Romieu A, Fedurco M, et al. A new class of cleavable fluorescent nucleotides: synthesis and optimization as reversible terminators for DNA sequencing by synthesis[J]. Nucleic Acids Res, 2008,36(4):e25.
- [32] Ju J, Kim D H, Bi L, et al. Four-color DNA sequencing by synthesis using cleavable fluorescent nucleotide reversible terminators [J]. Proc Natl Acad Sci USA, 2006,103(52):19635–19640.
- [33] Gupta P K. Single-molecule DNA sequencing technologies for future genomics research [J]. Trends Biotechnol, 2008,26(11):602–611.
- [34] Ozsolak F, Platt A R, Jones D R, et al. Direct RNA sequencing[J]. Nature, 2009,461(7265):814–818.
- [35] Acinas S G, Sarma-Rupavtarm R, Klepac-Ceraj V, et al. PCR-induced sequence artifacts and Bias: insights from comparison of two 16S rRNA clone libraries constructed from the same sample[J]. Appl Environ Microbiol, 2005,71(12):8966–8969.
- [36] Braslavsky I, Hebert B, Kartalov E, et al. Sequence information can be obtained from single DNA molecules[J]. Proc Natl Acad Sci USA, 2003,100(7):3960–3964.
- [37] Korlach J, Marks P J, Cicero R L, et al. Selective aluminum passivation for targeted immobilization of single DNA polymerase molecules in zero-mode waveguide nanostructures [J]. Proc Natl Acad Sci USA, 2008,105(4):1176–1181.
- [38] Driscoll R J, Youngquist M G, Baldeschwieler J D. Atomic-scale imaging of DNA using scanning tunnelling microscopy[J]. Nature, 1990,346(6281):294–296.
- [39] Ikai A. STM and AFM of biorganic molecules and structures [J]. Sur Sci Rep, 1990,26(8):261–332.
- [40] Tanaka H, Kawai T. Partial sequencing of a single DNA molecule with a scanning tunnelling microscope[J]. Nat Nanotechnol, 2009,4(8):518–522.
- [41] Postma H W. Rapid sequencing of individual DNA molecules in graphene nanogaps[J]. Nano Lett, 2010,10(2):420–425.
- [42] Gigliotti B, Sakizze B, Bethune D S, et al. Sequence-independent helical wrapping of single-walled carbon nanotubes by long genomic DNA[J]. Nano Lett, 2006,6(2):159–164.
- [43] Meng S, Maragakis P, Papaloukas C, et al. DNA nucleoside interaction and identification with carbon nanotubes [J]. Nano Lett, 2007,7(1):45–50.
- [44] Albertorio F, Hughes M E, Golovchenko J A, et al. Base dependent DNA-carbon nanotube interactions: activation enthalpies and assembly-disassembly control [J]. Nanotechnology, 2009,20(39):395101–395109.
- [45] Meller A, Nivon L, Branton D. Voltage-driven DNA translocations through a nanopore[J]. Phys Rev Lett, 2001,86(15):3435–3438.
- [46] Fologea D, Gershow M, Ledden B, et al. Detecting single stranded DNA with a solid state nanopore[J]. Nano Lett, 2005, 5(10):1905–1909.
- [47] Lagerqvist J, Zwolak M, Di Ventra M. Fast DNA sequencing via transverse electronic transport[J]. Nano Lett, 2006,6(4):779–782.