Sequencing technologies for the elusive DNA molecule: let's have a closer look

Evangelia Voulgari, Nicholaos Spanakis, Athanassios Tsakris
Department of Microbiology, Medical School, National and Kapodistrian University of Athens, Athens, Greece
DOI: https://doi.org/10.5281/zenodo.10020698

Summary

Over the past decade, impressive progress has been made in the field of genome sequencing due to the introduction of novel platforms capable of massive parallel sequencing. Herein we aim to present in a concise manner the extraordinary journey which has taken place and the following evolutionary technological pathway, beginning from the early days of Sanger sequencing and leading up to the newer second and third generation platforms. Several commonly used next generation sequencing (NGS) technologies will be presented and compared with Sanger sequencing and some of the challenges facing these newer platforms will also be debated.

Corresponding author
Athanassios Tsakris
Department of Microbiology, Medical School,
National and Kapodistrian University of Athens
75, Mikras Asias str., 11527 Athens, Greece
Tel. +30 210 7462011
Fax: +30 210 7462210
e-mail: atsakris@med.uoa.gr

Key words
sequencing, platforms, next generation sequencing

Introduction

Ever since 1869 and the first identification by Swiss physiological chemist Friedrich Miescher of what he called "nuclein" inside the nuclei of human white blood cells, the DNA molecule with its double helix structure, as later presented by Watson and Crick, has mystified the scientific community, which has sought out ways to 'get a closer look' at this elusive molecule and unravel its precise function.[1,2] Seeking to develop a 'molecular microscope' fit for genome analysis, has certainly been a painstaking process, it has demanded small consecutive steps, which have evolved from the introduction of electrophoresis as a visualization method, to the use of polymerase for DNA replication, restriction enzymes for cutting bulk DNA and enabling its manipulation and molecular techniques such as polymerase chain reaction (PCR).[3–6] However, in this overall effort it was the revolution of the introduction of DNA sequencing technologies, initially in the form of Sanger sequencing followed by a variety of newer next generation sequencing (NGS) methodologies that radically altered our understanding of the complexity of the DNA structure and its host interactions. In order to fully comprehend the immensity of the events that took place, we need to go back to the early years and understand the significance of the introduction of Sanger sequencing, otherwise known as first generation sequencing, which effectively monopolized the field for at least thirty years up to the gradual introduction of novel technologies from 2005. The goal of this review will also be to present in a clear and easily comprehendible fashion the technological concepts behind some of these NGS technologies regardless of whether they are currently commercially available. For further reading however, comprehensive and detailed reviews will be cited throughout.

First generation sequencing

In the 70s two novel non-enzymatic approaches were developed for DNA sequencing one by Sanger and Coulson and another by Maxam and Gilbert.[7,8] Difficulties associated with shorter reads and the use of hazardous chemicals in the Maxam and Gilbert approach hindered it from ever taking off and allowed for the domination of the Sanger methodology.

As mentioned, in contrast to previous non-enzymatic methods, Sanger sequencing uses a DNA polymerase to restructure a mix of complementary DNA molecules against a template strand and electrophoresis in polyacrylamide gels to separate the mix of resulting products according to their nucleotide sequence length. In these reactions, DNA is initially treated with restriction enzymes to provide the convenient length which is then used as a template. Synthesis of the complementary strand is terminated in a sequence-specific manner by supplying only one of the four nucleoside triphosphates ('plus' reactions), or else three of the four ('minus' reactions).[9,10,8] This 'plus and minus' method presented issues which were subsequently addressed by the 'dideoxy' method in 1977, which introduced the use of radioactively labeled chain terminating dideoxynucleotides (ddNTPs) in the synthesis process.[11] In its final form, the Sanger sequencing method combines the synthesis of a complementary DNA template using natural 2-deoxynucleotides (dNTPs) by a DNA polymerase and irreversible termination of the process with the incorporation of synthetic 2,3-dideoxynucleotides (ddNTPs). These ddNTP's are lacking the necessary OH group at the 3-position of the deoxyribose molecule for the formation of a phosphodiester bond, causing base specific termination and impediment of chain elongation (Figure 1). The sequencing specificity is controlled in each of the above methods by the use of specific or randomly annealing primers to the template DNA strands.

Competition between synthesis and termination processes results in the generation of a set of fragments variable in length. By fine tuning the ratio of ddNTP/dNTP in the reaction mix, usually at a 1% for each NTP, the frequency of chain termination and therefore the length of the terminated chains, can be tailored to produce a series of chain-terminated products, each ending with a selected NTP. The mixture of terminated DNA chains are subsequently electrophorised in four parallel acrylamide gels and manually analyzed in order of length to reveal the initial DNA sequence. Given the technical time consuming difficulties in this raw process, for Sanger sequencing to be established as a functional sequencing technology, additional breakthroughs had to be incorporated in its core concept. The automation which enabled the decrease in cost, manual labor, turnover times and thus established first generation sequencing, was gradually developed over the next ten years following the introduction of Sanger sequencing and was made possible only after additional scientific advances in the form of fluorescent dye development, automatic laser fluorescence detection, capillary electrophoresis and robotics were incorporated into the commercial platform.[9]

