

Editorial on the 20th Anniversary of the Genome Project Realization. The History of DNA Sequencing

Pier Paolo Piccaluga^{1,2,*}

¹ Department of Medical and Surgical Sciences, Bologna University School of Medicine, Bologna, Italy

² Biobank of Research, IRCCS Azienda Ospedaliera-Universitaria di Bologna, Bologna, Italy

*Correspondence: E-mail: pierpaolo.piccaluga@unibo.it

“... [A] knowledge of sequences could contribute much to our understanding of living matter.”

[Frederick Sanger [1]]

The order of nucleotides inside a DNA sequence is critical as it serves as the basis for the “Genetic Code” or “Code of Life”. It supports subsequent protein function and, in the case of mutations, disease manifestation. Identifying the order of nucleotides in biological samples is essential for various research applications.

This year, we celebrated the twentieth anniversary of the conclusion of the Human Genome Project, which aimed to sequence, for the first time, the complete human genome. The initiative followed the *Alta Summit Conference* held in 1984 by the Department of Energy and the International Commission for Protection Against Environmental Mutagens and Carcinogens. In that occasion it was discussed whether it could be possible to use new DNA sequencing methods to directly detect mutations and detect any increase in the mutation rate among survivors of the Hiroshima and Nagasaki bombings. In the subsequent year, Charles DeLisi, Robert Sinsheimer, and Renato Dulbecco individually proposed the concept of establishing a collaborative initiative to sequence the human genome with the aim of identifying genes linked to certain diseases [2, 3].

This unprecedented breakthrough in human genetics and biology spanned 13 years, starting in 1990 and ending in 2003 (International Human Genome Sequencing Consortium [4]). In our current era, after only two decades, we can realize the process of sequencing a complete genome in approximately five hours, at 1/1000 of the cost.

The achievement of such efficiency took, however, more than a century since Friedrich Mietscher first isolated DNA in 1869 [5]. Between 1871 and 1929, scientists including the same Friedrich Miescher, Walther Flemming, Albrecht Kossel, and Phoebus Levene established the fundamental principles of cell chemistry and

Citation

Pier Paolo Piccaluga (2023), Editorial on the 20th Anniversary of the Genome Project Realization. The History of DNA Sequencing. *Digital Medicine and Healthcare Technology* 2(1), 1–13.

DOI

<https://doi.org/10.5772/dmht.21>

Copyright

© The Author(s) 2023.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Published: 28 December 2023

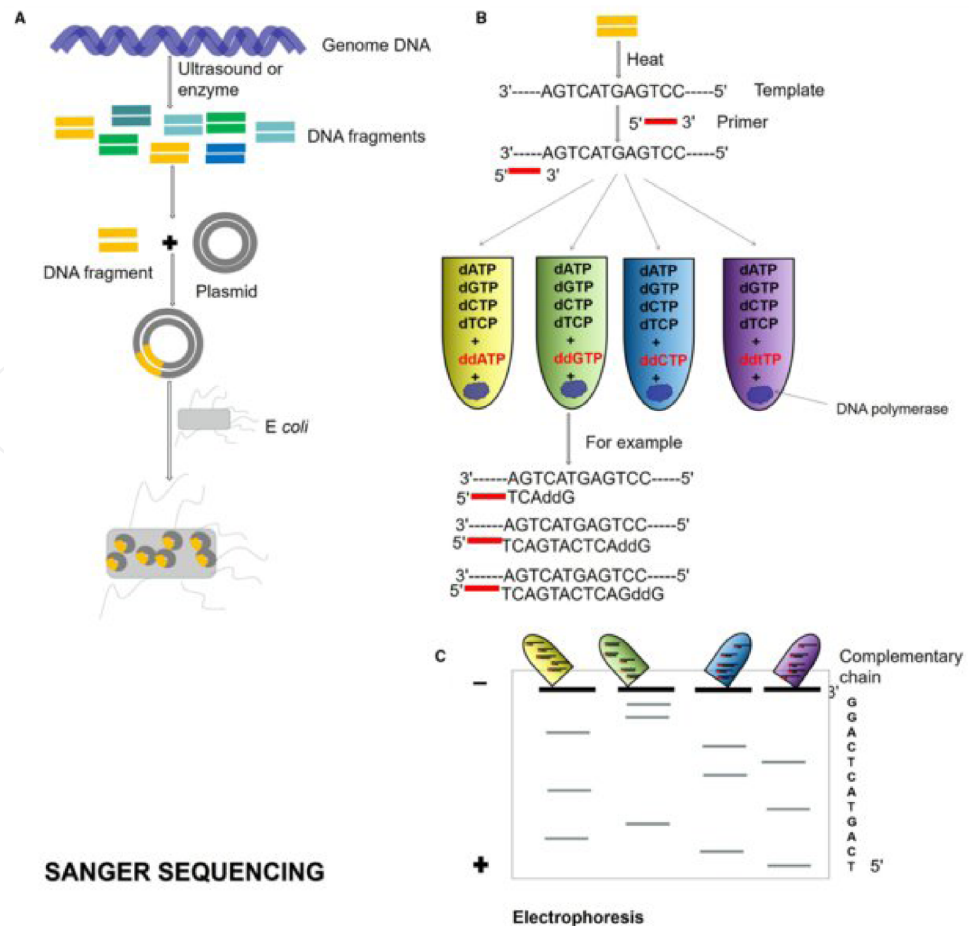
nucleic molecules, which facilitated the development of DNA sequencing. By 1929, it was determined that nucleic acids consisted of five nitrogen-containing bases: adenine (A), cytosine (C), guanine (G), thymine (T), and uracil (U). The DNA backbone is composed of alternating sugar and phosphate groups, with each sugar molecule being connected to a nitrogen-containing base, which can be either A, T, C, or G [2]. At that time, nonetheless, DNA was seen as a structurally insignificant molecule; proteins, by contrast, were believed to constitute the “genetic” material, due to their link with chromosomes and their intricate structure. In 1944, Oswald Avery, Colin MacLeod, and Maclyn McCarty published a paper that provided evidence showing that DNA, rather than proteins, had the ability to alter the characteristics of cells [6]. Based on this new evidence, the ongoing DNA research rapidly grew. In 1952, Erwin Chargraff and Linus Pauling made significant discoveries regarding the molecular characteristics of DNA and the nature of the chemical links that connect its molecules [7]. Chargraff’s findings established that the quantity of guanine (G) should be equivalent to the quantity of cytosine (C) in any living organism, while the quantity of adenine (A) should be equal to the quantity of thymine (T).

