



The Arrival of Next-Generation Sequencing: An Overview of Current Technologies

3

Khalid El Bairi, Falak Azzam, and Mariam Amrani

Abstract

Over the past decade, substantial progress has been achieved in understanding the molecular mechanisms of carcinogenesis, and several hallmarks defining cancer have been established. These advances have markedly impacted translational research and clinical practice following the arrival of the next-generation sequencing (NGS) technology. This innovative revolution in molecular biology has enabled a rapid interrogation of the cancer genomes even using small quantities of nucleic acids. In this chapter, we describe the advantages and limitations of current NGS platforms including those using sequencing by synthesis, sequencing by ligation, and real-time sequencing, as well as their significant impact in molecular oncology.

Keywords

Next-generation sequencing · Sequencing by ligation · Sequencing by synthesis · Real-time sequencing · Cancer genomes

K. El Bairi (✉)

Cancer Biomarkers Working Group, Oujda, Morocco

e-mail: k.elbairi@ump.ac.ma

F. Azzam

Faculty of Sciences, Laboratory of Biochemistry and Immunology, Mohamed V University, Rabat, Morocco

M. Amrani

Faculty of Medicine and Pharmacy, Department of Pathology, National Institute of Oncology, Université Mohamed V, Rabat, Morocco

3.1 Introduction

Historically, after the discovery of the DNA structure in 1953 by molecular biologists James Watson and Francis Crick (Watson and Crick 1953) and its sequencing based on autoradiography visualization in 1977 by Sanger et al. (1977) and by Maxam and Gilbert (Maxam and Gilbert 1977), major advances in molecular biology have allowed a better structure and function elucidation of this “magic” molecule. Later, around the 1990s, the first slab gel-based sequencer [ABI PRISM[®] 3700 DNA Analyzer (Thermo Fisher Scientific, Waltham, MA, USA)] was made available in parallel with the launch of the human genome sequencing project which was declared complete in 2003. Since then, a rapid and extraordinary evolution of this area has allowed more sophisticated, scalable, faster, and cheaper technologies for genome sequencing with a significant increase in fees related to “big data” management based on bioinformatic pipelines and associated errors. Following this after-Sanger era, Roche 454’s pyrosequencing system was the first marketed of the NGS platforms launched in 2005 based on light detection of pyrophosphate release in addition to QIAGEN[®] PyroMark Q series (Margulies et al. 2005; Müllauer 2017; Harrington et al. 2013). Compared to the old classical sequencing methods, NGS enables a simultaneously and massively increased sequencing rate ranging from few gigabases per run to 6000 gigabases and therefore a possible human genome sequencing within 1 week with only 999 US dollars according to Veritas[®] genomic company (Müllauer, 2017; Goodwin et al. 2016; <https://www.veritasgenetics.com/why-are-we-here>). Current NGS is categorized into (1) systems that use sequencing by synthesis chemistry [Illumina[®] platforms (Illumina[®], San Diego, CA, USA), Ion Torrent[®] platforms (Thermo Fisher Scientific, Waltham, MA, USA), QIAGEN GeneReader[®] (QIAGEN, Hilden, Germany), Roche[®] Sequencing platforms (Roche, Pleasanton, CA, USA)] and (2) systems that use sequencing by ligation [SOLiD[®] (Thermo Fisher, Waltham, MA, USA) and BGISEQ-500[®] (BGI (MGI) Tech, Shenzhen, China)] allowing short-read sequencing approaches (for review, see Goodwin et al. 2016). On the other hand, further recent technologies [Pacific BioSciences[®] platforms (PACBIO[®], California, USA) and 10X Genomics[®] platforms (10X Genomics, Pleasanton, CA, USA)] enable a long read and real-time sequencing advantages (Goodwin et al. 2016). Interestingly, novel “lab-on-a-chip” technologies such as the freshly introduced IBM[®] DNA Transistor (IBM[®], Armonk, New York, USA) and Oxford Nanopore Technologies (MinION, PromethION, SmidgION platforms; Oxford Nanopore Technologies[®], Oxford Science Park, UK) are revolutionizing this field beyond the current next-generation sequencers and enable genome sequencing in real-time conditions (Yang and Jiang 2017; Lu et al. 2016). NGS ranges from the whole-genome sequencing analyzing the totality of human genome to targeted exome sequencing and finally to focused single genetic alteration assays. Most of NGS technologies are still for research use only, but recently, some platforms have been validated and gained approval by the FDA for marketing and routine laboratory use.

3.2 Sequencing by Synthesis Platforms

3.2.1 Pyrosequencing Systems (Roche[®] and QIAGEN[®] PyroMark)

Pyrosequencing principle (Fig. 3.1) is based on single-nucleotide addition methods that quantify the liberated inorganic pyrophosphate (PPi) after incorporation of a nucleic base using a cascade of enzymatic reactions that produces detectable bioluminescence signals (Metzker 2010; Ronaghi et al. 1998). Instead of Sanger sequencing which needs addition of complementary nucleotides all together at the same time into the reaction medium, pyrosequencing incorporates sequentially each known deoxyribonucleotide triphosphate (dNTP) in the elongation single-stranded amplicon by DNA polymerase. A PPi is therefore released and captured by an ATP sulfurylase to produce an ATP molecule which in turn is coupled to a luciferin to generate an oxyluciferin and light signals by luciferase-mediated conversion. An apyrase is added to the reaction wells to degrade the excess of dNTPs, and a camera called charge-coupled device (CCD) enables high-resolution and sensitive detection of generated signals. Of note, recorded light peaks and intensity are proportional to the number of incorporated nucleotides and reveal DNA sequences using different programs (Fig. 3.1b). Before performing pyrosequencing using Roche[®] 454 platform, template preparation and amplification are required using a microfluidic emulsion PCR (EmPCR) technology that has the advantage to avoid loss of DNA