Other crucial advances included the introduction of PCR for DNA yield, which was addressed initially via cloning of the DNA fragments of interest and of co-

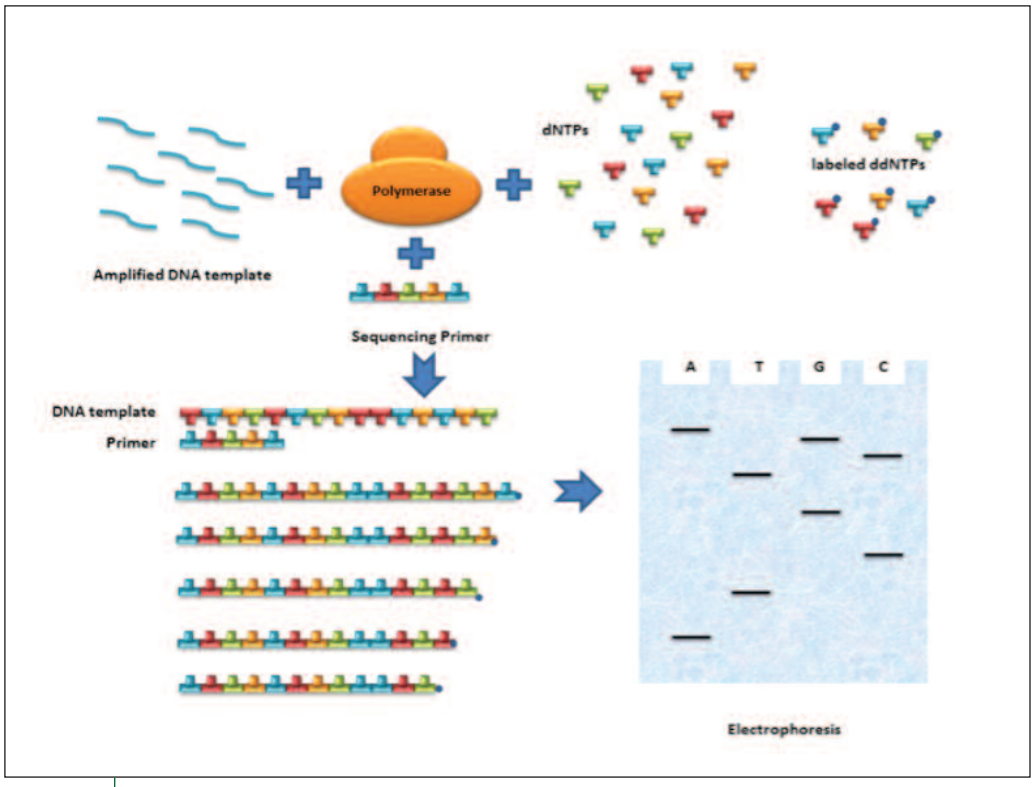

Figure 1 Sanger sequencing with the dideoxy method.

urse the use of bioinformatics for the interpretation of the acquired data. These incremental improvements to the original concept resulted in an more simple pre-analytical step and led to the introduction by Applied Biosystems of the automated, commercially viable platforms, which brought to a pass the 'Human Genome Project', an incredible undertaking to map the genome of a single human, in approximately 10 years' time and at a cost estimated from 0,5-1 billion dollars.[12]

Second and third generation sequencing

However, since 2005 novel technologies have emerged and have effectively challenged the supremacy of the 'dideoxy' method in the race for cheaper and faster methodologies and devices, addressing an audience not necessarily equipped with even a rudimentary bioinformatics background. These technologies while using different strategies have two common characteristics, the ability to produce an enormous amount of data cheaply though massive parallel sequencing and the use of benchtop platforms far more accessible to individual labs.

I) Template preparation

A very significant step in NGS is template preparation. In contrast to Sanger sequencing, NGS requires that DNA be fragmented in to small pieces preferably <1000bp and then adhered to a known region usually in the form of a vector or adaptor sequence to which a universal primer can bind so that the template can be attached or immobilized to a solid surface or support.[13] DNA templates can be prepared in NGS either by clonally amplifying a single DNA molecule or by preparing a single DNA molecule template. The former is achieved either by emulsion PCR (EmPCR) or solid-phase amplification.[14,15]

II) Clonal amplification templates

In EmPCR the key element is a population of agarose beads with attached oligonucleotides complementary to the adaptors which have been used to create the template library (Figure 2). The basic idea is that each bead attracts and attaches a single DNA fragment and that this bead/fragment complex afterwards is isolated in an individual oil/water droplet, creating a micro reactor that holds all the necessary reagents for the completion of multiple thermal PCR cycles.[16] As the reaction progresses, each DNA frag-