In 1953, James Watson and Francis Crick determined the 3D structure of DNA (“*Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid*”) [8] using crystallographic data given by Rosalind Franklin and Maurice Wilkins. However, the ability to read the DNA sequence took some years to develop. The approaches, based analytical chemistry, could only measure nucleotide content, not order.

1. First generation sequencing technologies

In 1965, Robert Holley and colleagues created the first complete nucleic acid of alanine tRNA from *Saccharomyces cerevisiae*, establishing solid bases for future DNA sequencing [9]. Meanwhile, Fred Sanger and colleagues developed a comparable approach that detects radiolabeled partial-digestion fragments [10]. Later, Sanger and Coulson proposed the “plus and minus” method [11]. Eventually, in 1977, Sanger, Nicklen and Coulson improved their method for DNA sequencing, which later became the widely accepted procedure, mixing DNA polymerase with a mixture of standard and chain-terminating nucleotides (ddNTPs) [12, 13]. When dNTPs and ddNTPs are mixed, unpredictable early termination of sequencing reactions occurs during PCR. Four parallel reactions are carried out, each containing one variant of a chain-terminating nucleotide. The sequence can be read off base by base by seeing this process using gel electrophoresis (Figure 1) [14]. This technique was groundbreaking at the time.

Int



SANGER SEQUENCING

Figure 1. Sanger sequencing process. (A) Genome DNA is cut into different DNA fragments by ultrasound or enzyme, connected into a plasmid, and transformed into Escherichia coli. (B) The Dideoxy chain termination method is used for sequencing. The reaction system is a mixture of template, primer, DNA polymerase, and dNTPs. Four kinds of ddNTP (ddATP, ddTTP, ddCTP, and ddGTP) with radioisotope marks are, respectively, added to four reaction systems. If ddNTP is incorporated into the oligonucleotide chain, the 2' and 3' of ddNTP do not contain hydroxyl groups and cannot form phosphodiester bonds, then the extension of the DNA chain is terminated. (C) Four PCR reaction system products are added to gel electrophoresis, and autoradiography was used to determine the DNA sequence according to the location of electrophoresis. (Adapted from Zhang L, Chen F, Zeng Z, Xu M, Sun F, Yang L, Bi X, Lin Y, Gao Y, Hao H, Yi W, Li M and Xie Y. *Advances in metagenomics and its application in environmental microorganisms. Front Microbiol.* 2021;12: 766364.)