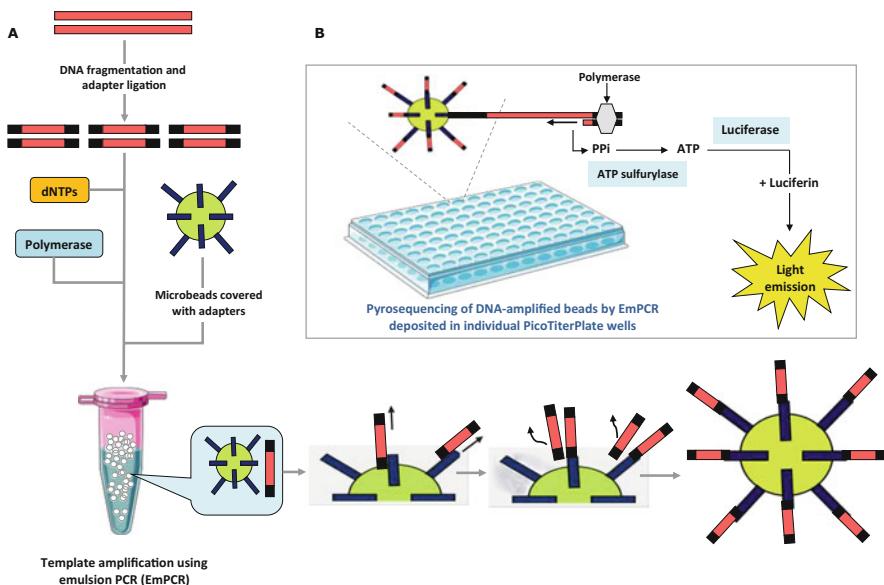


Fig. 3.1 Simplified diagram of (a) emulsion PCR and (b) pyrosequencing workflow. For comments, see text. *EmPCR* emulsion polymerase chain reaction, *DNA* deoxyribonucleic acid, *dNTPs* deoxynucleotides, *ATP* adenosine triphosphate, *PPi* pyrophosphate

sequences (for protocol review, see Kanagal-Shamanna 2016). In EmPCR (Fig. 3.1a), first, DNA templates are fragmented by sonication (or other methods), ligated to adapters and denatured followed by a capture in water-in-oil droplets. Each droplet contains DNA template with adapters, complementary adapters loaded on beads, primers, polymerase, and dNTPs. After amplification, millions of clonally amplified beads are placed and arrayed in PicoTiterPlate (PTP) microwells where massively parallel pyrosequencing reactions are performed (Metzker 2010; Goodwin et al. 2016). Despite their fast run times and improved read lengths (Roche[®] GS FLX Titanium and GS Junior), pyrosequencing machines had high error rates for sequencing homopolymer repeats and high reagent costs as well as difficulties in genome assembly. In 2013, Roche[®] discontinued its 454-based NGS platforms because of the arrival of highly competitive and coming of age technologies from Illumina[®] and Ion Torrent[®] (<https://www.fiercebitech.com/medical-devices/roche-to-close-454-life-sciences-as-it-reduces-gene-sequencing-focus>—accessed: 11/05/2018).

3.2.2 Illumina[®] Platforms

So far, Illumina is dominating the market of short-read NGS platforms as a result of its impressive high-throughput sequencing technology and low cost per base (van Dijk et al. 2014). The first NGS platform from Illumina (Genome Analyzer) was launched in 2006 by Solexa (acquired by Illumina one year later) allowing 1 gigabase/run (<https://emea.illumina.com/science/technology/next-generation-sequencing/illumina-sequencing-history.html>—accessed 18-05-2018). The foundation of Illumina instruments is based on sequencing by synthesis (base-by-base) technology using fluorescently labeled nucleotides (Fig. 3.2). In the first step, DNA is fragmented and ligated to adapters and bound to a solid support (glass flow cell) that contains immobilized primers (two types of oligos, forward and reverse) (Fig. 3.2a, b). The free end of DNA fragments interacts with close oligos, therefore creating bridges, and a clonal amplification PCR is used to generate the second strand. Finally, the bridge is denatured to form single-stranded DNA, the template is washed to remove reverse strands, and the process is repeated over again. In the second step, four differently labeled, fluorescent, and cleavable reversible terminator dNTPs (blockade of their 3'-OH group to prevent elongation) and DNA polymerase are added to the reaction (Guo et al. 2008; Goodwin et al. 2016). Every nucleotide is incorporated one by one into the elongating strand, unbound dNTPs are washed away, and a CCD camera is used to scan and identify which nucleotide is added and another cycle is repeated (Goodwin et al. 2016) (Fig. 3.2c). Illumina developed, refined, and optimized several NGS systems including MiniSeq series, MiSeq series, HiSeq series, HiSeq X series, NextSeq series, and the recently released NovaSeq 600 system that enable a tremendous increase in throughput and generate multiple terabases/run. Illumina MiSeq is designed as a personal sequencer with low run time and is adapted to small genomes. Illumina MiSeq seems to have superior position for metagenomic sequencing and molecular diagnostics laboratory. Moreover, Illumina

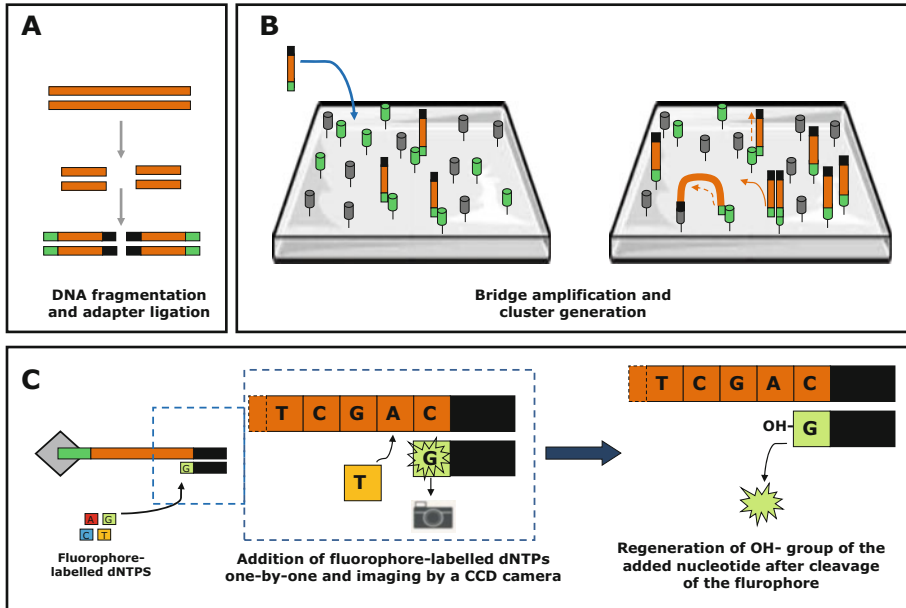


Fig. 3.2 Principle of Illumina sequencing: (a) template preparation, (b) amplification, and (c) sequencing. For comments, see text

HiSeq series are widely used for high-throughput applications such as large whole-genome sequencing and are more adapted to research use only. Substitution errors across Illumina platforms are the most frequent and are below 1%. In addition, Illumina technology has reduced homopolymer errors compared to other NGS systems using single-nucleotide addition strategies.