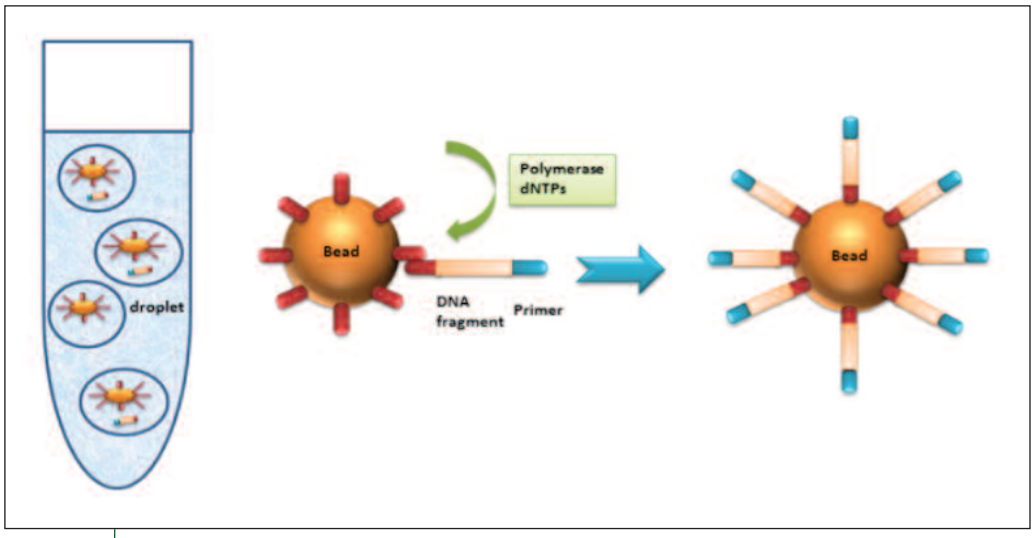

Figure 2 Emulsion PCR.

ment produces a complementary strand which is then dissociated and relocated upon the bead for a new amplification cycle to be initiated. By the end of this process each bead bears on its surface, at the sites of the complementary oligonucleotides, millions of copies of the initial DNA fragment, which can then be sequenced simultaneously. EmPCR is used for library preparation in several sequencing technologies such as pyrosequencing, semiconductor sequencing technologies and sequencing by ligation.

In solid phase amplification instead of a liquid environment as in emPCR, a glass slide is used as a surface for amplification. Upon this slide (usually a flow-cell) complementary forward and reverse primers are fixated at a high-density. The initial DNA fragments are adjusted with adaptors capable of attaching to the glass slide oligonucleotides and to sequencing primers as well (Figure 3). As the adaptor-bearing DNA fragment is attached to the 'docking' oligonucleotide, a DNA polymerase begins the synthesis of a complementary strand with the resulting dsDNA covalently attached to the glass surface. After this synthesis step the newly synthetized strand remains attached while the original DNA fragment is washed away. Clustering is then performed by bridge amplification. As the term denotes amplification is literally performed with the formation of a 'bridge'. The remaining DNA stand flips over and is attached to an adjacent complementary oligonucleotide on the slide, forming a 'bridge'. The polymerase then produces a new complementary strand forming a double stranded bridge which is then denatured leaving two single DNA strands tethered to the slide surface. As this process is repeated over and over a cluster of DNA is prepared, with free ends to which a universal sequencing primer can then be hybridized.

III) Single DNA molecule templates

Methodologies using single differ from the previous in that they to not implicate the use of PCR. The basic concept is to immobilize the single molecule template onto a solid surface prior to the NGS reaction. The initial DNA is fragmented into small sizes at the ends of which adaptors are attached. This template DNA is then captured by immobilized oligomers on glass cover slips.[13,17] After binding, these oligomers are used either as a primer for the template-directed primer extension that forms the basis of the sequence reading or for a template replication step which produces a complementary single stand with a hybridized common primer at its unattached end to be used later for sequencing.[17] Alternative procedures involve either the fixation of the polymerase molecules to a solid support and the adhesion to this polymerase of the primed template molecule or the fixation of motor enzymes and a tether to the template DNA and its attachment to an immobilized nanopore for unzipping and sequencing.[18,19]

Sequencing and imaging

Following DNA library preparation the goal is to sequence the resulting product. In order to clarify how the various sequencing technologies work, existing platforms regardless of their commercial availability, will presented in a simple manner.