Closely, Allan Maxam and Walter Gilbert published their chemical cleavage approach, which became the first widely used method for DNA sequencing for some time [15].

Originally, in Sanger method, the dideoxy terminator bases were labeled with ^{32}P using radioactivity. This included carrying out four distinct processes. However, the current method involves using modified bases that are tagged with one of four fluorescent markers. Applied Biosystems then commercialized and mechanized the process by 1987.

The devices designed for executing this DNA sequencing technique would generate reads that are less than one kilobase (typically, 500–1000 bp) [16]. To investigate extended segments of DNA, researchers can adopt shotgun sequencing technology. This technique revolves around fragmenting the DNA into multiple pieces, sequencing each fragment, and subsequently utilizing a computer software created by Rodger Staden to accurately align the resulting collection of gel readings into a single continuous sequence, commonly referred to as a *contig*. By this approach, in 1982, Sanger and his colleagues successfully obtained the whole DNA sequence of a bacteriophage, consisting of 48,502 base pairs (bp) [17]. This achievement coincided with the public availability of GenBank, the US (United States) National Institutes of Health's sequencing database.

Sanger's initial approach had been soon mechanized. This resulted in replacing large slab gels with acrylic-finer capillaries that might be seen on an electropherogram. This technology was critical to the Human Genome Project's completion in 2003. However, even after the Human Genome Project, the cost of capillary electrophoresis remained prohibitively expensive for large-scale sequencing programs. By the mid-2000s, there were various efforts seeking to reduce the costs of sequencing, and, across the globe, laboratories were putting novel methodologies and techniques for higher-throughput screening to the test.

2. High throughput “next generation” sequencing

The second generation of DNA sequencing arose after the introduction of pyrosequencers in 1996 [18]. The approach relied on sequencing-by-synthesis (SBS) technology, which involves the step-by-step incorporation and identification of individual nucleotides, quantifying the light produced during pyrophosphate synthesis. By detecting pyrophosphate generation as each nucleotide passed through the device, this method was utilized to infer sequences [18–20]. Pyrosequencing, introduced by Pal Nyérén's group, had several advantages, including the use of natural nucleotides (rather than the heavily modified dNTPs employed in chain termination), and could be monitored in real time [18–20].

Pyrosequencing was licensed to 454 Life Sciences (later acquired by Roche), where it became the first major commercially viable NGS (Next Generation Sequencing) technique. Water-in-oil emulsion PCR was used to attach individual

fragments (libraries) of DNA molecules to beads. Pyrosequencing can then take place by washing the plate with smaller bead-linked enzymes and dNTPs. In clonal amplification, primers bind to single-stranded DNA templates and a DNA polymerase triggers a response that produces many duplicates of the identical sequence. This parallelization (from which the term massive parallel sequencing as synonymous of NGS) significantly boosted the yield of sequencing efforts. During the sequencing reaction, nucleotides are introduced in a sequential manner, and the generated signal from each nucleotide is measured and categorized (Figure 2) [14]. This breakthrough led to the initiation of various genome sequencing programs, including the sequencing of James Watson's genome, who is known for co-discovering the Watson–Crick structure of DNA, that was finished in 2007 [21]. Further, the Neanderthal Genome Project led to the successful sequence of the Neanderthal genome, published in 2009 [22].

Subsequently, numerous sequencing platforms, including Illumina (formerly Solexa), IonTorrent, and Pacific Biosciences, did appear. The detected signal varies depending on the equipment used and may involve the release of pyrophosphate, fluorescence emission, or a change in electric current.

Solexa introduced a sequencing technique in which a single stranded DNA library is created and then passed over a flow cell that contains complementary oligonucleotides to one of the two adaptor sequences on the DNA fragments [23–25]. The creation of dense clusters of amplified fragments is then enabled by a process known as “bridge amplification”, and the original DNA strands result in clusters of clonal populations. This process involves the replication of DNA arching over to initiate the next polymerization of surrounding surface-bound oligonucleotides. As sequencing-by-synthesis occurs, a fluorescent signal can be recognized each time a single dNTP is added consecutively. The number of clusters read grows over time.

Solexa performed sequencing on the bacteriophage Φ X174, which was previously sequenced by Sanger utilizing the Sanger sequencing technique [23–25]. A single run of Solexa SBS technology yielded more than 3 million bases. In 2006, Solexa introduced the Genome Analyzer, which further enhanced sequencing capabilities by enabling the sequencing of 1 gigabase (Gb) of data in a single run. Illumina bought Solexa in 2007, establishing itself as a frontrunner in short read sequencing technology, also known as fragment-based approaches (Figure 3) [14]. In parallel, other methods such as Ion Torrent, which monitors the pH difference during polymerisation, and SOLiD, which employs sequencing-by-ligation rather than synthesis (e.g., catalyzed with a polymerase), appeared, integrating the NGS landscape.

