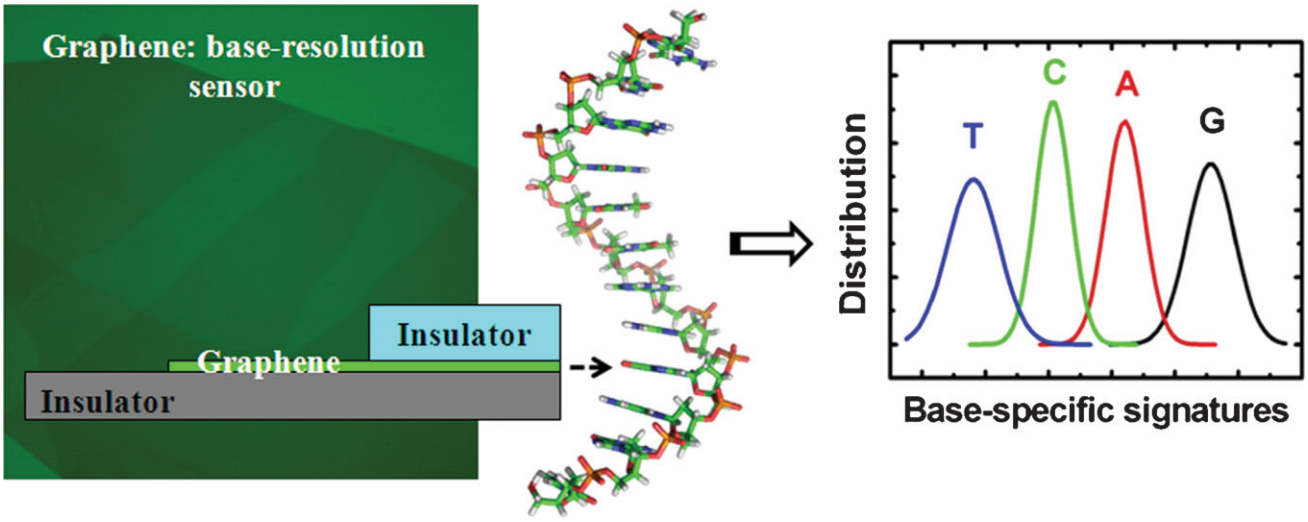
Perspectives and Challenges of Emerging Single-Molecule DNA Sequencing Technologies

Mingsheng Xu, Daisuke Fujita, and Nobutaka Hanagata*



From the Contents

The growing demand for analysis of the genomes of many species and cancers, for understanding the role of genetic variation among individuals in disease, and with the ultimate goal of deciphering individual human genomes has led to the development of non-Sanger reaction-based technologies towards rapid and inexpensive DNA sequencing. Recent advancements in new DNA sequencing technologies are changing the scientific horizon by dramatically accelerating biological and biomedical research and promising an era of personalized medicine for improved human health. Two single-molecule sequencing technologies based on fluorescence detection are already in a working state. The newly launched and emerging single-molecule DNA sequencing approaches are reviewed here. The current challenges of these technologies and potential methods of overcoming these challenges are critically discussed. Further research and development of single-molecule sequencing will allow researchers to gather nearly error-free genomic data in a timely and inexpensive manner.

small NANO MICRO

Table 1. New commercialized DNA sequencing platforms.

Platform	Read-length (bp)	Accuracy	Bases per run	Cost/human genome [US$]	Price of instrument [US$]
Massively parallel pyrosequencing by synthesis (Roche/454: GS FLX Titanium Series)[a]	400–500	Q20 read-length of 400 bases (99% at 400 bases and higher for prior bases)	400–600 million bases per 10 h run	1 million	500000
Sequencing by synthesis (Illumina/Solexa: Genome Analyzer *IIx*)[b]	2 × 75	Base call with Q30 (>70%)	14–18 Gigabases per 9.5-day run	60000	450000
Bead-based massively parallel clonal ligation based sequencing (Applied Biosystems/ Agencout: SOLiD 3 system)[c]	100	99.94%	20 Gigabases	60000	591000
Massively parallel single-molecule sequencing by synthesis (Helicos/Stanford Univ.)[d]	30–35	99.995% at >20 × coverage (raw error rate: ≤ 5%)	21–28 Gigabases per 8-day run	70000	1.35 million
Single molecule, real time sequencing by synthesis (Pacific BioSciences/Cornell Univ.)[e]	1000–1500[f]	99.3% at 15 × coverage (error of a single read: 15–20%)	–		

[a] http://www.454.com/, second-generation or next-generation sequencing technology. [b] http://www.illumina.com/, second-generation or next-generation sequencing technology. [c] http://www3.appliedbiosystems.com/AB_Home/index.htm, second generation or next generation sequencing technology. [d] http://www.helicosbio.com/, third generation or next-next generation sequencing technology. [e] http://www.pacificbiosciences.com/, third generation or next-next generation sequencing technology. [f] The 1000-1500 bases read length was obtained for repeated sequencing of small circular DNA.

1. Introduction

With the completion of the draft sequence[1,2] as well as subsequent finishing[3] of the human genome, genome-based medicine has come closer to reality. Despite this, many agencies, such as the US National Institute of Health/ National Human Genome Research Institute, and companies continue to sponsor many research projects aimed at developing new sequencing technologies to enable large-scale sequencing of complex genomes with low cost: the ''$1000 genome''.[4] Low-cost and faster ways to sequence DNA would revolutionize the use of genetic information and the nature of biological and biomedical research.

Sanger technology has been the workhorse of the DNA sequencing industry for some 30 years. Today, Sanger sequencing is a reliable, highly accurate method for sequencing, capable of reading sequences over 1000 base pairs (bp) in length and up to as high as 99.999% base calls. The technology initiated the process of deciphering genes, and eventually entire genomes, and the overwhelming majority of DNA sequence production to date has relied on the Sanger technology. The rapidly growing demand for throughput, with the ultimate goal of deciphering individual human genomes, has led to substantial improvements in the technique, as exemplified in automated capillary electrophoresis. To further reduce costs for this dideoxy sequencing, significant progress in miniaturization of the method and/or integration of a series of sequencing-related steps in a ''lab-on-chip'' format has been achieved.[5] However, the Sanger sequencing method is still far too slow and costly for reading personal genetic codes. It costs an estimated US$ 10–25 million to sequence a single human genome and $20000–$50000 to sequence a microbial genome.[6] The growing demand for analysis of the genomes of many species and cancers, for understanding the role of genetic variation among individuals in disease, and with the ultimate goal of deciphering individual human genomes has led to developing new technologies for DNA sequencing (Table 1), including 454 sequencing,[7] Solexa technology, and the SOLiD platform. The recently commercialized technologies, generally described as ''second generation'' or ''next generation'' sequencing systems,[8] have hugely reduced the cost of sequencing and have simultaneously yielded an increase in DNA sequencing speed. However, these technologies generate different base read-lengths, different error rates, and different errors profiles relative to traditional Sanger-based data and to each other. The limited read-length has substantial impact on certain applications such as de novo genome sequencing and assembly for whole-genome sequencing – a process that is sensitive to read length, base accuracy, and error type, and thus new methods are needed to analyze and integrate the massive data sets and then apply those results to various types of biological information.[9]

[*] Dr. M. S. Xu
International Center for Young Scientists
National Institute for Materials Science
1-2-1 Sengen, Tsukuba, Ibaraki 305-0047 (Japan)
E-mail: XU.Mingsheng@nims.go.jp
Prof. D. Fujita
International Center for Young Scientists
Advanced Nano Characterization Center
International Center for Materials Nanoarchitectonics
National Institute for Materials Science
1-2-1 Sengen, Tsukuba, Ibaraki 305-0047 (Japan)
Prof. N. Hanagata
Nanotechnology Innovation Center
National Institute for Materials Science
1-2-1 Sengen, Tsukuba, Ibaraki 305-0047 (Japan)

DOI: 10.1002/smll.200900976

To realize the prospect of medical care tailored to an individual's unique genetic identity, a revolutionary sequencing method is needed for the cost-effective and rapid interrogation of individual genomes. The sequencing of individual DNA molecules[10–12] has such potentials as those mentioned above. Cyclic-array approaches using fluorescent base incorporation with polymerases also has been adapted for single-molecule sequencing of molecules immobilized on surfaces.[13,14] In another approach, a processive exonuclease is used to digest a single DNA molecule and the released nucleotides are identified one at a time.[15] In one manifestation, fluorescent nucleotides are incorporated into the DNA and detected by single-molecule fluorescence upon release.[16,17]

This Review discusses several single-molecule DNA sequencing technologies, two of which have been launched recently and others that are under development by different companies. Single-molecule DNA sequencing represents an emerging wave of sequencing technologies. We focus on the basic principles and challenges as well as insights for overcoming the challenges in developing the emerging single-molecule technologies towards affordable and portable electronic DNA sequencing. We emphasize the rapid development of new DNA sequencing technologies and present some reviews on second-generation DNA sequencing,[8–11,18,19] their applications and impact, and the bioinformatics challenges of the sequencing technologies.[8,20,21] This Review is not intended to cover all relevant interesting and important works but instead illustrate the major challenges and advantages common to most of these approaches.