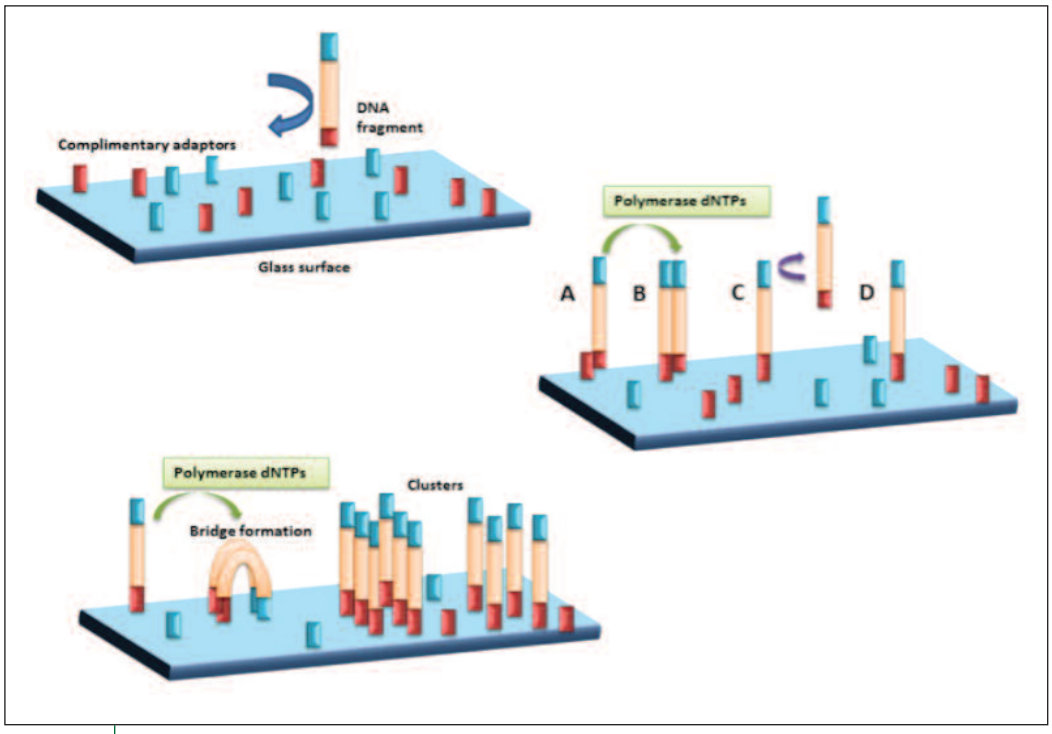

Figure 3 Bridge amplification forms clonally amplified clusters.

1) Pyrosequencing: 454 platform: Life sciences/Roche

In pyrosequencing, as applied in the Life sciences/Roche, 454 platform, the completion of the DNA library via emPCR is followed by the loading of the DNA amplified beads into individual PicoTiterPlate (PTP) wells where upon addition of beads, smaller in size, flood each well along with a complementary to the library adaptor universal primer, a DNA polymerase and a nucleotide degrading enzyme such as apyrase. Each well is designed to hold only one bead although some wells on the plate may hold no bead at all. The smaller beads are coated with two enzymes: ATP-sulphurylase and luciferase (Figure 4).[20] Individual dNTPs are streamed across the wells and dispensed in a predetermined sequential order in which the added nucleotides are degraded by apyrase prior to the addition to the next dNTP.[21] With the addition of a complementary dNTP by the polymerase to the single stranded DNA fragment on the bead, a pyrophosphate (PPi) is released. This production is equal to the molality of the incorporated dNTP and is thereafter indirectly quantified by the light produced in a luciferase-catalyzed reaction. In particular ATP-sulphurase converts these PPis into ATP which is translated by luciferase into light and oxyluciferin. It is this light omission which is then quantified by a charged coupled device camera (CCD) and translated into sequencing data in a flowgram.[21]

2) Semiconductror sequencing: Ion torrent: Life technologies/ThermoFisher

The semiconductor method as used in Life Technologies Ion Torrent's semiconductor applies emPCR for library preparation, uses microwells upon which the beads are distributed where again a sequencing by

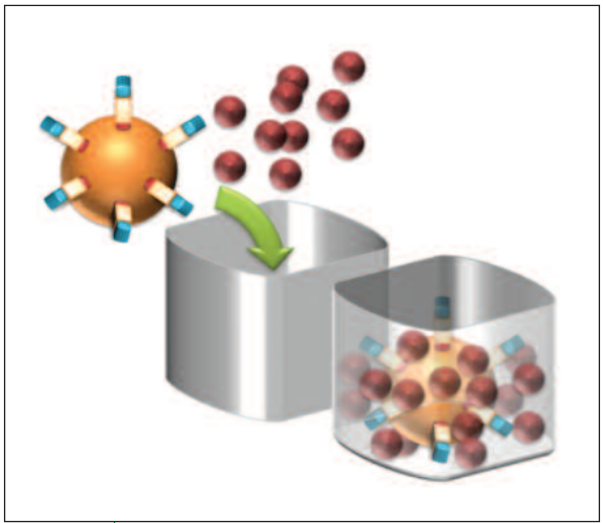

Figure 4 In pyrosequencing the agarose beads are distributed into small wells, one bead per well, which are then flooded by smaller beads bearing sulphurylase and luciferase.

synthesis reaction takes place but unlike pyrosequencing which relies on light detection, the Torrent's semiconductor measures pH changes induced by the release of hydrogen ions during DNA extension.[22,23] These pH changes are detected by a sensor positioned at the bottom of the microwell and converted into a voltage signal. The voltage signal is proportional to the number of bases incorporated, and the sequential addition of individual nucleotides during each sequencing cycle allows base discrimination.[22]