In general, the implementation of these next-generation sequencers greatly diminished the expense of sequencing, driving forward scientific investigation and

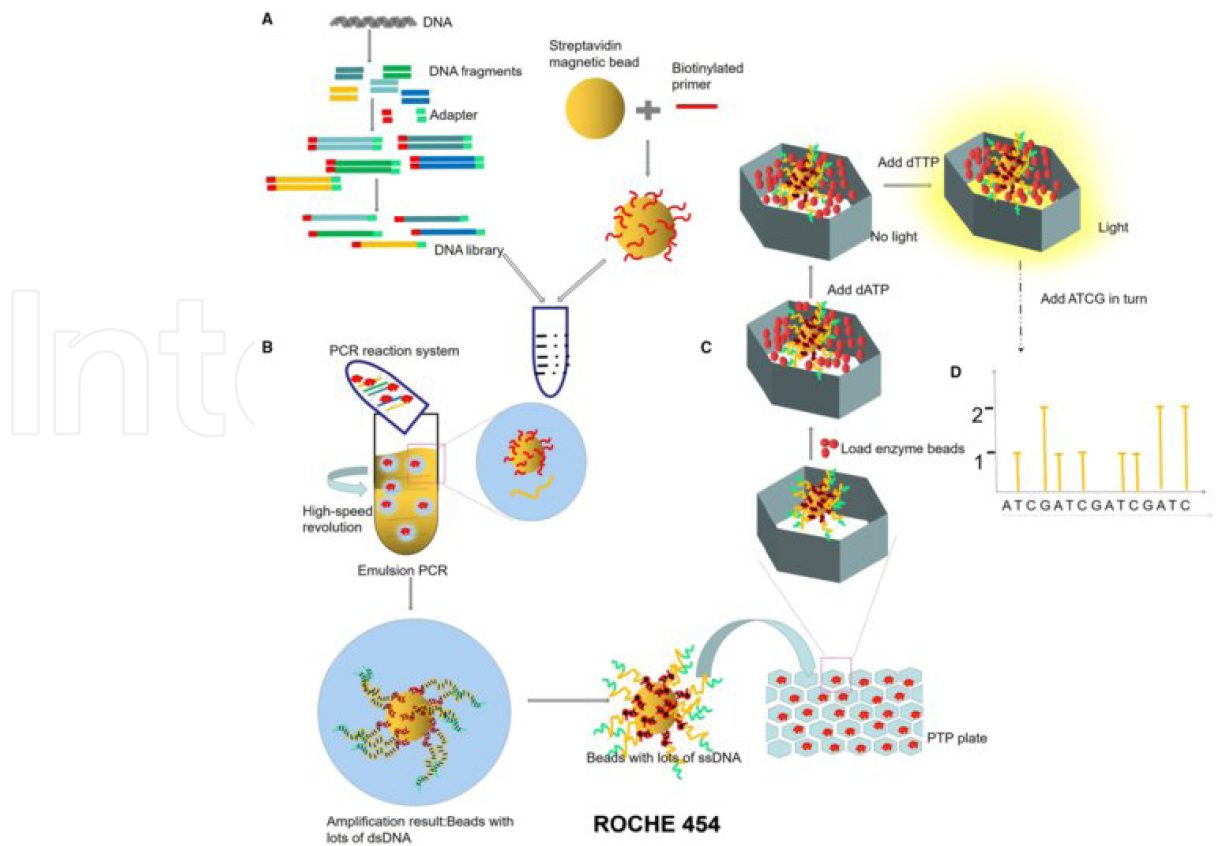


Figure 2. ROCHE 454 sequencing process. (A) Construction of micro-reaction system. DNA library is constructed by adding an adapter at the ends of DNA fragments. The magnetic beads combined with primers are mixed in the PCR reaction system (DNA fragments, enzymes, dNTPs). (B) Emulsion PCR. PCR microreaction systems are formed by injecting water (PCR mixture) into the oil, and each system contained only one template and one bead. (C) Pyrophosphate sequencing. Pyrophosphate technology is an enzyme cascade chemiluminescence reaction in the same reaction system catalyzed by four enzymes. In each circle of the sequencing reaction, only one dNTP is added. If it just matches the next base of the DNA template, it will be added to the 3' end of the sequencing primer under the action of DNA polymerase and release a molecular PPI at the same time. Under the action of ATP sulfurylase, luciferase, and apyrase, an enzyme cascade chemiluminescence reaction is triggered by the PPI. (D) Light signal reception. (Adapted from Zhang L, Chen F, Zeng Z, Xu M, Sun F, Yang L, Bi X, Lin Y, Gao Y, Hao H, Yi W, Li M and Xie Y. *Advances in metagenomics and its application in environmental microorganisms. Front Microbiol.* 2021;12: 766364.)

industrial utilization of sequencing into a novel epoch. However, those methods all have a limitation in read length, as typically yielded reads ranging between 50 to 500 bp.