Single-molecule DNA sequencing is a method used to determine the base sequences along individual DNA strands by detecting base-specific features of the four bases, including size, and optical (fluorescence), electrical, and magnetic properties. Based on detection techniques, single-molecule sequencing can be classified as fluorescence sequencing (Helicos and Pacific Biosciences), direct sequencing by using transmission electron microscopy (TEM), and electronic sequencing by using nanopore, scanning tunneling microscopy (STM), and nano-knife-edge probes. The direct and electronic sequencing are physical approaches[22,23] to detecting DNA at the single-base level, in striking contrast to currently available sequencing technologies. Success of single-molecule sequencing relies heavily not only on detection techniques but also on the preparation of DNA samples, which is evident in the platform of Pacific Biosciences.

Mingsheng Xu received his Ph.D in electronic engineering from The Chinese University of Hong Kong. He joined the Nanoelectronics Collaborative Research Center at The University of Tokyo in 2003 and remained there until 2007. He also joined the Venture Business Laboratory at Chiba University (Japan) in 2005. Since 2007 he has worked at the National Institute for Materials Science (NIMS), Japan. He is currently a researcher of the International Center for Young Scientists, NIMS. His research interests include scanning probe microscopy, π-electron materials and opto-electronic devices, and nanobiological interactions.

Daisuke Fujita is the managing director of Advanced Nano Characterization Center of National Institute for Materials Science (NIMS), and the principal investigator of the Center for Materials Nanoarchitectonics (WPI-MANA). His research field is nanometer-scale materials science on metals, semiconductors, superconductors, nanoclusters, and molecules in extreme environments such as low or high temperature, high magnetic field, ultrahigh vacuum, stress–strain field, and so on. Novel quantum phenomena and nanofunctionality of nanostructures observable at extreme fields are his main research targets. Development of novel nanostructured materials such as size-controlled nanoclusters, nanowires, nanotubes, and nanosheets is also among his research interests.

Nobutaka Hanagata joined National Institute for Materials Science (NIMS) as a senior researcher in 2005, and he currently supervises the Nanotechnology Innovation Center in NIMS, an open facility for interdisciplinary studies including nanobiology. He also joined the Graduate School of Life Science at Hokkaido University as a professor in 2008. Recently, he has been working in the interdisciplinary area between biology and materials science, particularly on the effect of surface structure of materials on cellular function. His other interests include discovery genomics using DNA microarrays, expressed sequence tags, differential displays, and LC-MS techniques to find novel functional genes or proteins.

2. True Single-Molecule Sequencing (Helicos BioSciences)

Helicos BioSciences[24] demonstrated the first single-molecule sequencing of the M13 viral genome, which is approximately 1 million times smaller than the human genome, using Helicos's proprietary true single-molecule sequencing (tSMS) technology. The method monitors the synthesis of a single strand of DNA using a highly sensitive fluorescence detection system via sequencing by synthesis. The technology is based on work by Quake et al.[13] and relies on cyclic interrogation of a dense array of sequencing features, with no requirement of clonal amplification. A DNA polymerase sequentially adds labeled nucleotides to surface-immobilized primer-template duplexes in a stepwise fashion, and the asynchronous growth of individual DNA molecules is monitored by fluorescence imaging. A brief description of the method is illustrated in Figure 1.

To validate the tSMS technology, Harris et al.[24] re-sequenced the M13 virus genome, examining more than 280000 strands of captured DNA and directly visualizing the sequential incorporation of individual labeled nucleotides. The accuracy of the consensus sequence was 100%. To assess the accuracy and robustness of mutation detection, they introduced single-nucleotide changes into the reference M13 virus genome sequence. The tSMS technology correctly found 98% of 500 simulated mutations with zero false-positive errors.

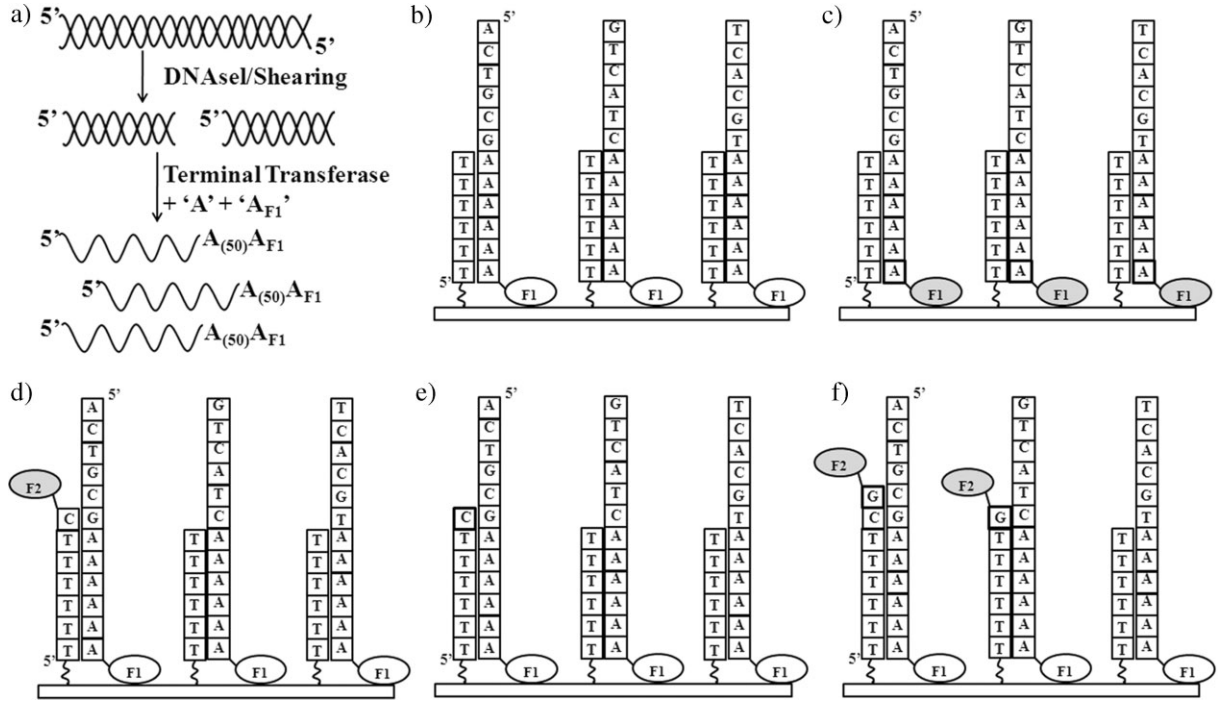


Figure 1. The first single-molecule sequencing technology (tSMS, Helicos), including sample synthesis, wash, imaging, and cleavage steps. a) Genomic DNA is prepared for sequencing by fragmentation and 3′ poly(A) tail addition, labeling, and blocking by terminal transferase. b) Hybridization capture of these templates onto a surface with covalently bound 5′ "down" dT(50) oligonucleotide. c) Imaging of the captured templates to establish sites for sequencing by synthesis. d) Incubation of this surface with one fluorescently labeled nucleotide and polymerase mixture, followed by rinsing of the synthesis mixture and direct imaging of the Cy5 labels exciting at 647 nm. e) Chemical cleavage of the dye–nucleotide linker to release the dye label. f) Addition of the next nucleotide and polymerase mixture. F1 and F2 represent cyanine-3 temperate labels and cyanine-5 nucleotide labels, respectively. Reproduced with permission from Reference [24]. Copyright 2008, AAAS.

The tSMS technology avoids polymerase chain reaction (PCR) amplification, which reduces the error rate and dramatically lowers the cost of individual genomic and genetic analyses. The tSMS approach performs a "two-pass" sequencing process (i.e., sequencing each individual template twice, giving two reads from the same position on the same strand. In the first pass, a template is sequenced as usual (pass 1); the primers are then melted off and the same template is sequenced a second time (pass 2), substantially improving raw sequencing accuracy, and the technology can easily be scaled and replicated to meet the requirements of large and complex experiments, overcoming the workflow bottlenecks encountered by methods such as PCR or cloning. Depending on reversible terminators, tSMS technology may be limited to the analysis of short DNA fragments; the present average read-length of the technology is around 30–35 bases. It is a challenge to detect base repeats and homopolymers, such as a string of consecutive cytosine bases.