3.2.3 Thermo Fisher Ion Torrent[®] Platforms

Ion Torrent[®] systems share sequencing by synthesis strategy used by other platforms such as pyrosequencing and employ a unique pH-mediated non-optical sequencing (Rothberg et al. 2011). Similar to pyrosequencing, Ion Torrent[®] uses EmPCR to prepare templates (Fig. 3.1a). DNA-amplified beads are incubated in microwells where sequencing takes place. Nucleotides are added into the reaction one species at a time, and if the dNTP incorporated in the elongation strand is complementary, hydrogen ions (H^+) are released and induce pH changes which are detected by ion sensors [CMOS (complementary metal-oxide semiconductor) and ISFET (ion-sensitive field-effect transistor)] placed in the microwells and converted to voltage signals; the residual dNTPs are washed away and another cycle begins (Fig. 3.3). Basically, a voltage signal is proportional to the number of sequential dNTPs added to the elongating strand. Moreover, DNA templates may have homopolymer repeats; thus, multiple dNTPs are added in a distinct cycle and a strong

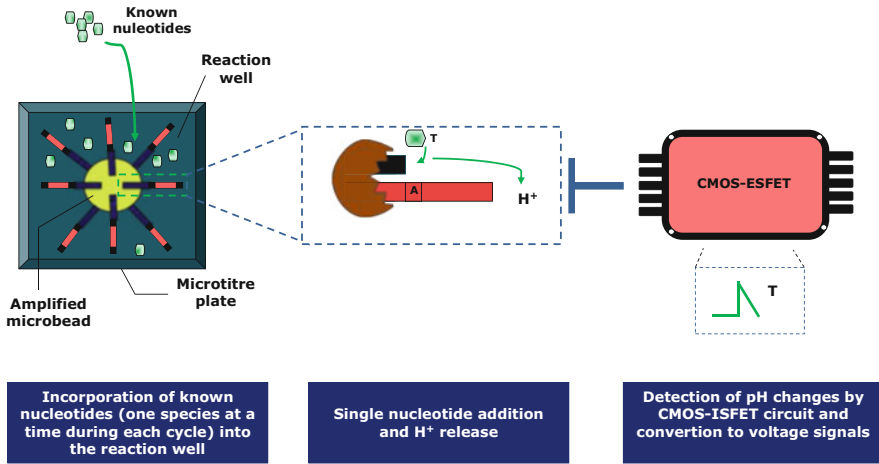


Fig. 3.3 Principle of Ion Torrent sequencing. For comments, see text. *CMOS* complementary metal-oxide semiconductor, *ISFET* ion-sensitive field-effect transistor

voltage signal is then detected which may limit the strength of this NGS by increasing the error rates (especially indels). However, this non-optical NGS has the advantage to distinguish between incorporated dNTPs during sequencing cycles and therefore enables fast runs and reduces reagents costs. Ion Torrent has marketed two platforms: Ion Torrent PGM which delivers 400 bp of read lengths and 2–7 h run time and Ion Proton system with a read length of 200 bp and a run time between 2 and 4 h. Ion Torrent PGM seems to be the best choice for affordable targeted sequencing panels (Lupini et al. 2015; Haley et al. 2015; Malapelle et al. 2015; Algars et al. 2017) compared to Ion Proton that is more practical for exome and transcriptomic sequencing (Brown et al. 2017).

3.2.4 QIAGEN[®] GeneReader

QIAGEN[®] introduced its all-in-one NGS system named GeneReader in 2015 (Karow 2015). The GeneReader was developed to perform all the sequencing steps from nucleic acid extraction and clonal amplification using the QIAcube system until data analysis and interpretation workflow. Template enrichment during the preparation phase uses EmPCR as the one used by Roche[®] pyrosequencing, SOLiD[®], and Ion Torrent[®] platforms. Typically, the GeneReader sequences incorporated fluorescent nucleotides by Illumina platforms and detects signals with imaging by TIRF (total internal reflection fluorescence) microscopy using laser channels (Goodwin et al. 2016) (Fig. 3.4). Sequencing of DNA from FFPE samples from CRC subjects using this NGS system was recently validated with reference to PCR, pyrosequencing, and Illumina MiSeq (Darwanto et al. 2017). Until this time, the GeneReader is intended for cancer clinical research use only.

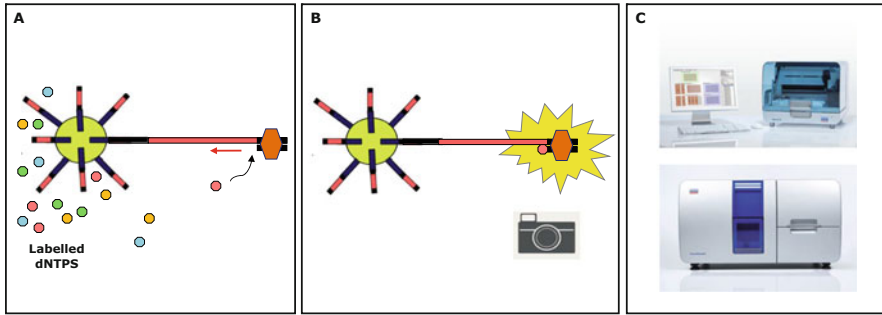


Fig. 3.4 Principle of Qiagen GeneReader platform: (a) addition of fluorophore-labeled dNTPs to hybridize with the complementary strand, (b) after the incorporation of fluorophore-labeled dNTP and the cleavage of the fluorophore to regenerate the OH group, the unit is imaged using four laser channels and another cycle begins, and (c) top: the QIAcube system, bottom: the Qiagen® GeneReader platform (reused with permission from Qiagen®)