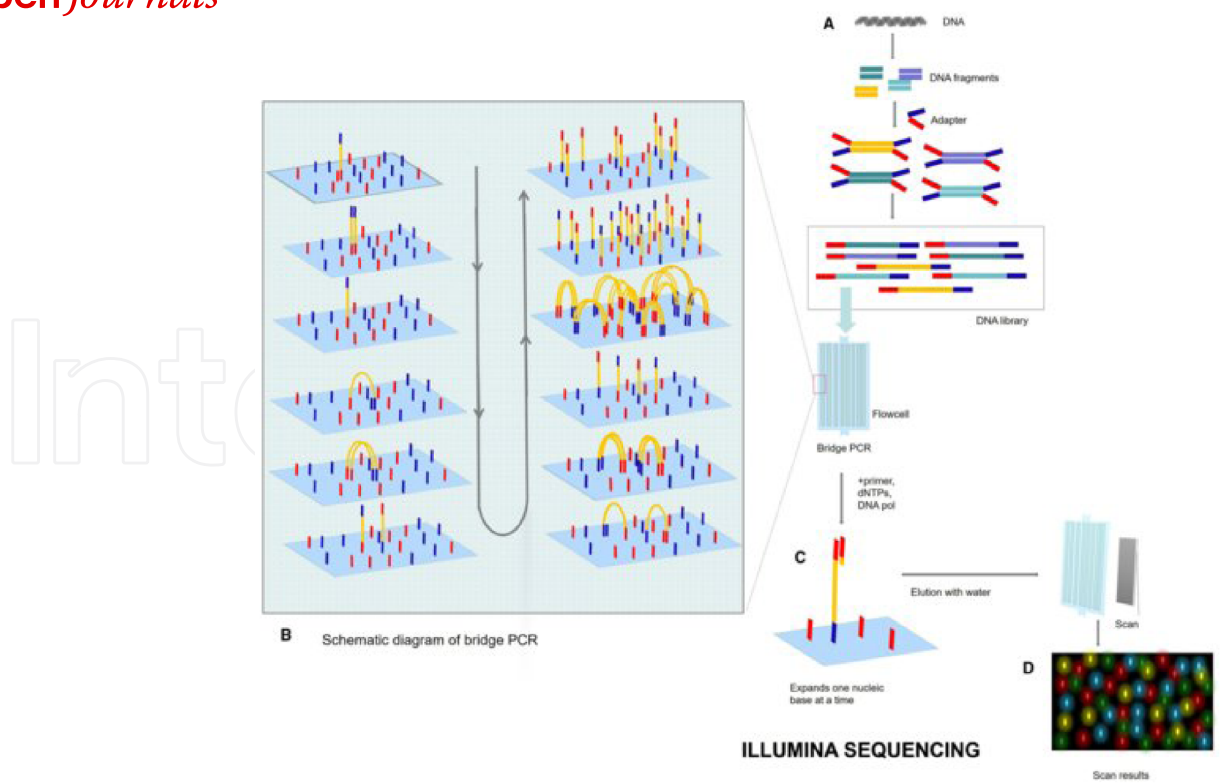


Figure 3. Illumina sequencing process. (A) DNA library. Breaks the genome DNA to form DNA fragments, adds adapters at both ends, and constructs a single-stranded DNA (ssDNA) library. After adding DNA library into flowcell, DNA fragments will be attached to the surface of flowcell when passing through flowcell. (B) Bridge PCR. Each DNA fragment is clustered in its position. After bridge amplification, each cluster contains many copies of the ssDNA template. (C) SBS: DNA polymerase, primers, and four dNTPs with specific fluorescence are added to the reaction system at the same time. The 3' end of these dNTPs relates to an azide group, which can block the incorporation of the next base, so only one base can be extended at a time. After washing off the remaining dNTP and enzyme with water, photos are scanned. After scanning, add chemical reagent, cut off the azide group and quenching fluorescence, and next cycle. (D) Fluorescent signal reception. (Adapted from Zhang L, Chen F, Zeng Z, Xu M, Sun F, Yang L, Bi X, Lin Y, Gao Y, Hao H, Yi W, Li M and Xie Y. Advances in metagenomics and its application in environmental microorganisms. *Front Microbiol.* 2021;12: 766364.)