3) Sequencing by ligation : SOLiD: Applied Biosystems

Sequencing by ligation which also relies on emPCR for library preparation, was commercially launched in the SOLiD platform from Applied Biosystems. In this method the emPCR resulting beads are randomly fixated to a glass surface. Then a universal primer is attached to one of the library adaptors and the flowcell is flooded by a library of 1,2 di base- fluorescently labelled probes consisting of 8 nucleotides in length (Figure 5).[24] These probes consist of 2 interrogation bases, 3 degenerate bases and 3 universal bases to which the fluorescent dye is attached. Each probe is labelled by a different dye representing 4 out of 16 possible dinucleotide combinations. The 3' interrogating probe end is ligated to the universal primer and its labelled 3 base tail is then cleaved off reducing the probe to a 5 nucleotide size allowing for a three base interval. Fluorescent emission is measured and the ligation process is repeated nine times-cycles. Following the completion of the initial cycle the resulting DNA strand is removed and the next ligation round is initiated beginning with the addition of another universal primer, off set however this time by one base. Each template is interrogated twice and compiled into a string with bits of color intermitted by space data.[13] These reads are then aligned to decode the DNA.

4) Sequencing with reversible terminators: Illumina/Solexa

One of the most commercially successful vendors in the sequencing industry is Solexa/Illumina with its Illumina platform. The system's library preparation is based on the creation of DNA clusters by bridge amplification deriving from a single initial fragment and sequencing is performed by synthesis on the solid surface of the flow cell.[16] The Illumina system also utilizes a 'wash and scan' method as in the previous cases. In contrast to the aforementioned methodologies, the four fluorescently labelled nucleotides are added simultaneously to the flow cell along with a polymerase, responsible for incorporating them into the oligo-primed cluster fragments. Each fixated DNA fragment bears at its free end the universal adaptor to which a primer is attached. Each nucleotide incorporation is a unique event, since the fluorescent label at the 3'-OH group acts as a chemical blocker, ensuring that only one nucleotide is incorporated and detected at a time. This chemical 'stop' signal is only removed after the dye has been cleaved and the emitted fluorescence detected, leaving a regenerated 3'-OH end for the elongation of the strand to be continued. In this way four color images are acquired by total internal reflection fluorescence (TIRF) imaging using two lasers.

5) Single molecule detection: Helicoscope: Helicos BioSciences

Another platform by Helicos BioSciences, the Helicscope, also uses this cyclic reversible termination method, however in this case the nucleotides are connected to a single dye and are not dispensed simultaneously but in a predefined order. Again the dye is cleaved, imaging is performed and the wells are washed in order for the next dye labelled dNTP to be added. During the imaging process, four lasers illuminate 1100 Fields of View (FOV) per channel with pictures taken by four CCD (Charge-coupled device) cameras via a powerful confocal microscope.

6) Single molecule sequencing: PacBioRS: Pacific Biosciences

A different approach is used in the sequencing technology currently referred to as second and a half generation, which has been commercialized by Pacific Biosciences and in which the single molecule real time (SMRT) methodology is applied. The main difference with the second generation technologies lies in the fact that library preparation does not require amplification neither in the form of emPCR or clusters. Instead a single digested DNA fragment with hairpin adaptors is used for generating a template.[18] To increase accuracy this fragment can be sequenced multiple times. The polymerase is fixated at the bottom of zeptoliter-sized chambers called zero mode waveguides (ZMWs) which are located on an SMRT chip, in such a way that each chamber contains only one molecule of polymerase.[22,25] These ZMWs are in fact dark cylindrical holes in which the polymerase is confined in a tiny volume at the bottom of the well (Figure 6). In this real-time sequencing method no reversible terminators are used. Instead it is more a process of directly monitoring the incorporation of nucleotides within the DNA strand during DNA synthesis.[18] These nucleotides are of course dye-labelled (phospholinked hexaphosphate nucleotides) and are added simultaneously so that the reaction progresses in a continuous manner.[26] Fluorescence is detected solely