3.3 Sequencing by Ligation Platforms

3.3.1 Thermo Fisher SOLiD®

SOLiD (Sequencing by Oligonucleotide Ligation and Detection) NGS system was launched by Applied Biosystems Inc. in 2007 (purchased later by Thermo Fisher®) and is based on the use of two-base color encoding and sequencing by ligation strategies (Goodwin et al. 2016; Valouev et al. 2008) allowing a maximum read length of 75 bp (Goodwin et al. 2016). Following DNA amplification generated by EmPCR, 3'-modified beads are deposited to be covalently attached in the surface of the flowchips (glass slides). In each flowchip, a sequence of bases (anchor primer) binds to the adapter and probes containing two first known labeled nucleotides attached to six other bases with a fluorophore hybridized to the strand template using a DNA ligase and the complex is imaged (Goodwin et al. 2016; Meldrum et al. 2011; Shendure et al. 2005) (Fig. 3.5). After this step, cleavage of the fluorophore is performed together with three bases of the probe, and another round of ligation, imaging, and cleavage is completed to recognize two out of every five nucleotides (probe extension). Finally, other sequencing cycles using this time progressive offset primers ($n - 1$, one base shifted) to decode the rest of the strand and therefore allowing an accurate double-sequencing strategy. However, substitution errors and difficulties in sequencing palindromic regions are the drawbacks of this technology (Huang et al. 2012). SOLiD short-read NGS platforms were discontinued as of May 1, 2016, and are no longer available for sale (https://www.thermofisher.com/content/dam/LifeTech/Documents/PDFs/5500_DiscontinuanceLetter_November2015.pdf—accessed 22-05-2018).

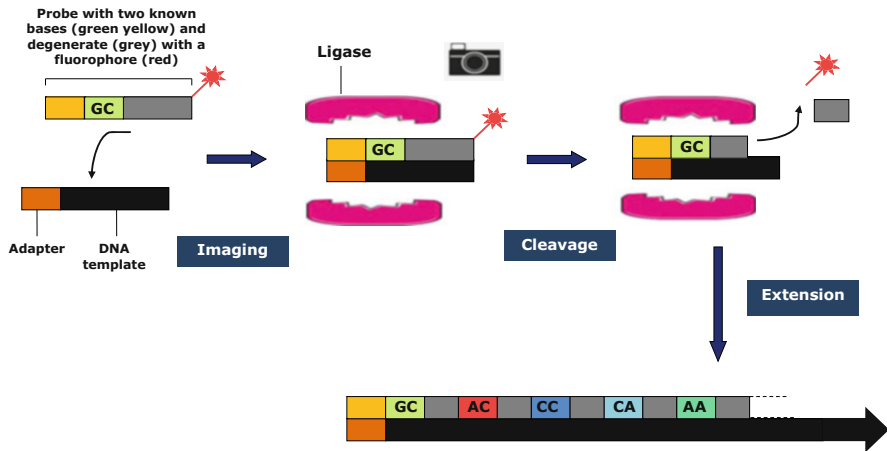


Fig. 3.5 Principle of SOLiD sequencing. For comments, see text

3.3.2 BGI Complete Genomics Platforms (BGISEQ-500[®] and BGISEQ-50[®])

BGISEQ sequencers are provided by the life sciences company “Complete Genomics” and use sequencing by ligation based on DNA nanoballs. In this technology, template preparation is performed utilizing a process called rolling circle amplification in which DNA undergoes repeated ligation, cleavage, and circularization (Goodwin et al. 2016) (Fig. 3.6a). After adapter ligation, template DNA is circularized and then cleaved downstream the adapter using endonucleases to bind other adapters in three additional cycles. Finally, the DNA is amplified to generate billions of circular structures that contain four adapters called nanoballs (Fig. 3.6b) to be deposited on sequencing flow cells (Goodwin et al. 2016; Drmanac et al. 2010). First, a complementary probe with single known base in addition to supplementary degenerate nucleotides and a fluorophore hybridize to the nanoball template via the sequences of the four ligated adapters. The complex is imaged and the probe is removed to enable hybridization of other new probes with another known base ($n + 1$) in other rounds of sequencing cycles (Goodwin et al. 2016) (Fig. 3.6c). The company claims to have 99.999% accuracy in sequencing complete human genomes with only \$600 (Drmanac et al. 2010; <https://www.bgi.com/us/human-whole-genome-sequencing-from-600>—accessed 27-05-2018). However, this technology is found to underrepresent AT-rich regions (Goodwin et al. 2016; Rieber et al. 2013). Using the BGISEQ-500 platform, some authors were able to show concordant results with Illumina HiSeq X10 in whole-genome sequencing of somatic and germline variants of pleural mesothelioma (Patch et al. 2018). Recently, a miniaturized and compacted desktop machine of BGISEQ-500 called BGISEQ-50 was released and designed for clinical sequencing laboratories with an output of 8 gigabases per run and a read length of 50 bp (<https://www.genomeweb.com/>