3. Sequencing technologies of the third and fourth generations

The third- and fourth-generation sequencing methods involve single-molecule technologies capable of producing reads of over 10,000 base pairs, known as long

read technologies. These technologies offer substantial enhancements in addressing the difficulties associated with *de novo* genome assembly and structural analysis. The concept of third generation sequencing emerged in 2008–2009, as researchers convened to exchange knowledge regarding their most recent technology [26] (Table 1). Currently, there are various commercial platforms that utilize third-generation DNA sequencing methods, including Pacific Biosciences (PacBio) Single Molecule Real Time (SMRT) sequencing, and Illumina Tru-seq Synthetic Long-Read technology. Each of these approaches has an average size ranging from 5 to 15 kilobase pairs (kbp) and can surpass 100,000 bp [27]. The PacBio SMRT cell technology employs a SMRT sequencing method that utilizes a SBS approach to identify fluorescent nucleotides during their incorporation into the template molecules. It employs miniaturized wells known as zero-mode waveguides in which a single polymerase inserts labeled nucleotides and light emission is measured in real time. PacBio machines may also generate extremely lengthy reads of up to and over 10 kb in length. The device has currently generated approximately 100,000 bp with a daily output of about 8 Gb. These reads are especially beneficial for genome assembly from scratch. Despite the high error-rate, the combination of various algorithms and deep sequencing of numerous template copies significantly enhances the accuracy to more than 99.99% [28].

The Oxford Nanopore sequencing platform (ONT), conversely, is considered a fourth-generation technology [29]. Its potential, however, was initially recognized even before second-generation sequencing, as electrophoresis had previously established that single-stranded RNA or DNA may be driven across a lipid bilayer via large-haemolysin ion channels. ONT uses a nanopore that is put into a membrane with electrical resistance. The current disruption caused by the passage of bases via the pore is monitored in real time to determine exact sequences of single molecules [30]. As far as DNA library preparation is concerned, in ONT sequencing, a hairpin structure is attached to the double stranded DNA, allowing the machine to read both strands in a single uninterrupted sequence. ONT sequencing can produce exceptionally long reads, exceeding 300 kilobases in length [31]. ONT later on created a lot of buzz with their nanopore platforms, which include GridION and MinION. The MinION, released in 2014, is a compact, mobile phone-sized USB gadget that has been used in the field, allowing rapid run times and small design also allow for decentralized sequencing. This method has greatly enhanced the process of identifying and screening for diseases such as the Ebola and Lassa viruses [32]. Nanopore sequencers, with future developments, could revolutionize not only the composition of data produced, but also where, when, and by whom it can be produced.

Table 1. Overview of sequencing technologies over time.

First generation sequencing	
1972	Sanger stated work on DNA sequencing
1977	Sanger developed Di_deoxy chain termination method of DNA sequencing
1977	Maxam and Gilbert developed chemical degradation method of DNA sequencing
1977	First DNA based genome sequenced (Φ X174 bacteriophage)
1995	First bacterium <i>Haemophilus influenzae</i> was sequenced by shot gun method
1996	Applied Biosystems developed automated DNA sequencing based on Sanger's method
1996	First eukaryotic genome (<i>Saccharomyces cerevisiae</i>) was sequenced
2001	First human genome draft was published by two different independent teams
Second generation sequencing	
2005	First NGS platform released Roche 454 GS-20
2006	Introduction of second NGS platform – Solexa Genome Analyzer
2006	Initiation of 1000 genome project
2007	Introduction of Roche 454 GS-FLX & ABI-SOLiD sequencer
2008	Development of Illumina GA-II
2009	Introduction of Roche 454 GS-FLX Titanium
2010	Introduction of Roche 454 GS-Junior
2011	Introduction of SOLiD 550 W & Illumina MiniSeq
2012	Introduction of Illumina HiSeq
2013	Introduction of SOLiD 5500xl W & Illumina MiniSeq
2014	Introduction of Roche 454 GS-Junior+, Illumina NextSeq 500 & Illumina HiSeq X Ten
2017	Introduction of Illumina iSeq 100
Third generation sequencing	
2008	Development of first commercial platform of third generation technology (Heliscope Biosciences)
2010	Ion Torrent released the Personal Genome Machine (PGM)
2011	Introduction of PacBIO RS C1/C2
2012	Introduction of PacBIO RS C2XL & PacBIO RS II C2 XL, Ion Torrent released Ion Proton
2013	Introduction of PacBIO RS II C2 XL
2014	Introduction of PacBIO RS II P5 C3 & PacBIO RS II P6 C4
2015	Introduction of Ion S5/S5XL 520/530/540
2016	Introduction of PacBIO sequel
Fourth Generation Sequencing	
2014	Release of MinION platform by Oxford Nanopore Technologies
2017	Release of ProMethION, GridION & SmidgION X5 platforms by Oxford Nanopore Technologies
2018	Commercialization of ProMethION platform by Oxford Nanopore Technologies