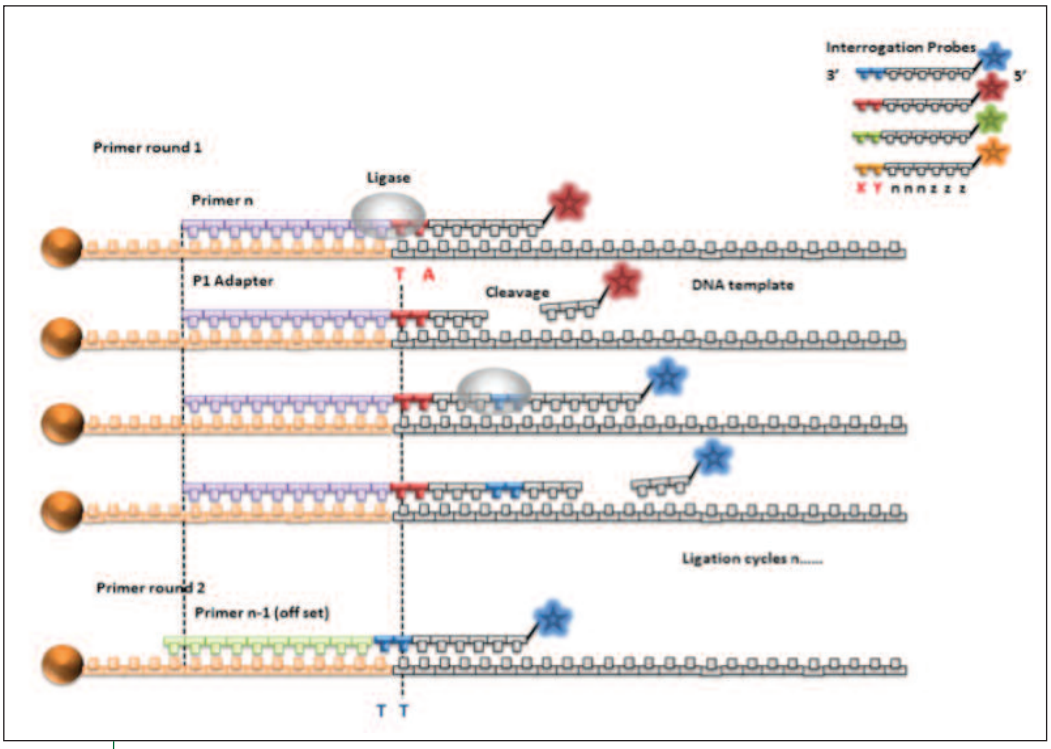

Figure 5 Sequencing by ligation.

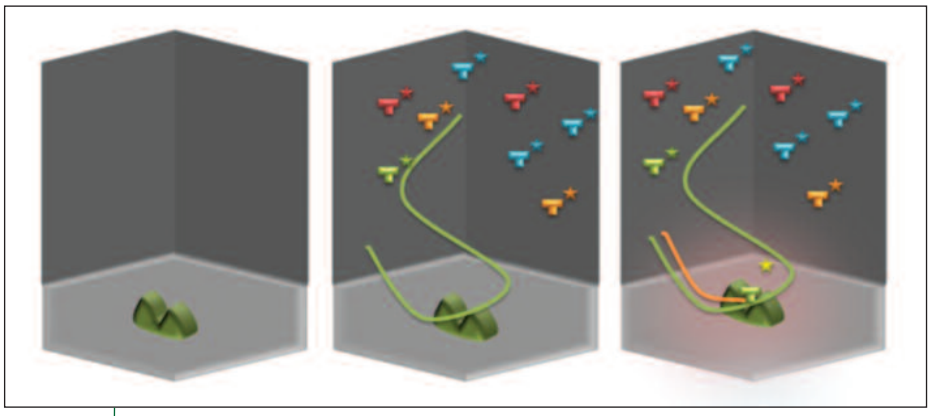

Figure 6 Single molecule sequencing.

within the tiny area occupied by the polymerase and following incorporation the dyes are diffused into the dark background.[24]

7) Sequencing with nanopores: Minion: Oxford Nanopore technologies

Nanopore sequencing also bypasses issues concerning emPCR and cluster formation. Libraries are created by single DNA templates ligated with two adapters (Figure 7). Like the SMRT sequencing this can be done with or without PCR amplification.[22] The first adaptor is bound with a motor enzyme as well as a tether, whereas the second adaptor is a hairpin oligo that is bound by the so called HP motor protein. The concept is that this DNA molecule is attached to a tine pore (nanopore) stabilized on a synthetic bilayer, where it is unzipped by the motor protein and as it passes through the nanopore the changes in induced current areused to discriminate bases and sequence the strand.[27,28] All four nucleotides are provided in a step-