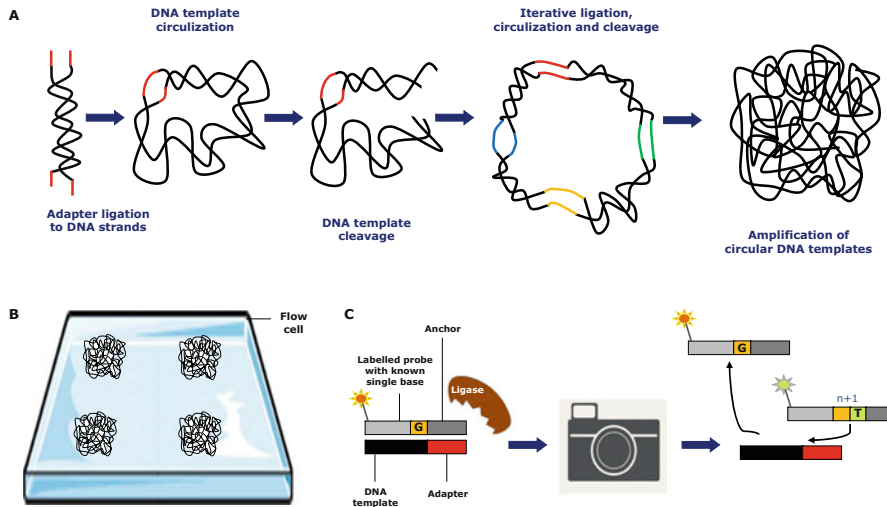


Fig. 3.6 Principle of BGI Complete Genomics sequencing platforms: (a) template preparation, (b) immobilization of amplified DNA templates (known as nanoballs) on flow cells and hybridization, (c) hybridization of single-base probe to DNA template (nanoball) followed by imaging of the whole complex to identify the labeled base, removal of anchor-probe, and a new process begins with a new base ($n + 1$ position). For additional comments, see text

[sequencing/bgi-launches-new-desktop-sequencer-china-registers-larger-version-cfda#.WwsyczTRB0w](#)—accessed 27-05-2018).

3.4 Real-Time Sequencing Platforms

The advent of single-molecule real-time sequencing technology used by Pacific BioSciences[®] and Oxford Nanopore[®] is based on considerably longer read generation of data without interruptions between read steps compared to the previously discussed technologies which produce short-read sequences (Goodwin et al. 2016; Bleidorn 2017).

3.4.1 Pacific BioSciences[®] (PacBio) Platforms

In PacBio technology, template preparation avoids clonal amplification by using direct sequencing of modified DNA (Rhoads and Au 2015). DNA templates are ligated to two hairpin barcoded adapters (Fig. 3.7a) followed by a removal of templates with inadequate size using a selection process (Goodwin et al. 2016). Templates and fluorescently labeled dNTPs are then deposited in picoliter wells called zero-mode waveguide cells containing each single DNA polymerase immobilized at the bottom that can bind the hairpin adapters (Rhoads and Au

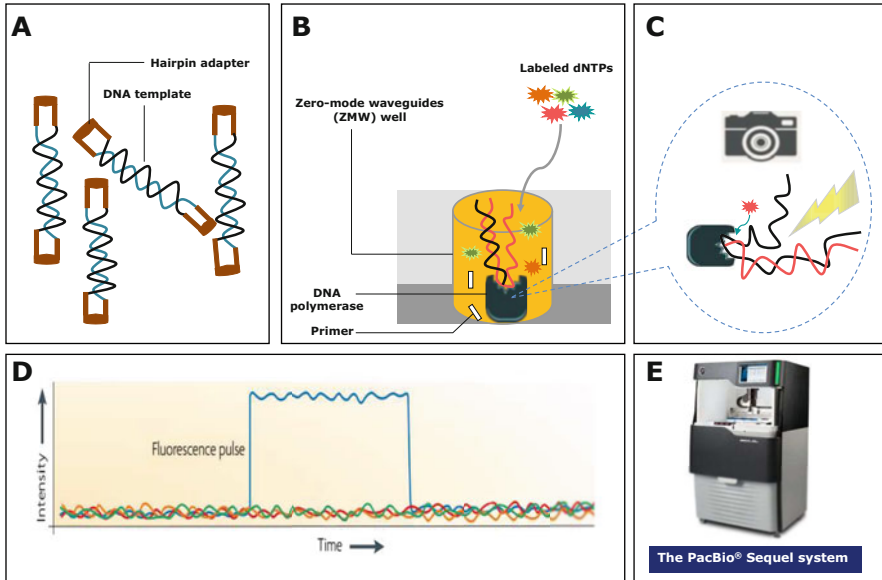


Fig. 3.7 Principle of PacBio sequencing platform: (a) template preparation (ligation of hairpin adapters), (b, c) addition of prepared template into the zero-mode waveguide cells where real-time sequencing takes place, (d) example of a recorded fluorescence pulse (reprinted from Nat Rev Genet, 11, Metzker ML, Sequencing technologies-the next generation, 31–46, Copyright (2010), with permission from Springer Nature), (e) the recently launched PacBio Sequel system (reused with permission from Pacific Biosciences®). For comments, see text. *DNA* deoxyribonucleic acid, *dNTPs* deoxynucleotides, *ZMW* zero-mode waveguides

2015) (Fig. 3.7b, c). Resulting light pulses (Fig. 3.7d) corresponding to the colors emitted by the incorporated tagged nucleotides during amplification are detected and visualized using a camera and matched tags are cleaved off (Rhoads and Au 2015). With a great long read length estimated at ~20 Kb, PacBio RS II platform is the most commonly used for this purpose, and it seems to be the gold standard for de novo assembly of genome projects (Giordano et al. 2017; Goodwin et al. 2016; Gordon et al. 2016). However, this system is dominated by random indel errors, and their cost per gigabase is still high (Goodwin et al. 2016). To improve these drawbacks, PacBio has recently launched the PacBio Sequel system (Fig. 3.7e) that significantly ameliorated the sequencing throughput (~7× that of PacBio RS II) (Goodwin et al. 2016).