4. DNA sequencing in cancer diagnostic

Since the first recognition of specific genetic lesions in human cancer in the seventies, DNA sequencing has become more important in molecular pathology for entities definition, prognostication, and prediction. The major example is represented by the current classification of hematopoietic and lymphoid neoplasms. Indeed, several entities among leukemias and lymphomas are diagnosed by

identifying specific translocations or gene mutations [33, 34]. In addition, genetic testing is used for prognostication and patients' selection for targeted therapies, as in the case of *FLT3* (*FMS Like Tyrosine kinase 3*)-positive acute myeloid leukemias, *TP53* mutated chronic lymphocytic leukemias, and chronic myeloid leukemias carrying *ABL1* mutations [33, 34].

As far as solid tumors are concerned, several new therapeutic options are offered based on the mutation profile, and so-called companion tests (i.e., genetic testing mandatory for prescribing certain drugs) have been developed. Among others, mutations affecting *EGFR* in lung cancer, *KRAS* in colorectal, and *BRAF* in both the previous as well as in malignant melanoma, have been studied for several years to select patients to be treated with novel inhibitors. Nowadays, several institutions offer NGS panels, evaluating 30 to 400 genes, are routinely offered to cancer patients for an optimal profiling [35].

5. Microbial identification using DNA sequencing

The identification and characterization of microorganisms have evolved with the progress of various technologies and procedures. Until recently, bacteria were categorized using the DNA–DNA hybridization technique [36]. The process was arduous and severely constrained by the absence of a central database. The development of sequencing technology and the utilization of ribosomal genes, such as the 16S rRNA gene, have facilitated the establishment of a comprehensive database for precise microbiological identification. Due to the decreasing cost of sequencing and the availability of advanced algorithmic methods for analyzing microbiological databases, numerous organism-specific databases have been developed and utilized. These databases can even offer identification at the strain level [2].

It should be noted, however, that despite the adoption of NGS in most recent studies, Sanger sequencing still played a crucial role in supplying comparative data. Recent developments in NGS have enabled researchers to characterize and investigate species and organisms that cannot be cultivated under typical laboratory conditions. Using both short and long reads in whole genome sequencing, along with the quick sequencing technique offered by ONT, has had substantial effects on microbiology study and advancement. We are on the verge of applying these technical developments in an industrial environment, where scientific instruments are utilized to enhance our understanding of microbes and offer quick diagnostic and screening techniques. Accugenix® is eagerly anticipating the outcomes that will arise from the integration of these advanced methods in microbial identification.

6. Summary and perspectives

Over the last 50 years, researchers from all over the world have worked hard to develop and improve the technologies that support DNA sequencing. Future perspectives, first, include a further drop in NGS costs making those technologies the new routine and instruments in medical diagnostic, agri-food industry, and scientific research. Furthermore, a parallel implementation of bioinformatic approach is needed. The tremendous amount of added information provided by NGS needs a refined bioinformatic deconvolution to become clinically useful. It is the case, for instance, of intron characterization deriving from whole genome sequencing, which currently still represents an unsolved issue. Nonetheless, it must be acknowledged that NGS will soon become an integral part of routine diagnostic for most if not all cancer patients and most infectious diseases, this contributing an essential step toward precision medicine.

Conflict of interest

The author declares no conflict of interest.