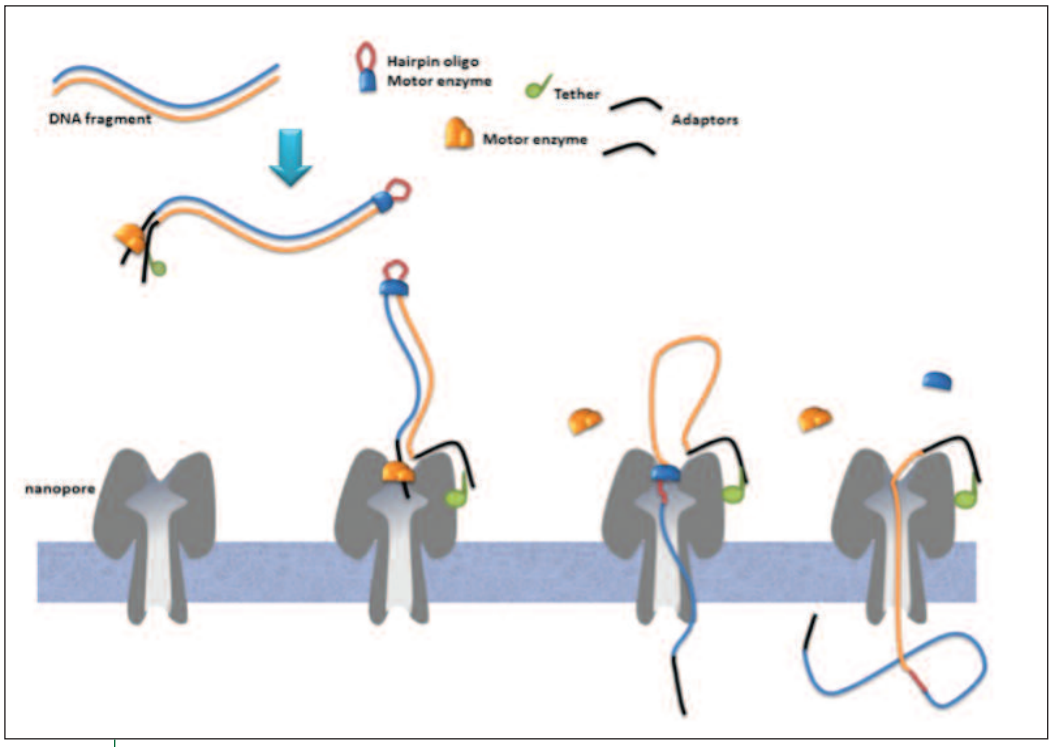

Figure 7 Sequencing with nanopores.

wise fashion during an automated run.[23] The library design with the use of two motor proteins, one at each end, allows sequencing of both strands of DNA from a single molecule (two-direction reads) increasing accuracy.[28]

Concluding remarks

Given the importance of DNA sequencing to life sciences, biotechnology and medicine the primary goal became the development of more scalable and lower cost solutions. This demand escalated with the completion of 'The Human Genome Project', given the vast amount of time and resources that had been required. Furthermore, it became apparent that it was essential that the throughput, that is the DNA that could be processed, which in Sanger sequencing adhered to the one DNA molecule one PCR concept, had to be increased in order to avoid the factory-like centers with hundreds of sequencing instruments which were necessary for parallel processing.[29] As the introduction of Sanger sequencing technology altered the status quo, the need for a 'democratization' of the field was made evident, with the concept of accessibility and affordability for all laboratories and clinical settings, becoming the primary objective.

The introduction of the novel NGS technologies with their massive parallel sequencing potential made this possible thus having a profound impact on research. The new sequencing methods share three major improvements. First of all bacterial cloning of DNA fragments is not required since they rely on the preparation of NGS libraries in a cell free environment. Secondly multiple sequencing reactions are carried out in parallel on a massive scale, resulting in an extraordinary increase in throughput and finally there is no need for electrophoresis for base interrogation.[29] These breakthroughs came with a substantial reduction in overall cost and the development of bench-top devices oriented towards smaller laboratories. To gain perspective concerning the decrease in question, bearing in mind the cost of the 'Human Genome Project', Illumina with its newer platform, is currently claiming to have reached the 1000$ per genome threshold.[29] However, the desire still lingers to further drop costs at an further exponential rate consistent with the semiconductor industry's Moore's Law.[30]

Nevertheless, these technologies are not without flaws and problems. Second generation devices to date continue to generate relatively short read lengths (40-450nt), although third generation platforms have overcome this obstacle with PacBio producing extremely long reads up to 20kb long. Furthermore,

the enormous increase in throughput has required significant advances in data handling, management and storage placing huge demands on bioinformatic tools and cloud storage solutions. Finally, as more and more of these bench top devices are acquired by labs, without specialized bioinformatics personnel, the need for development of simpler bioinformatic tools is also evident.

Another aspect of this industry however, is that it is in a constant volatile state. Since 2005, the launching year of the 454 platform, many things have changed and even more are expected to change. This decade has seen these new technologies spur, but has also witnessed their demise, as in the case of Helicos Biosciences which effectively went bankrupt and the 454 platform which will only be supported until mid-2016. In retrospect, over the past decade we have witnessed a second post-Sanger revolution and can anticipate more to come in the near future in this extra-competitive field.

Περίληψη

Τεχνολογίες αλληλούχισης DNA επόμενης γενιάς: μία ενδελεχής ανασκόπηση

Ευαγγελία Βούλγαρη, Νικόλαος Σπανάκης, Αθανάσιος Τσακρής
Εργαστήριο Μικροβιολογίας, Ιατρική Σχολή, Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών

Η τελευταία δεκαετία έχει υπάρξει καταλυτική στον τομέα των τεχνολογιών αλληλούχισης DNA, με την εισαγωγή νεοτέρων μηχανημάτων ικανών να πραγματοποιούν παράλληλη αλληλούχιση σε μαζική πλέον κλίμακα. Στόχος του άρθρου αυτού αποτελεί η περιεκτική απεικόνιση της συγκλονιστικής αυτής τεχνολογικής διαδικασίας ξεκινώντας από τις αρχικές μέρες της αλληλούχισης κατά Sanger και φτάνοντας μέχρι τις νεότερες πλατφόρμες δεύτερης και τρίτης γενιάς. Στην ανασκόπηση περιγράφονται ορισμένες από τις πιο ευρέως διαδεδομένες τεχνολογίες αλληλούχισης επόμενης γενιάς συγκριτικά με την αλληλούχιση κατά Sanger, ενώ επίσηςσυζητούνται και οι προκλήσεις που απορρέουν από αυτές τις καινούργιες τεχνολογικές εφαρμογές.