3.4.2 Oxford Nanopore Technologies® Platforms

Oxford Nanopore Technologies® (ONT) is a rising star in real-time sequencing using pocket-sized devices. Compared to the other platforms that detect secondary signals (pH changes, light emission, or color) revealing the composition of DNA, the

technology behind these long-read sequencers directly sequences DNA fragments during their passage through a biological protein nanopore fixed on a microwell (Goodwin et al. 2016; Clarke et al. 2009). Before sequencing, DNA is fragmented (8–10 kb) and ligated to two different adapters to form a leader-hairpin structure, a desired conformation that increases the interaction between the DNA and the α -hemolysin pore and facilitates its passage using a motor protein (Goodwin et al. 2016). Once the DNA is translocated through the pore, a characteristic disruption in the electric current is detected and enables a discrimination of nucleotides in question (Fig. 3.8a). In 2014, the company released its first attracting super-portable platform known as MinION (Fig. 3.8b) only with a price of \$900, and able to sequence ~70 bp/s and adapted to personal laptops (Yang and Jiang 2017; Goodwin et al. 2016). Following its successful development, the company marketed two other multiple sequencing devices known as PromethION and GridION with up to 5–48 flow cells, respectively, which have increased dramatically its throughput (<https://nanoporetech.com/how-it-works>—accessed 04-06-2018). Very recently, the company has developed the VoITRAX, a small USB-powered manual device designed for automated library preparation without the need of a molecular biology laboratory and skilled sequencing teams (Fig. 3.8b). Moreover, another device called SmidgION for smallest sequencing purposes is being developed to be adapted for smartphone-based sequencing and will be launched soon. Importantly, Minervini et al. assessed *TP53* mutations in chronic lymphocytic leukemia by nanopore

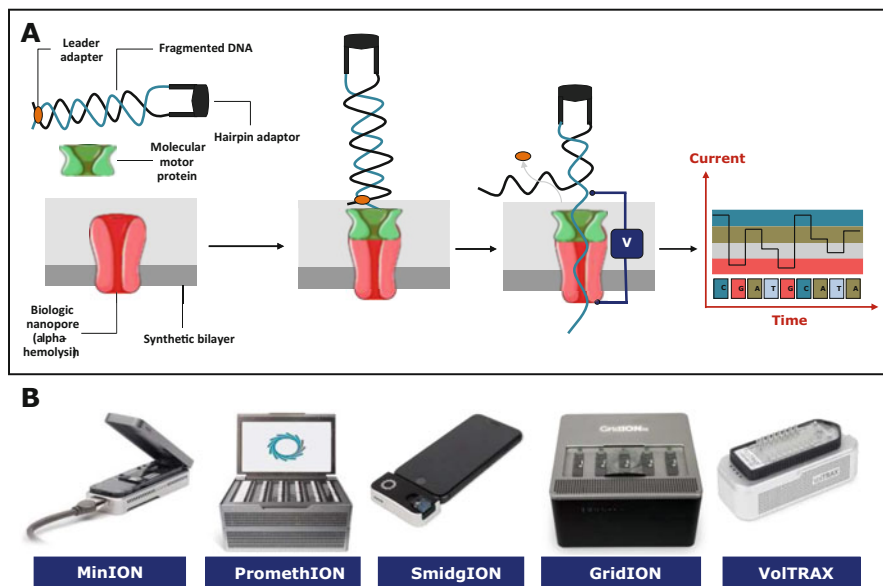


Fig. 3.8 Principle of Oxford Nanopore sequencing: (a) summary of platforms sequencing principle, (b) pocket-sized devices developed recently by the company. For comments, see text

MinION and showed correlation, more sensitivity, and less expensiveness compared to Sanger sequencing (Minervini et al. 2016). However, despite these impressive advances, this nanopore sequencing is still suffering from high indel errors (other emerging sequencing technologies are listed in Table 3.1).

Table 3.1 Other emerging next-generation sequencing technologies

Company	NGS platform	Principle	Website
10X Genomics ^a	Chromium and GemCode systems	Synthetic long-read and emulsion-based sequencing	https://www.10xgenomics.com/
Direct Genomics	GenoCare	Single-molecule direct sequencing using TIRF ^b imaging for parallel detection of multiple fluorescently tagged single molecules	http://www.directgenomics.com/
Bionano Genomics	Saphyr and Irys systems	High-resolution imaging of a linearized and labeled DNA in nanochannels	https://bionanogenomics.com/
NanoString Technologies	Hyb and Seq	Library-free, amplification-free, single-molecule direct sequencing using cyclic DNA hybridization of fluorescent molecular barcodes	https://www.nanostring.com/
GnuBio (Bio-Rad)	GnuBIO platform	Droplet microfluidics-based sequencing	http://gnubio.com/
Genia (Roche)	In development	Single-molecule semiconductor-based DNA sequencing using nanopore technology	https://sequencing.roche.com/en/technology-research/technology/nanopore-sequencing.html
GenapSys	Genius	Electronic DNA sequencing	http://www.genapsys.com/
Electron Optica	In development	Electron microscopy-based sequencing ^c	http://www.electronoptica.com/Electron_optica/HOME.html
IBM ^d	The DNA transistor (in development)	Nanopore-based sequencing	http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/dnatransistor/
NABsys	In development	Solid-state nanodetectors-based sequencing	http://nabsys.com/
Electronic BioSciences	In development	Nanopore-based sequencing	http://electronicbio.com/

^aRelated publications can be found at: [10xgenomics.com/resources](https://www.10xgenomics.com/resources)

^bTotal internal reflection fluorescence

^cDetails can be found at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4117835/> and <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0154707>

^dIn collaboration with Roche

3.5 Conclusion

In conclusion, according to the current literature, the patents, and approvals for marketing, Illumina and Ion Torrent platforms seem to be the best mature sequencing devices to be used for clinical laboratory practice. Moreover, they are the most utilized for analyzing CRC genomics (see the next chapter for details; educative videos about NGS technologies can be found in Box 3.1). For further reading and useful websites, see Box 3.2.

Box 3.1 Useful Links and Educative Videos About Next-Generation Sequencing Platforms and Technologies

Sequencing company	Website	Links for educative videos
Pyrosequencing (Roche®)	Discontinued	https://www.youtube.com/watch?v=KzdWZ5ryBIA
QIAGEN® PyroMark	https://www.qiagen.com/us/	https://www.youtube.com/watch?v=bNKEhOGvcaI https://www.youtube.com/channel/UCPXwu_KIrSKWMilWgiQuVaw https://www.jove.com/video/50405/pyrosequencing-for-microbial-identification-and-characterization
Illumina® platforms	https://www.illumina.com/	https://emea.illumina.com/science/technology/next-generation-sequencing/sequencing-technology.html https://sapac.illumina.com/company/video-hub/view-all-videos.html https://www.youtube.com/user/IlluminaInc
Thermo Fisher Ion Torrent®	https://www.thermofisher.com/ma/en/home/life-science/sequencing/next-generation-sequencing/ion-torrent-next-generation-sequencing-technology.html	https://www.youtube.com/watch?v=WYBzbxIfuKs
QIAGEN® GeneReader	https://www.qiagen.com/us/	https://www.youtube.com/watch?v=HQhw5Ihp8IA