3. Single-Molecule Real-Time Sequencing (Pacific Biosciences)

With the aim of uninterrupted sequencing, Pacific Biosciences is developing a transformative single-molecule real-time (SMRT) DNA sequencing technology based on the natural DNA synthesis by a DNA polymerase with phospholink nucleotides as it occurs in a continuous, processive manner (Figure 2). This approach is enabled by two key innovations: nanophotonic visualization chambers (SMRT chips) and phospholinked nucleotides. Each SMRT chip contains thousands of zero-mode waveguides (ZMWs).[25] Arrays of ZMWs are manufactured cylindrical holes with a 30 to 70-nm diameter in approximately 100-nm-thick aluminum film deposited on a SiO_2 substrate,[14,26] providing a detection volume of about 20 zeptoliters, which may localize the excitation light to dimensions unobtainable by conventional techniques such as confocal or total internal reflection microscopy.[27] The ZMW provides a window for watching DNA polymerase as it performs sequencing by synthesis.

In contrast to most sequencing-by-synthesis methods using nucleotides with fluorophores attached directly to the bases, the phosholinked nucleotides of SMRT sequencing carry their fluorescent label on the terminal phosphate[28,29] rather than the base, and each of the four nucleotides, A, C, G, and T, is labeled with a different colored fluorophore. As a natural step in the synthesis process, the phosphate chain is cleaved when the nucleotide is incorporated into the DNA strand. Thus, upon incorporation of a phospholinked nucleotide, the DNA polymerase naturally cleaves the dye molecule from the nucleotide when it cleaves the phosphate chain. The phosphate-chain–dye complex quickly diffuses the short distance out of the detection volume, ensuring the background signal remains at a low level. Phospholinked nucleotides enable the

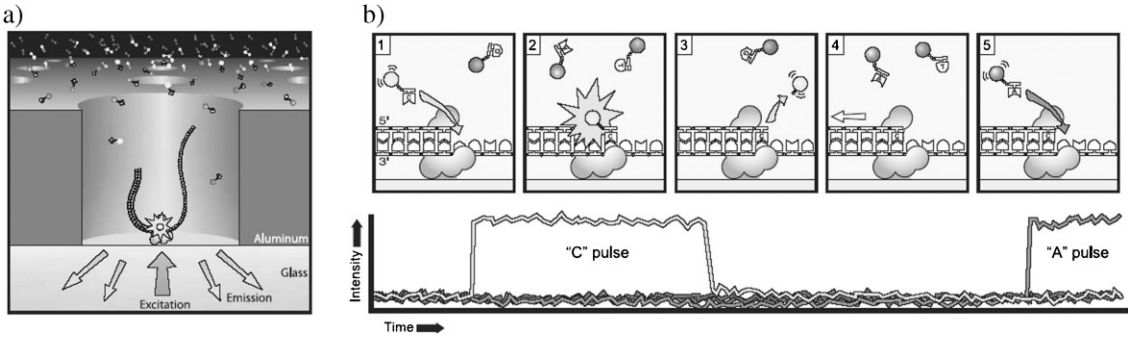


Figure 2. Principle of single-molecule, real-time DNA sequencing (SMRT, Pacific Bioscience). a) Experimental geometry. b) Schematic event sequence of the phospholinked deoxyribonucleoside triphosphate (dNTP) incorporation cycle, with a corresponding expected time trace of detected fluorescence intensity from the ZMW. 1) A phospholinked nucleotide forms a cognate association with the template in the polymerase active site, 2) causing an elevation of the fluorescence output on the corresponding color channel. 3) Phosphodiester bond formation liberates the dye-linker-pyrophosphate product, which diffuses out of the ZMW, thus ending the fluorescence pulse. 4) The polymerase translocates to the next position, and 5) the next cognate nucleotide binds the active site, beginning the subsequent pulse. Reproduced with permission from Reference [25]. Copyright 2009, AAAS.

polymerase to synthesize DNA in a fast and processive manner. When the DNA polymerase encounters the nucleotide complementary to the next base in the template, it is incorporated into the growing DNA chain. The polymerase advances to the next base and the process continues to repeat. With the use of phospholinked nucleotides, a long, natural strand of DNA is produced.

Regarding the fluorescence label, VisiGen Biotechnologies uses a similar nucleotide modification but base identification is by detecting the fluorescence resonance energy transfer (FRET) between a fluorescently tagged polymerase and a nucleotide labeled with a unique acceptor fluorescent moiety tagging the terminal phosphate of a nucleotide.[30]

The SMRT chip enables observation of individual fluorophores against a dense background of labeled nucleotides by maintaining a high signal-to-noise ratio, and it enables parallel and simultaneous sequencing with thousands of ZMWs. The technology uses a holographic optical element, a confocal pinhole array, and a dispersive optical component to acquire spectroscopic information from each of the sample locations in real time. The impressive demonstrated sequencing rate for 740 single-molecule reads was about 5 bases s^{-1}. The read accuracy of the 150-base linear DNA strand template (158 total bases in the alignment) was reported to be about 83% but improved to more than 99% by sequencing the same template molecule 15 times,[25] which would increase the time and cost of sequencing. The sequencing errors were mainly attributed to very short interphase intervals, dissociation of the complementary nucleotide before phosphodiester-bond formation, and spectral misassignment of fluorescent dyes exhibiting significant emission overlap. Such an accuracy level can produce alignment and consensus adequate for resequencing application. However, challenges would include de novo assembly or alignment into highly repetitive DNA. Performing circular consensus sequencing several times, i.e., sequencing a circularized template, could improve accuracy and read-length. As an example, the company sequenced a small synthetic DNA circle more than 10 times using a strand-displacing polymerase and reached a total read length of 1500 bases. The long reads will be a great advantage for certain applications, for example, de novo sequencing.

4. Electron Microscopy for Sequencing (ZS Genetics, Inc.)

ZS Genetics is developing a TEM platform to directly image modified individual DNA bases.[31] The technology involves a PCR to label nucleotides with unique atoms or molecules, attaching a sample to a substrate, and using TEM to identify the labeled nucleotides. Because normal DNA has only light elements such as carbon, nitrogen, oxygen, hydrogen, and phosphorus, natural DNA is essentially invisible to TEM analysis. Thus, in order to sequence DNA by TEM, DNA nucleotides need modification. ZS Genetics uses heavy elements to label nucleotides in DNA molecules. The labels are selected for easier detection and identification than that of nucleotides, including elements with an atomic number greater than 55 such as bromine and iodine, and thus create contrast. Essentially, it needs four kinds of unique label to correspondingly represent the four DNA bases. After the PCR-based modification, pieces of the DNA samples are attached to a substrate and stretched using fluid flow (e.g., molecular combing[32]). The substrate is generally thin to enable sufficient particle beam transmission analysis. By distinguishing the size and intensity of the atomic labels with nucleotides, TEM can, thus, directly sequence individual DNA molecules.

This method involves one PCR step to label but not amplify the DNA bases. This step might induce atomic labeling errors, and bias and errors resulting from PCR even if conducted once not 30 times as many technologies used for amplification on a small sample to get a strong enough signal. Attaching and straightening DNA pieces to a substrate for TEM observation may be one technical challenge since sample preparation is critical to a TEM study. The imaging cycle may be limited by the speed of the digital camera, currently about 1.5 s $shot^{-1}$. This technology has a high capital cost. An electron microscope with sufficient resolution in the subangstrom

NANO MICRO small

regime costs on the order of US$1 million. However, it is estimated that sequencing an entire human genome will cost $5000–$10000 in consumables and labor.

5. Nanopores for Sequencing

Electronic DNA sequencing has attracted a lot of interest. The prerequisite for electronic sequencing of a whole genome is the identification of robust electronic signatures of the four bases one by one, in particular in the context of single-stranded DNA (ssDNA). Because the bases have similar chemical structure, the difference in electrical signals among them may not be large.[22] To detect small signals, approaches are needed to reduce interference from the measurement environment, and/or to amplify the signals by modifying nucleotides or employing an optical, electric, or magnetic stimulus.

The underlying principle of nanopore sequencing is that an ssDNA molecule is electrophoretically driven through a nanoscale pore (1.5–5 nm) [22,23,33–35] in such a way that the molecule passes through the pore in strict linear sequence. A change in electrical signals, such as ionic current blockages, transverse tunneling currents, or capacitance, is recorded to discriminate DNA sequences. The proof of the concept for nanopore sequencing was first demonstrated by Deamer et al.[35] in 1996. Since this demonstration, the prospect of a rapid and inexpensive direct physical approach to massive sequencing capacity has stimulated nanopore research using either protein pores or solid-state nanopores. However, despite the enormous research activity, substantial lengths of DNA have yet to be sequenced with a nanopore. In this section, we consider the major challenges and advantages of nanopore sequencing. We don't intentionally separate protein pore from solid-state nanopores, though they have different characteristics. For instance, solid-state nanopores show superior chemical, thermal, and mechanical stability, tunability over biological counterpart, and have the potential of integration into devices, whereas biological pores exhibit lower noise level than solid-state nanopores.