Λέξεις κλειδιά
τεχνολογίες αλληλούχισης επόμενης γενιάς, next generation sequencing.

References

1. Dahm R. Friedrich Miescher and the discovery of DNA. *Dev Biol* 2005; 278: 274–88.
2. Watson JD, Crick FH. The structure of DNA. *Cold Spring Harb Symp Quant Biol* 1953; 18: 123–31.
3. Kornberg A, Zimmerman SB, Kornberg SR, Josse J. Enzymatic synthesis of deoxyribonucleic acid influence of bacteriophage T2 on the synthetic pathway in host cells. *Proc Natl Acad Sci U S A* 1959; 45: 772–85.
4. Saiki RK, Bugawan TL, Horn GT, Mullis KB, Erlich HA. Analysis of enzymatically amplified beta-globin

and HLA-DQ alpha DNA with allele-specific oligonucleotide probes. *Nature* 1986; 324: 163–6.
5. Smith HO, Welcox KW. A Restriction enzyme from Hemophilus influenzae: I. Purification and general properties. *J Mol Biol* 1970; 51: 379–91.
6. Tiselius A. A new apparatus for electrophoretic analysis of colloidal mixtures. *Trans Faraday Soc* 1937; 33: 524–31.
7. Maxam AM, Gilbert W. A new method for sequencing DNA. *Proc Natl Acad Sci U S A* 1977; 74: 560–4.
8. Sanger F, Coulson AR. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol* 1975; 94: 441–8.
9. Hutchison CA. DNA sequencing: bench to bedside and beyond. *Nucleic Acids Res* 2007; 35: 6227–37.
10. Sanger F. The Croonian Lecture, 1975: Nucleotide Sequences in DNA. *Proc R Soc Lond B Biol Sci* 1975; 191: 317–33.
11. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 1977; 74: 5463–7.
12. Venter JC, Adams MD, Myers EW, et al. The Sequence of the Human Genome. *Science* 2001; 291: 1304–51.
13. Metzker ML. Sequencing technologies - the next generation. *Nat Rev Genet* 2010; 11: 31–46.
14. Dressman D, Yan H, Traverso G, Kinzler KW, Vogelstein B. Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. *Proc Natl Acad Sci U S A* 2003; 100: 8817–22.
15. Fedurco M, Romieu A, Williams S, Lawrence I, Turcatti G. BTA, a novel reagent for DNA attachment on glass and efficient generation of solid-phase amplified DNA colonies. *Nucleic Acids Res* 2006; 34: e22.
16. Mardis ER. Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet* 2008; 9: 387–402.
17. Harris TD, Buzby PR, Babcock H, et al. Single-molecule DNA sequencing of a viral genome. *Science* 2008; 320: 106–9.
18. Eid J, Fehr A, Gray J, et al. Real-time DNA sequencing from single polymerase molecules. *Science* 2009; 323: 133–8.
19. Schneider GF, Dekker C. DNA sequencing with nanopores. *Nat Biotechnol* 2012; 30: 326–8.
20. Margulies M, Egholm M, Altman WE, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 2005; 437: 376–80.
21. Ronaghi M, Uhlén M, Nyrén P. A sequencing method based on real-time pyrophosphate. *Science* 1998; 281: 363, 365.
22. Reuter JA, Spacek DV, Snyder MP. High-throughput sequencing technologies. *Mol Cell* 2015; 58: 586–97.
23. Rothberg JM, Hinz W, Rearick TM, et al. An integrated semiconductor device enabling non-optical genome sequencing. *Nature* 2011; 475: 348–52.
24. Morey M, Fernández-Marmiesse A, Castiñeiras D, Fraga JM, Couce ML, Cocho JA. A glimpse into past, present, and future DNA sequencing. *Mol Genet Metab* 2013; 110: 3–24.
25. Levene MJ, Korlach J, Turner SW, Foquet M, Craighead HG, Webb WW. Zero-Mode Waveguides for Single-Molecule Analysis at High Concentrations. *Science* 2003; 299: 682–6.
26. Korlach J, Bibillo A, Wegener J, et al. Long, processive enzymatic DNA synthesis using 100% dye-labeled terminal phosphate-linked nucleotides. *Nucleosides Nucleotides Nucleic Acids* 2008; 27: 1072–83.
27. Seeman NC. Nucleic acid junctions and lattices. *J Theor Biol* 1982; 99: 237–47.
28. Wang Y, Yang Q, Wang Z. The evolution of nanopore sequencing. *Genomic Assay Technol* 2015; 5: 449.
29. van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C. Ten years of next-generation sequencing technology. *Trends Genet TIG* 2014; 30: 418–26.
30. Moore GE. Cramming more components onto integrated circuits. *Electronics* 1965; 38: 114–7.