(continued)

Box 3.1 (continued)

Thermo Fisher SOLiD [®]	https://www.thermofisher.com/ma/en/home/life-science/sequencing/next-generation-sequencing/solid-next-generation-sequencing/solid-next-generation-sequencing-systems-reagents-accessories.html	https://www.thermofisher.com/ma/en/home/life-science/sequencing/next-generation-sequencing/solid-next-generation-sequencing/solid-next-generation-sequencing-systems-reagents-accessories.html https://www.youtube.com/watch?v=YLT-DUeaLms
BGI Complete Genomics platforms	http://www.seq500.com/en/	http://www.seq500.com/en/portal/videos.shtml
Pacific BioSciences [®] (PacBio) platforms	https://www.pacb.com/	https://www.pacb.com/smrt-science/smrt-resources/video-gallery/
Oxford Nanopore Technologies [®] platforms	https://nanoporetech.com/	https://nanoporetech.com/resource-centre/videos https://www.youtube.com/channel/UC5yMIYjHSgFfZ37LYq-dzig

Additional videos about NGS can be found in JoVE (the *Journal of Visualized Experiments*): <https://www.jove.com/>

Box 3.2 Useful Bioinformatic Tools, Websites, and Databases

GeneCards [®] : The Human Gene Database	http://www.genecards.org/
Online Mendelian Inheritance in Man [®] (OMIM) database	https://www.omim.org/
The Cancer Genome Atlas Clinical Explorer ^a	http://genomeportal.stanford.edu/pan-tcga
The Catalogue Of Somatic Mutations In Cancer (COSMIC)	https://cancer.sanger.ac.uk/cosmic
Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer	https://cgap.nci.nih.gov/Chromosomes/Mitelman
Sequence Variant Nomenclature	http://varnomen.hgvs.org/
ClinVar database ^b	https://www.ncbi.nlm.nih.gov/clinvar/
Variant Annotation and Filter Tool ^c	http://varaft.eu/
PharmGKB [®]	https://www.pharmgkb.org/
GenomeWeb ^d	https://www.genomeweb.com
Cochrane Library	http://www.cochranelibrary.com/

(continued)

Box 3.2 (continued)

The U.S. National Library of Medicine clinical trials database	https://www.clinicaltrials.gov/
The Human Gene Mutation Database (HGMD [®])	http://www.hgmd.cf.ac.uk/ac/index.php
Guidelines for diagnostic next-generation sequencing	http://www.irdirc.org/guidelines-for-diagnostic-next-generation-sequencing/
The International Society for Gastrointestinal Hereditary Tumours (InSiGHT)	https://www.insight-group.org/
ASCO guidelines for molecular testing in colorectal cancer	https://www.asco.org/practice-guidelines/quality-guidelines/guidelines/gastrointestinal-cancer#/15831
Educative videos about genomics	https://www.yourgenome.org/video
Colorectal Cancer Atlas	http://www.colonatlas.org/
CoReCG ^c	http://lms.snu.edu.in/corecg/
CBD: a biomarker database for colorectal cancer	http://sysbio.suda.edu.cn/CBD/
Colon Cancer Alliance	https://www.ccalliance.org/
The Human Pathology Atlas	http://www.proteinatlas.org/humanpathology/
The Cancer Genome Atlas	https://cancergenome.nih.gov/
IGSR: The International Genome Sample Resource	http://www.internationalgenome.org/

^aA web and mobile interface for identifying clinical–genomic driver associations

^bA database about genomic variations and their relationship to human health

^cDetails can be found in: <https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gky471/5025894>

^dAn online news website focusing on genomics and emerging technologies

^eA comprehensive database of genes associated with colon-rectal cancer

Authors' Contribution KE and FA wrote the chapter and MA supervised the writing process.

References

- Algars A, Sundström J, Lintunen M et al (2017) EGFR gene copy number predicts response to anti-EGFR treatment in RAS wild type and RAS/BRAF/PIK3CA wild type metastatic colorectal cancer. *Int J Cancer* 140(4):922–929. <https://doi.org/10.1002/ijc.30507>
- Bleidorn C (2017) Sequencing techniques. In: *Phylogenomics*. Springer, Cham, pp 43–60. https://doi.org/10.1007/978-3-319-54064-1_3
- Brown RB, Madrid NJ, Suzuki H, Ness SA (2017) Optimized approach for ion proton RNA sequencing reveals details of RNA splicing and editing features of the transcriptome. *PLoS One* 12(5):e0176675. <https://doi.org/10.1371/journal.pone.0176675>
- Clarke J, Wu HC, Jayasinghe L et al (2009) Continuous base identification for single-molecule nanopore DNA sequencing. *Nat Nanotechnol* 4(4):265–270. <https://doi.org/10.1038/nnano.2009.12>