For most nanopore-based electrical sequencing approaches,[22,23] that is, measurement of ionic current blockades,[35–37] transverse tunneling currents,[38,39] or capacitance,[40,41] the control of DNA velocity and orientation during translocation through the nanopore is a major challenge, as well as possible nanoparticle or nanobubble trapping (which might clog the pore and introduce noise) in the pore and a great deal of noise in the electrical signal. Precise control of the DNA movement through the nanopore is an essential feature of the sequencing system. Typically, DNA passes through a nanopore at 0.05–1 Mb s^{-1} at an electrostatic potential of ≈100 mV. To detect individual bases thus requires sampling rates of 1 MHz or more. Substantial reduction of the translocation rate can be achieved with processive DNA enzymes,[42-44] which, however, adds extra steps to the sequencing procedure.

In the case of measuring ionic current blockages (Figure 3a), a finite thickness (channel length ≈5 nm) or even electric "read" region[45] (≈3 nm) of the channel of a nanopore might place a fundamental restriction on the single-base resolution because at least 10–15 nucleotides of ssDNA extend through a channel and thus all these nucleotides together contribute to the ionic current blockage. To detect the effect of individual nucleotides on the ionic conductance, a nanopore with channel length comparable to that of a single nucleotide (≈0.4 nm) is required, or a base-recognition element would be required to install at a specific point inside the pore. As we address later, monolayer graphene sandwiched between two insulating layers may meet this requirement.

In addition, noise in the ionic blockade, due to both intrinsically ionic fluctuations and structural fluctuations of the nucleotides, is likely to be too large to successfully distinguish the bases. For instance, the fluctuation of the ionic blockade current for homogeneous sequences is about 30% of the

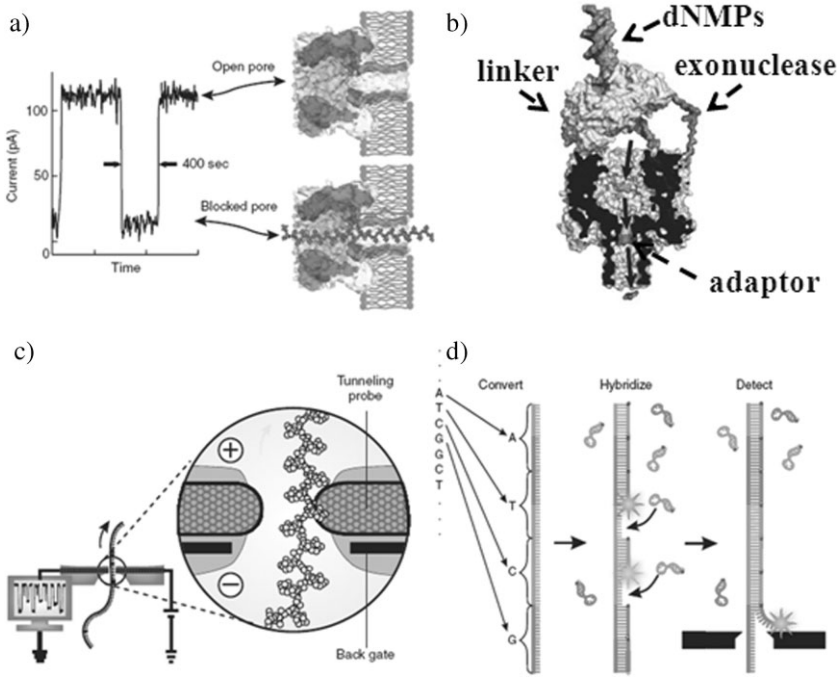


Figure 3. Four concepts to nanopore sequencing. a) Strand-sequencing using ionic current blockage. b) Exonuclease-sequencing by modulation of the ionic current. An exonuclease attached to the top of an α-hemolysin pore through a genetically encoded, or chemical, linker sequentially cleaves deoxynucleoside monophosphates (dNMPs) off the end of a DNA strand (in this case, one strand of a double-stranded DNA). A dNMP's identity (A, T, G, or C) is determined by the level of the current blockade it causes when driven into an aminocyclodextrin adaptor lodged within the pore. After a few milliseconds, the dNMP is released and exits on the opposite side of the bilayer. c) Strand-sequencing using transverse electron currents. d) Nanopore sequencing using synthetic DNA and optical readout. Adapted with permission from Reference [23]. Copyright 2008, Nature Publishing Group.

average current,[36] and fluctuations due to structural changes are larger than the differences between bases.[45] Thus, without amplification of the bases, it is nearly impossible to obtain single-base resolution by sequencing a strand of bases with bare picoampere ionic currents. However, there is potential to go beyond the normal current blockages by engineering nanopores with specific molecules. By covalently attaching an adapter molecule to a protein nanopore[12] (Figure 3b), Oxford Nanopore reported identification of an unlabeled nucleotide 5′-monophoshate molecule with average accuracy of 99.8%.[46] The Oxford Nanopore identified bases[47] but not yet in sequence. The system involved passing chopped-up DNA rather than a complete strand through the pore. It is important for the nanopores–exonuclease approach[15] to assure that 100% of the exonuclease-released deoxynucleoside monophosphates[48] is captured in the pore and efficiently expelled on the opposite side of the membrane, and the sequence of independent readings reflects the order in which the bases are cleaved from the DNA.

Instead of measuring ionic current blockage, Di Ventra et al.[38,39] proposed a protocol for identifying the nucleobases of ssDNA by measurement of transverse tunneling currents through nucleotides that are driven through a nanopore articulated with conductive probes (Figure 3c).[49] Prior to the simulation study of transverse tunneling currents through electrode–nucleotide–electrode by Di Ventra el al.,[38] Golovchenko et al. [50] mentioned the possibility of articulating the solid-state nanopores with electrically conducting electrodes to measure electronic tunneling and perform near-field optical studies. The concept is based on STM, as discussed later, but differs in the use of paired conductive probes. Single-base resolution could be resolved if the transverse tunneling current is from a conductive probe with as small an apex as the tip used in STM. This current origin is not from the nucleotide occupancy through the entire length of the nanopore channel. Control of the electrode–nucleotide–electrode transverse alignment during the DNA translocation through a nanopore is essential for achieving the sequencing of the DNA bases from electrical conductance measurements. The theoretical studies[38,39,49] show that the transverse conductance of a nucleotide is sensitive to its location, orientation, and geometric shape. The variation in the conductance due to the geometry of the molecule relative to the electrodes could overwhelm the difference between different types of nucleotide. Along with the conformation, orientation, and speed control of ssDNA translocation through a nanopore, construction of a suitable nanopore with embedded and precisely aligned paired conductive probes is a formidable challenge. Such nanoelectrodes might be fabricated by the method of programmable pulse electrolytic metal deposition or depletion.[51] For single-base resolution, the dimension of conductive probes must be at the angstrom level to avoid interference from neighbors, and should be robust. Sandwiching single-layer, bilayer, or few-layer graphene or other lamellar conductive materials between two insulating layers may offer a solution. In the sandwich structures, the edge of graphene nanoribbon serves as an electric detector rather than the atomic plane of the sheet, and thus base resolution could be achieved because the thickness of graphene is ≈0.35 nm.[52] Furthermore, the detector of graphene nanoribbon is robust because of the sandwich structure.

It has been observed on ionic blockade current that solid-state nanopores showed higher-level noise than biological pores.[53] The signal-to-noise ratio (SNR) for DNA translocation through solid-state nanopores is functioned as salt concentration and diameter of nanopore, and it is found that the smaller nanopore, the better the SNR.[54] Two dominant noise sources of ionic blockade revealed in silicon nitride nanopores are a high-frequency noise associated with the capacitance of the silicon support chip (dielectric noise), and a low-frequency current fluctuation with $1/f^{\alpha}$ characteristics (flicker noise), where f is frequency. Dielectric noise could be reduced by modifying the SiNx membrane side of the support chip with polydimethylsiloxane,[53] and fabricating Al_2O_3 membrane-based nanopores.[55] At low frequency, the noise of solid-state nanopores greatly exceeds the thermal and shot-noise limits, and varies as $1/f^{\alpha}$ with $0.95 < \alpha < 1.1$ for the range of bias from —400 to +400 mV,[53] showing strong pore-to-pore variations. This low-frequency conductivity ($\sigma = nq\mu$) fluctuation, where n, q, and u are the carrier density, elementary charge, and mobility, respectively, with an $1/f^{\alpha}$ characteristic may be explained by the number fluctuation theory (Δn) [56] due to surface effects and fluctuations in the number of charge carriers due to trapping at surface states, and the mobility fluctuation theory ($\Delta\mu$).[57] It has been demonstrated that large $1/f^{\alpha}$ noise in nanopores could be largely suppressed via surface chemical modification.[58–60] In the case of measurement of transverse tunneling current, fluctuations of the environment, such as ionic and DNA motion, would introduce important scattering processes and in turn electronic noise, and may thus affect the ability to distinguish the bases. Through a simulation, Di Ventra et al.[61] found that such noise would likely not affect the statistical distinguishability of the current distributions obtained from measuring the transverse electronic current of the different DNA nucleotides, due to the off-resonant tunneling nature of the noise through the nucleotide. However, tunneling transport (non-resonant or resonant tunneling) strongly depends on the coupling of a nucleotide to the electrodes, which controls current distributions and might be dominated by non-resonant tunneling.

If these challenges could be overcome, nanopore sequencing has the great potential to enable ultrarapid[39] and cost-effective sequencing of population of DNA molecules. Unlike other methods, such as the Sanger method, the second-generation technologies and those discussed above, the ideal nanopore sequencing approach would not require fluorescent labels or element labels and would use unamplified genomic DNA, thus eliminating enzymes, cloning, and amplification steps. Sample preparation for nanopore sequencing may only require minimal chemistry or enzyme-dependent amplification but eliminates the need for nucleotides and polymerases or ligases during readout. Furthermore, the possibly extremely long reads of ≈50000 bases would greatly simplify the genome assembly and annotation. If so, the cost of a successful nanopore sequencing would be dominated by the cost of the disposable chip and the amortized cost of the instrument, which is estimated to total less than $1000 per mammalian genome. [23]

small NANO MICRO

With the aim of overcoming the challenges of nanopore electrical sequencing, such as the need to engineer nanoelectrodes in the nanopore and the very high translocation rate and electrical noise that can prevent discrimination among single bases, Soni and Meller[62] proposed the development of nanopore single-molecule optical detection approach (Figure 3d). The method involves three basic steps: 1) converting the quaternary DNA code of A, T, G, and C into a binary code in which each base is represented by a pair of 12-mer oligos (A and B); 2) hybridizing the converted DNA mixture with a mixture containing two different 12-mer oligos ''molecular beacons'' designed to complement either A or B; and 3) optically detecting the briefly emitted fluorescence of the molecular beacons as they are sequentially stripped off the complementary converted DNA strand through a <2-nm-diameter nanopore. Although it is difficult to accurately attach a fluorescent probe to each and every nucleotide, and it is a great challenge presently to fabricate 1.7- to 2-nm-diameter nanopores[63] required to strip off the complementary, fluorescently labeled 12-mer oligos, this method is highly suitable for massive parallel readout using high-density arrays of nanopores, which can be simultaneously probed using a single imaging device such as a high-resolution electron-multiplying charge-coupled device camera. Preliminary estimates suggest that arrays of 100 by 100 nanopores could sequence an entire human genome in an hour, driving the cost per genome to $100000 apiece.

6. Scanning Tunneling Microscopy for Sequencing

The application of the atomic resolution of STM began with the legendary images of silicon[64] in 1983 and quickly spread to almost every class of conductive materials as well as DNA molecules.[65–67] The fundamental principle of STM is conceptually rather simple. An atomically sharp metal tip is brought into such close proximity (<1 nm) to a sample surface that an overlap occurs between the tip and the surface electronic wave functions, which decay exponentially in the junction gap. If a small bias voltage is applied to the sample or the tip, electrons can tunnel elastically from filled tip states into sample states or vice versa, depending on the polarity of the applied voltage. This vacuum tunneling establishes a small tunneling current within the nanoampere range. At low bias voltage, STM images represent images of the local density of states at the Fermi level projected to the tip apex above the sample surface. Thus, although STM lacks chemical sensitivity, it can, in principle, be used to detect DNA sequences if there is a difference in electronic characteristics among the four bases. Furthermore, the inelastic electron tunneling spectroscopy (IETS) technique can probe the vibrational modes of a molecule down to the scale of a single bond. The excitation of vibrations is a consequence of inelastic scattering processes that can happen during the tunneling process. If the electrons have sufficient energy to excite a vibrational mode, an additional transport channel opens. Usually this leads to a small but abrupt increase in the tunnel current exactly at the onset of the threshold value of excitation. Thus, IETS can serve as a fingerprinting technique for chemical analysis.

In 2007, Xu et al. reported base-specific electronic signatures of DNA bases.[67] They measured for the first time the four DNA bases deposited separately on a Au (111) surface by ultrahigh vacuum (UHV) STM and noted electronic fingerprints such as bandgap and molecular energy levels of the four bases statistically, along with first-principles calculation. However, overlap of the electronic signatures among the different bases was found. This may be due to conformation variation in the self-assembled films. In 2009, Tanaka and Kawai[68] showed that some bases appear brighter in the conductance image (the derivative of the current–voltage (I–V)) obtained at a chosen bias voltage than the other nucleic acid bases of the stretched single-stranded M13mp18 phage DNA by using low-temperature UHV/STM. By comparing to the known sequence of M13mp18, the brighter bases were identified as the guanine bases in the DNA molecule. When DNA bases interacted with a single-walled carbon nanotube (SWCNT), Kaxiras et al.[69] found theoretically 100% base identifications based on the electronic characteristics among the four DNA bases. These reports indicate the possibility of sequencing DNA using the STM technique. However, the different electronic properties of carbon nanotubes (CNTs) depend on their chirality and diameter, and the DNA wrapping around CNTs (orientation issue) make it experimentally difficult to access for probing individual bases along the DNA molecule. Thus, for electronic DNA sequencing by STM, it is preferable to engineer DNA molecules on a flat surface, and tunneling spectrum acquired by sweeping bias in a wide voltage range can provide many more electronic characteristics[66,67] than current or conductance image generally obtained at a chosen bias voltage.[68]

Enhanced tunneling current was observed between a base-specific modified STM tip and its complementary nucleobase.[70] The current–distance response was used to identify nucleobases with chemically functionalized STM tips in solution.[71] These studies demonstrated the ability to identify chemicals based on selective chemical interactions. For sequencing DNA by a chemical-selective STM tip, however, the reliable preparation of readable samples and nucleobase-modified tips would be challenging. A large molecule used for tip functionalization makes an originally sharp tip blunt and thus loses the single-base resolution. In addition, measurement in solution is a blind observation without temporal identification.

STM-based sequencing (Figure 4) faces a great challenge: how to prepare a suitable ssDNA sample on a solid-state surface. Such a critical issue has not yet been addressed in detail. To date, DNA molecules for STM investigation were directly laid on conductive surfaces such as Au, Cu, and highly ordered pyrolytic graphite (HOPG),[65,72–75] and a few reported high-resolution images.[68,76] Besides helical structures of DNA, the most likely reason for the lack of base resolution is electron coupling[77] of the DNA molecules with the substrate electrons. Note that the size of the assigned guanine bases along the DNA strand by Tanaka and Kawai[68] is not uniform, and there is no obvious gap between neighboring bases, which may indicate the presence of electron coupling. And the conductance (brightness) difference may originate from electron coupling

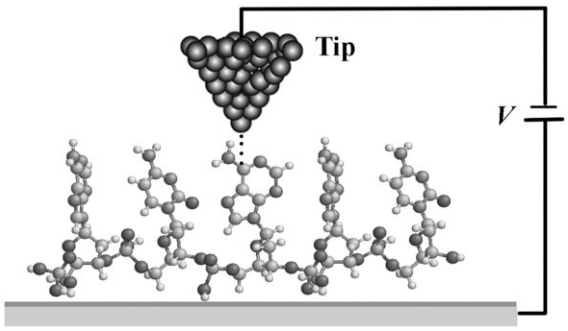


Figure 4. Schematic images of STM-based DNA sequencing; an ssDNA molecule with bases regularly aligned is prepared on an ultrathin insulating layer that is deposited on a conductive substrate.

and conformation fluctuation. To understand and probe the electronic properties of individual DNA molecules, an electronic decoupling of the molecules from the supporting substrate is important. Electronic decoupling of a molecule from a conductive surface can be achieved by inserting an ultrathin insulating layer between the molecules and substrate. For example, ultrathin oxide films, ionic salts, nitride, or alkanethiol layers have been used successfully to decrease the electronic overlap between a molecule and a metal surface.[78] Also, hydrogenation of a clean silicon surface resulted in the passivation of Si surface states.[79]

To straighten DNA molecules on a surface, molecular combing and fluidic biochip approaches may be used. However, no report has shown base resolution and uniformity and consistency for the stretched and linearized DNA molecules. Most works only reported a straight linear shape of the stretched DNA molecules on a surface. In molecular combing,[32] DNA is elongated by means of a receding fluid meniscus as a droplet of solution is allowed to dry on a surface. The solute will migrate towards the boundaries of the droplet in a phenomenon known as the ''coffee-stain effect''. This method of stretching is not inherently well controlled, and it is difficult to predict the degree and uniformity of stretching or placement of the molecules on a surface. Linearization of individual DNA molecules using fluidic biochip platforms also proved relatively inefficient in effecting the desired linearization. Elongation of DNA molecules has been reported using a fluidic nanochannel chip.[80] However, quality of the linearization needs further evaluation. Tanaka and Kawai[68] used an oblique pulse-injection method to deposit the single-stranded M13mp18 phage DNA, showing base-resolution STM image, where the solution of the DNA was injected directly onto a tilted and cleaned Cu surface in an ultrahigh-vacuum chamber with a pulse valve. This method can avoid possible contamination incurred in air. However, it is still unclear whether or not the pulse-injection method could destroy DNA molecules, such as bond and chain breaking, and how to control orientation of the nucleotides is an open question. Such factors can influence conductance as observed in the images.