- Darwanto A, Hein AM, Strauss S et al (2017) Use of the QIAGEN GeneReader NGS system for detection of KRAS mutations, validated by the QIAGEN Therascreen PCR kit and alternative NGS platform. *BMC Cancer* 17(1):358. <https://doi.org/10.1186/s12885-017-3328-z>
- Drmanac R, Sparks AB, Callow MJ et al (2010) Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science* 327(5961):78–81. <https://doi.org/10.1126/science.1181498>
- Giordano F, Aigrain L, Quail MA et al (2017) De novo yeast genome assemblies from MinION, PacBio and MiSeq platforms. *Sci Rep* 7:3935. <https://doi.org/10.1038/s41598-017-03996-z>
- Goodwin S, McPherson JD, McCombie WR (2016) Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet* 17(6):333–351. <https://doi.org/10.1038/nrg.2016.49>
- Gordon D, Huddleston J, Chaisson MJP et al (2016) Long-read sequence assembly of the gorilla genome. *Science* 352(6281):aae0344. <https://doi.org/10.1126/science.aae0344>
- Guo J, Xu N, Li Z et al (2008) Four-color DNA sequencing with 3'-O-modified nucleotide reversible terminators and chemically cleavable fluorescent dideoxynucleotides. *Proc Natl Acad Sci USA* 105(27):9145–9150. <https://doi.org/10.1073/pnas.0804023105>
- Haley L, Tseng LH, Zheng G et al (2015) Performance characteristics of next-generation sequencing in clinical mutation detection of colorectal cancers. *Mod Pathol* 28(10):1390–1399. <https://doi.org/10.1038/modpathol.2015.86>
- Harrington CT, Lin EI, Olson MT et al (2013) Fundamentals of pyrosequencing. *Arch Pathol Lab Med* 137(9):1296–1303. <https://doi.org/10.5858/arpa.2012-0463-RA>
- <https://emea.illumina.com/science/technology/next-generation-sequencing/illumina-sequencing-history.html>. Accessed 18 May 2018
- <https://www.bgi.com/us/human-whole-genome-sequencing-from-600>. Accessed 27 May 2018
- <https://www.genomeweb.com/sequencing/bgi-launches-new-desktop-sequencer-china-registers-larger-version-cfda#WwsyczTRB0w>. Accessed 27 May 2018
- <https://nanoporetech.com/how-it-works>. Accessed 04 June 2018
- <https://www.fiercebitech.com/medical-devices/roche-to-close-454-life-sciences-as-it-reduces-gene-sequencing-focus>. Accessed 11 May 2018
- https://www.thermofisher.com/content/dam/LifeTech/Documents/PDFs/5500_DiscontinuanceLetter_November2015.pdf. Accessed 22 May 2018
- <https://www.veritasgenetics.com/why-are-we-here>. Accessed 20 Jan 2018
- Huang Y-F, Chen S-C, Chiang Y-S, Chen T-H, Chiu K-P (2012) Palindromic sequence impedes sequencing-by-ligation mechanism. *BMC Syst Biol* 6(Suppl 2):S10. <https://doi.org/10.1186/1752-0509-6-S2-S10>
- Kanagal-Shamanna R (2016) Emulsion PCR: techniques and applications. *Methods Mol Biol* 1392:33–42. https://doi.org/10.1007/978-1-4939-3360-0_4
- Karow J (2015) Qiagen launches GeneReader NGS System at AMP; presents performance evaluation by broad. *GenomeWeb*. <https://www.genomeweb.com/molecular-diagnostics/qiagen-launches-geneaderngs-system-amp-presents-performance-evaluation>. Accessed 4 Nov 2015
- Lu H, Giordano F, Ning Z (2016) Oxford nanopore MinION sequencing and genome assembly. *Genomics Proteomics Bioinformatics* 14(5):265–279. <https://doi.org/10.1016/j.gpb.2016.05.004>
- Lupini L, Bassi C, Mlcochova J et al (2015) Prediction of response to anti-EGFR antibody-based therapies by multigene sequencing in colorectal cancer patients. *BMC Cancer* 15:808. <https://doi.org/10.1186/s12885-015-1752-5>
- Malapelle U, Vigliar E, Sgariglia R et al (2015) Ion torrent next-generation sequencing for routine identification of clinically relevant mutations in colorectal cancer patients. *J Clin Pathol* 68(1):64–68. <https://doi.org/10.1136/jclinpath-2014-202691>
- Margulies M, Egholm M, Altman WE et al (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437(7057):376–380. <https://doi.org/10.1038/nature03959>
- Maxam AM, Gilbert W (1977) A new method for sequencing DNA. *Proc Natl Acad Sci USA* 74(2):560–564

- Meldrum C, Doyle MA, Tothill RW (2011) Next-generation sequencing for cancer diagnostics: a practical perspective. *Clin Biochem Rev* 32(4):177–195
- Metzker ML (2010) Sequencing technologies – the next generation. *Nat Rev Genet* 11(1):31–46. <https://doi.org/10.1038/nrg2626>
- Minervini CF, Cumbo C, Orsini P et al (2016) TP53 gene mutation analysis in chronic lymphocytic leukemia by nanopore MinION sequencing. *Diagn Pathol* 11(1):96
- Müllauer L (2017) Next generation sequencing: clinical applications in solid tumours. *Memo* 10(4):244–247. <https://doi.org/10.1007/s12254-017-0361-1>
- Patch A-M, Nones K, Kazakoff SH et al (2018) Germline and somatic variant identification using BGISEQ-500 and HiSeq X ten whole genome sequencing. *PLoS One* 13(1):e0190264. <https://doi.org/10.1371/journal.pone.0190264>
- Rhoads A, Au KF (2015) PacBio sequencing and its applications. *Genomics Proteomics Bioinformatics* 13(5):278–289. <https://doi.org/10.1016/j.gpb.2015.08.002>
- Rieber N, Zapatka M, Lasitschka B et al (2013) Coverage bias and sensitivity of variant calling for four whole-genome sequencing technologies. *PLoS One* 8(6):e66621. <https://doi.org/10.1371/journal.pone.0066621>
- Ronaghi M, Uhlén M, Nyrén P (1998) A sequencing method based on real-time pyrophosphate. *Science* 281(5375):363,365
- Rothberg JM, Hinz W, Rearick TM et al (2011) An integrated semiconductor device enabling non-optical genome sequencing. *Nature* 475(7356):348–352. <https://doi.org/10.1038/nature10242>
- Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74(12):5463–5467
- Shendure J, Porreca GJ, Reppas NB et al (2005) Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* 309(5741):1728–1732
- Valouev A, Ichikawa J, Tonthat T et al (2008) A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. *Genome Res* 18(7):1051–1063. <https://doi.org/10.1101/gr.076463.108>
- van Dijk EL, Auger H, Jaszczyszyn Y et al (2014) Ten years of next-generation sequencing technology. *Trends Genet* 30(9):418–426. <https://doi.org/10.1016/j.tig.2014.07.001>
- Watson JD, Crick FH (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* 171(4356):737–738
- Yang N, Jiang X (2017) Nanocarbons for DNA sequencing: a review. *Carbon* 115:293–311. <https://doi.org/10.1016/j.carbon.2017.01.012>