A well-ordered and stretched ssDNA on a suitable surface that allows electron decoupling between the DNA molecule and substrate is essential for STM investigations and potential STM-based electronic DNA sequencing. The STM-based method does not require imaging of the internal atomic structure of bases, but it does require single-base resolution, so that the STM tip can be located on top of each base individually. Similar to the sample preparation for nanopore sequencing, the genomic DNA needs only to be isolated and separated into ssDNA, and modifications such as cloning of DNA fragments with subsequent amplification and fluorescent labeling are not required. An STM-based approach may support ultrafast sequencing, considering that a scan speed of 15000 nm s^{-1} has already been achieved on flat metal and semiconductor surfaces.[81] If this proves to be the case, we can envision an STM probe with single-base resolution gliding along ssDNA and recording tunneling spectra with sequencing speed up to an estimated 40000 bases s^{-1}.[82] However, at the current stage of a typical STM system, to obtain the scanning tunneling spectrum it is necessary to stop the scanning and feedback of the STM probe and then measure the *I–V* characteristic. As a result, it takes a period of time ranging from 1 s to 1 min for each point. Also, similar to the TEM system, it takes time to find interesting sample locations, and a typical low-temperature UHV/STM instrument costs about US$0.6 million. STM and TEM systems are not so complex; training for a couple of hours enables one to know how to use TEM and UHV/STM.

7. Nano-Knife-Edge Probes for Sequencing (Reveo, Inc.)

Reveo has developed a proprietary method for directly reading sequences using nano-knife-edge probes. The operating methodology of the real-time sequencer (Figure 5a and b) includes forcing straightened and stretched ssDNA into microchannels (≈10-μm width), moving the channel platform at a constant speed, detecting hybridization events in real time with nano-knife-edge probes, and using a gated bias array that controls synchronized current measurement, applied bias voltage, DNA strand motion, or other excitation such as electric stimulus to maximize single nucleotide reaction.[83] In the sequencer system, multiple nano-knife-edge probe arrays/nanonozzles pass over a stretched and immobilized ssDNA in a channel and the potential hybridization events are detected through elastic tunneling, inelastic tunneling, resonantly enhanced tunneling, and/or capacitance by STM-like techniques. The key challenges of the nano-knife-edge probe detection system are to manufacture and functionalize the probe arrays/nanonozzles and to incorporate optical, magnetic, and electric stimuli into the probes (Figure 5c) to enhanced detecting signals. To reduce detection errors, Reveo proposed to measure each nucleotide with 64 sets of the four nano-knife edge probes and to use a fifth probe to detect a methyl group (Figure 5d).

Faris provided many ideas of how to fabricate the probe arrays and nanonozzles with a monolayer of lamellar materials, such as graphene, by folding techniques and described possible methods to detect hybridization events with the nano-knife-edge probes.[83] The probe is characterized by a tip thickness of less than 0.5 nm and a large width. For example, the probe width

NANO MICRO small

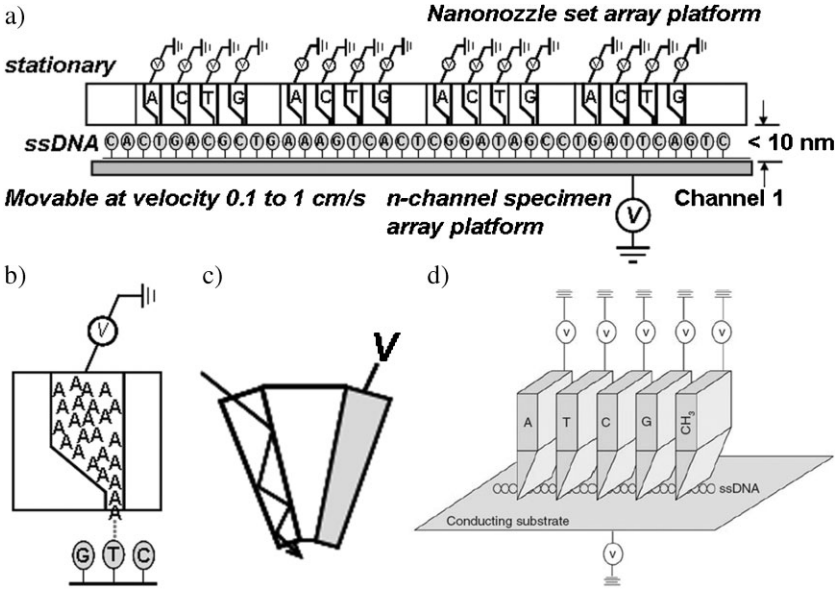


Figure 5. a) An ultrafast real-time sequencer that has a series of probe sets (a probe set including probes or nozzles) for detecting hybridization events. A DNA strand is moved along the channel, and nucleotides from the reservoirs interact with the bases of the strand through the nozzle. The hybridization even provides measurable and detectable current pulses, thereby allowing identification of the base. b) An enlarged view of a nanonozzle with a nucleotide reservoir of A, C, G, or T. c) A probe includes a detection probe and a light nozzle for facilitating detections of hybridization. d) An array of 5 nano-knife-edge probes, in which each nano-edge probe is tuned to recognize a particular base (A, C, T, G) and CH_3 (epigenome, i.e., methylation code) in the stretched DNA strand. Adapted with permission from Reference [86]. Copyright 2008, Nature Publishing Group.

is 10 μm, the same as the channels, so that the probes always land on the ssDNA in the channel to solve the landing error problems that arise because the system cannot directly see the ssDNA. Despite these ideas, no prototype probes have been produced to date.

Note that such an edge probe differs from the probes used in a real STM system, with an atomic apex for atomic resolution. Thus, other theoretical models for the edge probe may be needed to explain the measured results. With respect to typical STM, the most prominent model of an atomic sharp tip is the *s*-wave-tip model. The tip is regarded as a protruding piece of metal with a radius of curvature, and is assumed that only the *s*-wave solutions of the quantum mechanical problem (spherical potential well) are important. Thus, at low bias the tunneling current is proportional to the local density of states of the sample. The nano-knife-edge probes should be robust enough to sequence a long strand, like embedded nanoelectrodes in a nanopore and a tip of STM-based sequencing, not only several bases. The lifetime of such a probe-based detector is a fundamental issue of the sequencing platforms. Furthermore, the probe height and thickness should be uniform enough to detect, for example, tunneling current and force, because tunneling current is exponentially dependent on probe-sample separation (in the STM case) and contact force (in the conductive atomic force microscopy case). If using nucleotide-modified cantilever arrays to measure the interaction force, calibration of the spring constant of the array is needed.

8. Summary and Outlook

The commercial tSMS and SMRT single-molecule sequencing platforms can produce an unprecedented amount of data quickly. The short read-lengths (30–35 bases) resulting from the reversible terminators of tSMS platform, however, limits its application to the sequencing of small and less complex genomes or genome regions and poses challenges for genome assembly and annotation. By contrast, labeling nucleotides with four distinguishable fluorescent dyes on the terminal phosphate group makes SMRT platform possible achieve read-length of many thousands of bases. Improvement in accuracy of this platform, however, is needed by improving the camera detection and signal collection techniques, and modifying fluorescent dyes and polymerase.

Accurately modifying each and every nucleoside in DNA is a common challenge to sequencing individual DNA molecules by detecting the element-labeled nucleotides with TEM and fluorescence of beacon-labeled oligos with nanopores. These chemical modifications change the structures of DNA nucleotides, and they add extra steps that slow sequencing procedure. However, they can add enhanced distinguishability as well.

Because the electronic properties of DNA and bases are sensitive to conformational variation,[49,84,85] it is essential to control base orientation in DNA for electronic sequencing. The nanopore electrical methods would especially support rapid sequencing, but they have to cope with additional solution and base rotation effects by using a recognition element that will both recognize and orient the base, and slow the translocation rate of DNA through the pore. A question of STM-based sequencing is how to get suitable samples of ssDNA with bases regularly aligned for accurate sequencing. Stretching of anchored ssDNA by molecular combing, flow in nanofluidics, or oblique pulse-injection may constitute an inexpensive means of achieving this. Nano-knife-edge probe sequencing-based techniques pose the difficulties of fabricating uniform functionalized probes as well as preparation of single DNA molecules such as the STM-based approach. Robust probes are required for sequencing long DNA strands by probe-based sequencing platforms.

The proliferation of large-scale DNA sequencing projects for applications in clinical medicine and health care has driven the development of alternative methods to reduce time and costs. These new sequencing platforms are providing significantly improved throughput over Sanger methods, as well as second-generation sequencing technologies. Advances in the DNA sequencing field are changing the scientific horizon and promising an era of personalized medicine for improved human health. In this Review, we have examined a variety of different

technologies for inexpensive and rapid single-molecule DNA sequencing, while critically describing their challenges and possible solutions to overcome these. These technologies use either conventional fluorescent or physical approaches to detecting single bases in a sequence. Single-molecule sequencing methods avoid the long timescales, error, and cost associated with sample preparation and amplification. Therefore, these methods would allow researchers to gather genome data faster and at lower cost than presently available sequencing methods. Single-molecule DNA sequencing provides novel direct methods for deciphering the composition of DNA molecules. For example, genomic rearrangements such as insertions, deletions, and inversions that are often associated with cancers or variations within the transcriptome of specific genes may be difficult to detect with conventional sequencing strategies.

The development of single-molecule sequencing technology is just beginning, with a growing number of companies taking very different approaches.[86] An individual human genome has already been sequenced by using Helicos technology.[87] The method of directly detecting electronic differences between the bases is fundamentally different to the methods currently used for sequencing. The similarity of chemical structures and composition and, maybe, of electronic properties of the four bases raises key challenges for the fabrication of sequencing devices, development of detection techniques, preparation of a suitable DNA sample, and control of DNA conformation and orientation during measurement. Ingenious techniques and methodologies are needed to overcome the challenges. Because of the dramatic increase in interest in new sequencing methods, it is difficult to peer even a few years into the future. But, single-molecule DNA sequencing technology is believed to hold great promise for low-cost and rapid sequencing and detection, originating from minimal sample preparation and extremely long read-length.

Acknowledgements

This work was partially supported by the budget for commission of MEXT, Japan, and by the World Premier International Research Center Initiative (WPI Initiative) on Materials Nanoarchitectonics, MEXT, Japan.

Keywords:

DNA sequencing · fluorescence · nanopores · single-molecule studies

[1] *Human Genome Sequencing Consortium* . *Nature* **2001**, *409*, 860–921.
[2] J. C. Venter, *Science* **2001**, *291*, 1304–1351.
[3] F. S. Collins, E. S. Lander, J. Rogers, R. H. Waterston, *Nature* **2004**, *431*, 931–945.
[4] http://grants.nih.gov/grants/guide/rfa-files/RFA-HG-04-003.html.
[5] R. G. Blazej, P. Kumaresan, R. A. Mathies, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 7240–7245.
[6] D. Gresham, M. J. Dunham, D. Botstein, *Nat. Rev. Genet.* **2008**, *9*, 291–302.
[7] M. Margulies, M. Egholm, W. E. Altman, S. Attiya, J. S. Bader, L. A. Bemben, J. Berka, M. S. Braverman, Y. Chen, Z. Chen, S. B. Dewell, L. Du, J. M. Fierro, X. V. Gomes, B. C. Godwin, W. He, S. Helgesen, C. H. Ho, G. P. Irzyk, S. C. Jando, M. L. I. Alenquer, T. P. Jarvie, K. B. Jirage, J. Kim, J. R. Knight, J. R. Lanza, J. H. Leamon, S. M. Lefkowitz, M. Lei, J. Li, K. L. Lohman, H. Lu, V. B. Makhijani, K. E. McDade, M. P. McKenna, E. W. Myers, E. Nickerson, J. R. Nobile, R. Plant, B. P. Puc, M. T. Ronan, G. T. Roth, G. J. Sarkis, J. F. Simons, J. W. Simpson, M. Srinivasan, K. R. Tartaro, A. Tomasz, K. A. Vogt, G. A. Volkmer, S. H. Wang, Y. Wang, M. P. Weiner, P. Yu, R. F. Begley, J. M. Rothberg, *Nature* **2005**, *437*, 376–380.
[8] J. Shendure, H. L. Ji, *Nat. Biotechnol.* **2008**, *26*, 1135–1145.
[9] A. Kahvejian, J. Quackenbush, J. F. Thompson, *Nat. Biotechnol.* **2008**, *26*, 1125–1133.
[10] H. Bayley, *Curr. Opin. Chem. Biol.* **2006**, *10*, 628–637.
[11] P. K. Gupta, *Trends in Biotechnol.* **2009**, *26*, 602–611.
[12] Y. Astier, O. Braha, H. Bayley, *J. Am. Chem. Soc.* **2005**, *128*, 1710–1715.
[13] I. Braslavsky, B. Herbert, E. Kartalov, S. R. Quake, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3960–3964.
[14] M. J Levene, J. Korlach, S. W. Turner, M. Foquet, H. G. Craighead, W. W. Webb, *Science* **2003**, *299*, 682–686.
[15] J. H. Jett, R. A. Keller, J. C. Martin, B. L. Marrone, R. K. Moyzis, R. L. Ratliff, N. K. Seitzinger, E. B. Shera, C. C. Stewart, *J. Biomol. Struct. Dyn.* **1989**, *7*, 301–309.
[16] M. Sauer, B. Angerer, W. Ankenbauer, Z. Földes-Papp, F. Göbel, K. T. Han, R. Rigler, A. Schulz, J. Wolfrum, C. Zander, *J. Biotechnol.* **2001**, *86*, 181–201.
[17] J. H. Werner, H. Cai, J. H. Jett, L. Reha-Krantz, R. A. Keller, P. M. Goodwin, *J. Biotechnol.* **2002**, *102*, 1–14.
[18] C. P. Fredlake, D. G. Hert, E. R. Mardis, A. E. Barron, *Electrophoresis* **2006**, *27*, 3689–3702.
[19] E. Pettersson, J. Lundeberg, A. Ahmadian, *Genomics* **2009**, *93*, 105–111.
[20] M. Pop, S. L. Salzberg, *Trends in Genet.* **2008**, *24*, 142–149.
[21] E. R. Mardis, *Trends in Genet.* **2008**, *24*, 133–141.
[22] M. Zwolak, M. Di Ventra, *Rev. Mod. Phys.* **2008**, *80*, 141–165.
[23] D. Branton, D. W. Deamer, A. Marziali, H. Bayley, S. A. Benner, T. Butler, M. Di Ventra, S. Garaj, A. Hibbs, X. Huang, S. B. Jovanovich, P. S. Krstic, S. Lindsay, X. S. Ling, C. H. Mastrangelo, A. Meller, J. S. Oliver, Y. V. Pershin, J. M. Ramsey, R. Riehn, G. V. Soni, V. Tabard-Cossa, M. Wanunu, M. Wiggin, J. A. Schloss, *Nat. Biotechnol.* **2008**, *26*, 1146–1153.
[24] T. D. Harris, P. R. Buzby, H. Babcock, E. Beer, J. Bowers, I. Braslavsky, M. Causey, J. Colonell, J. DiMeo, J. W. Efcavitch, E. Giladi, J. Gill, J. Healy, M. Jarosz, D. Lapen, K. Moulton, S. R. Quake, K. Steinmann, E. Thayer, A. Tyurina, R. Ward, H. Weiss, Z. Xie, *Science* **2008**, *320*, 106–109.
[25] J. Eid, A. Fehr, J. Gray, K. Luong, J. Lyle, G. Otto, P. Peluso, D. Rank, P. Baybayan, B. Bettman, A. Bibillo, K. Bjornson, B. Chaudhuri, F. Christians, R. Cicero, S. Clark, R. Dalal, A. deWinter, J. Dixon, M. Foquet, A. Gaertner, P. Hardenbol, C. Heiner, K. Hester, D. Holden, G. Kearns, X. Kong, R. Kuse, Y. Lacroix, S. Lin, P. Lundquist, C. Ma, P. Marks, M. Maxham, D. Murphy, I. Park, T. Pham, M. Phillips, J. Roy, R. Sebra, G. Shen, J. Sorenson, A. Tomaney, K. Travers, M. Trulson, J. Vieceli, J. Wegener, D. Wu, A. Yang, D. Zaccarin, P. Zhao, F. Zhong, J. Korlach, S. Turner, *Science* **2009**, *323*, 133–138.
[26] M. Foquet, K. T. Samiee, X. Kong, B. P. Chauduri, P. M. Lundquist, S. W. Turner, J. Freudenthal, D. B. Roitman, *J. Appl. Phys.* **2008**, *103*, 034301.
[27] P. M. Lundquist, C. F. Zhong, P. Zhao, A. B. Tomaney, P. S. Peluso, J. Dixon, B. Bettman, Y. Lacroix, D. P. Kwo, E. McCullough, M. Maxham, K. Hester, P. McNitt, D. M. Grey, C. Henriquez, M. Foquet, S. W. Turner, D. Zaccarin, *Opt. Lett.* **2008**, *33*, 1026–1028.

[28] J. Korlach, A. Bibillo, J. Wegener, P. Peluso, T. T. Pham, I. Park, S. Clark, *Nucleosides, Nucleotides, Nucleic Acids* **2008**, *27*, 1072–1083.
[29] A. Sood, S. Kumar, S. Nampalli, J. R. Nelson, J. Macklin, C. W. Fuller, *J. Am. Chem. Soc.* **2005**, *127*, 2394–2395.
[30] S. Hardin, X. L. Gao, J. Briggs, R. Willson, S. C. Tu, Pat. US7329492.
[31] W. R. Glover, III, Pat. US7288379 B2.
[32] A. Bensimon, A. Simon, A. Chiffaudel, V. Croquette, F. Heslot, D. Bensimon, *Science* **1994**, *265*, 2096–2098.
[33] C. Dekker, *Nat. Nanotechnol.* **2007**, *2*, 209–215.
[34] S. M. Iqbal, D. Akin, R. Bashir, *Nat. Nanotechnol.* **2007**, *2*, 243–248.
[35] J. J. Kasianowiez, E. Brandin, D. Branton, D. W. Deamer, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 13770–13773.
[36] A. Meller, L. Nivon, E. Brandin, J. Golovchenko, D. Branton, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 1079–1084.
[37] N. Ashkenasy, J. Sanchez-Quesada, H. Bayley, M. R. Ghadiri, *Angew. Chem. Int. Ed.* **2005**, *44*, 1401–1404.
[38] M. Zwolak, M. Di Ventra, *Nano Lett.* **2005**, *5*, 421–424.
[39] J. Lagerqvist, M. Zwolak, M. Di Ventra, *Nano Lett.* **2006**, *6*, 779–782.
[40] M. E. Gracheva, A. Xiong, A. Aksimentiev, K. Schulten, G. Timp, J. Leburton, *Nanotechnology* **2006**, *17*, 622–633.
[41] G. Sigalov, J. Comer, G. Timp, A. Aksimentiev, *Nano Lett.* **2008**, *8*, 56–63.
[42] B. Hornblower, A. Coombs, R. D. Whitaker, A. Kolomeisky, S. J. Picone, A. Meller, M. Akeson, *Nat. Methods* **2007**, *4*, 315–317.
[43] S. Benner, R. J. A. Chen, N. A. Wilson, R. Abu-Shumays, N. Hurt, K. R. Lieberman, D. W. Deamer, W. B. Dunbar, M. Akeson, *Nat. Nanotechnol.* **2007**, *2*, 718–724.
[44] S. L. Cockroft, J. Chu, M. Amorin, M. R. Ghadiri, *J. Am. Chem. Soc.* **2008**, *130*, 818–820.
[45] A. Aksimentiev, J. B. Heng, G. Timp, K. Schulten, *Biophys. J.* **2004**, *87*, 2086–2097.
[46] J. Clarke, H. C. Wu, L. Jayasinghe, A. Patel, S. Reid, H. C. Bayley, *Nat. Nanotechnol.* **2009**, *4*, 265–270.
[47] D. Stoddart, A. Heron, E. Mikhailova, G. Maglia, H. Bayley, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 7702–7707.
[48] H. C. Wu, Y. Astier, G. Maglia, E. Mikhailova, H. Bayley, *J. Am. Chem. Soc.* **2007**, *129*, 16142–16148.
[49] X. G. Zhang, P. S. Krstic, R. Zikic, J. C. Wells, M. Fuentes–Cabrera, *Biophys. J.* **2006**, *91*, L04–L06.
[50] J. L. Li, M. Gershow, D. Stein, E. Brandin, J. A. Golovchenko, *Nat. Mater.* **2003**, *2*, 611–615.
[51] J. W. Lee, T. G. Thundat, Pat. US6905586 B2.
[52] A. K. Geim, K. S. Novoselov, *Nat. Mater.* **2007**, *6*, 183–191.
[53] V. Tabard-Cossa, D. Trivedi, M. Wiggin, N. N. Jetha, A. Marziali, *Nanotechnology* **2007**, *18*, 305505.
[54] R. M. M. Smeets, U. F. Keyser, N. H. Dekker, C. Dekker, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 417–421.
[55] B. M. Venkatesan, B. Dorvel, S. Yemenicioglu, N. Watkins, I. Petrov, R. Bashir, *Adv. Mater.* **2009**, *21*, 2771–2776.
[56] A. L. McWhorter, PhD Dissertation MIT Cambridge, MA, **1955**.
[57] F. N. Hooge, *Phys. Lett. A* **1969**, *29*, 139.
[58] P. Chen, T. Mitsui, D. B. Farmer, J. Golovchenko, R. G. Gordon, D. Branton, *Nano Lett.* **2004**, *4*, 1333–1337.
[59] C. Danelon, C. Santschi, J. Brugger, H. Vogel, *Langmuir* **2006**, *22*, 10711–107115.
[60] J. Nilsson, J. R. I. Lee, T. V. Ratto, S. E. Letant, *Adv. Mater.* **2006**, *18*, 427–431.
[61] M. Krems, M. Zwolak, Y. V. Pershin, M. Di Ventra, arXiv:0903.0670v2.
[62] G. V. Soni, A. Meller, *Clin. Chem.* **2007**, *53*, 1996–2001.
[63] A. F. Sauer-Budge, J. A. Nyamwanda, D. K. Lubensky, D. Branton, *Phys. Rev. Lett.* **2003**, *90*, 238011.
[64] G. Binnig, H. Rohrer, C. Gerber, E. Weibel, *Phys. Rev. Lett.* **1983**, *50*, 120–123.
[65] W. M. Heckl, G. Binnig, *Ultramicroscopy* **1992**, *42*, 1073–1078.
[66] M. S. Xu, S. Tsukamoto, S. Ishida, M. Kitamura, Y. Arakawa, R. G. Endres, M. Shimoda, *Appl. Phys. Lett.* **2005**, *87*, 083902.
[67] M. S. Xu, R. G. Endres, Y. Arakawa, *Small* **2007**, *3*, 1539–1543.
[68] H. Tanaka, T. Kawai, *Nat. Nanotechnol.* **2009**, *4*, 518–522.
[69] S. Meng, P. Maragakis, C. Papaloukas, E. Kaxiras, *Nano Lett.* **2007**, *7*, 45–50.
[70] T. Ohshiro, Y. Umezawa, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 10–14.
[71] J. He, L. Lin, P. Zhang, S. Lindsay, *Nano Lett.* **2007**, *7*, 3854–3858.
[72] D. J. Driscoll, M. G. Youngquist, J. D. Baldeschwieler, *Nature* **1990**, *346*, 294–296.
[73] C. R. Clemmer, T. P. Beebe, *Science* **1991**, *251*, 640.
[74] N. J. Tao, J. A. Derose, S. M. Lindsay, *J. Phys. Chem.* **1993**, *97*, 910–919.
[75] E. Shapir, J. Yi, H. Cohen, A. B. Kotlyar, G. Cuniberti, D. Porath, *J. Phys. Chem. B* **2005**, *109*, 14270–14274.
[76] H. Tanaka, T. Kawai, *Surf. Sci.* **2003**, *539*, L531.
[77] J. Repp, G. Meyer, S. M. Stojkovic, A. Gourdon, C. Joachim, *Phys. Rev. Lett.* **2005**, *94*, 026803.
[78] K. J. Franke, G. Schulze, N. Henningsen, I. Fernandez-Torrente, J. I. Pascual, S. Zarwell, K. Ruck-Braun, M. Cobian, N. Lorente, *Phys. Rev. Lett.* **2008**, *100*, 036807.
[79] A. Bellec, F. Ample, D. Riedel, G. Dujardin, C. Joachim, *Nano Lett.* **2008**, *8*, 144–147.
[80] H. Cao, J. Tegenfeldt, R. Austin, S. Y. Chou, *App. Phys. Lett.* **2002**, *81*, 3058.
[81] R. Curtis, T. Mitsui, E. Ganz, *Rev. Sci. Instrum.* **1997**, *68*, 2790–2796.
[82] M. S. Xu, R. G. Endres, Y. Arakawa, *Perspectives from Physics, Chemistry, and Biology*, (Ed.: Tapash Chakraborty,), Springer, Berlin **2007**, p. 205–220.
[83] S. M. Faris, Pat. US 2007/0082459 A1.
[84] R. G. Endres, D. L. Cox, R. R. P. Singh, *Rev. Mod. Phys.* **2004**, *76*, 195–214.
[85] M. S. Xu, R. G. Engres, S. Tsukamoto, M. Kitamura, S. Ishida, Y. Arakawa, *Small* **2005**, *1*, 1168–1172.
[86] N. Blow, *Nat. Methods* **2008**, *5*, 267–274.
[87] D. Pushkarev, N. F. Neff, S. R. Quake, *Nat. Biotechnol.* **2009**, *27*, 847–850.

Received: June 8, 2009