References

- 1 Sanger F. Frederick Sanger—Biographical, 1980. Available from: http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1980/sanger-bio.html.
- 2 Heather JM, Chain B. The sequence of sequencers: the history of sequencing DNA. *Genomics*. 2016 Jan;107(1):1–8. Available from: doi:10.1016/j.jygeno.2015.11.003. Epub 2015 Nov 10. PMID: 26554401; PMCID: PMC4727787.
- 3 Holley RW, et al. Structure of a ribonucleic acid. *Science*. 1965;147: 1462–1465. Available from: <http://www.sciencemag.org.libproxy.ucl.ac.uk/content/147/3664/1462.full.pdf>.
- 4 International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature*. 2004 Oct 21;431(7011):931–945. Available from: doi:10.1038/nature03001. PMID: 15496913.
- 5 Heather JM, Chain B. The sequence of sequencers: the history of sequencing DNA. *Genomics*. 2016 Jan;107(1):1–8. Available from: doi:10.1016/j.jygeno.2015.11.003. Epub 2015 Nov 10. PMID: 26554401; PMCID: PMC4727787.
- 6 Avery OT, MacLeod C, McCarty M. Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. *J Exper Med*. 1944;79: 137–158.
- 7 Hargittai I. Paradigms and paradoxes: complementarity in chemical structures—a tribute to Erwin Chargaff. *Struct Chem*. 2022;33: 1003–1005. Available from: doi:10.1007/s11224-022-01915-1.
- 8 Watson J, Crick F. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature*. 1953;171: 737–738. Available from: doi:10.1038/171737a0.
- 9 Holley RW, Apgar J, Merrill SH, Zubkoff PL. Nucleotide and oligonucleotide compositions of the alanine-, valine-, and tyrosine-acceptor soluble ribonucleic acids of yeast. *J Am Chem Soc*. 1961;83: 4861–4862. Available from: <http://pubs.acs.org/doi/abs/10.1021/ja01484a040>.

- 10 Sanger F, Brownlee G, Barrell B. A two-dimensional fractionation procedure for radioactive nucleotides. *J Mol Biol.* 1965;**13**: 373–398.
- 11 Sanger F, Coulson AR. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol.* 1975;**94**: 441–446.
- 12 Holley RW, Madison JT, Zamir A. A new method for sequence determination of large oligonucleotides. *Biochem Biophys Res Commun.* 1964;**17**: 389–394.
- 13 Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA.* 1977 Dec;**74**(12):5463–5467. Available from: doi:10.1073/pnas.74.12.5463. PMID: 271968; PMCID: PMC431765.
- 14 Zhang L, Chen F, Zeng Z, Xu M, Sun F, Yang L, et al. Advances in metagenomics and its application in environmental microorganisms. *Front Microbiol.* 2021;**12**: 766364.
- 15 Maxam AM, Gilbert W. A new method for sequencing DNA. *Proc Natl Acad Sci USA.* 1977 Feb;**74**(2):560–564. Available from: doi:10.1073/pnas.74.2.560. PMID: 265521; PMCID: PMC392330.
- 16 Akintunde O, Tucker T, Carabetta VJ. The evolution of next-generation sequencing technologies [Internet]. arXiv; 2023. <https://arXiv:2305.08724v1>. PMID: 37292469; PMCID: PMC10246072.
- 17 Sanger F, Coulson AR, Hong GF, Hill DF, Petersen GB. Nucleotide sequence of bacteriophage lambda DNA. *J Mol Biol.* 1982 Dec 25;**162**(4):729–773. Available from: doi:10.1016/0022-2836(82)90546-0. PMID: 6221115.
- 18 Nyrén PL. Enzymatic method for continuous monitoring of DNA polymerase activity. *Anal Biochem.* 1987;**238**: 235–238. Available from: <http://www.sciencedirect.com/science/article/pii/0003269787901588>.
- 19 Ronaghi M, Karamohamed S, Pettersson B, Uhlen M, Nyrén PL. Real-time DNA sequencing using detection of pyrophosphate release. *Anal Biochem.* 1996;**242**: 84–89.
- 20 Ronaghi M, Uhlen M, Nyrén PL. A sequencing method based on real-time pyrophosphate. *Science.* 1998;**281**: 363–365. Available from: <http://www.sciencemag.org/cgi/doi/10.1126/science.281.5375.363>.
- 21 Check E. James Watson's genome sequenced. *Nature.* 2007; Available from: doi:10.1038/news070528-10.
- 22 Dalton R. Neanderthal genome to be unveiled. *Nature.* 2009;**457**: 645.
- 23 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. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1363783&tool=pmcentrez&rendertype=abstract>.
- 24 Bentley DR, et al. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature.* 2008;**456**: 53–59. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2581791&tool=pmcentrez&rendertype=abstract>, <http://www.nature.com/nature/journal/v456/n7218/full/nature07517.html>.
- 25 Gupta N, Verma VK. Next-generation sequencing and its application: empowering in public health beyond reality. *Microbial Technol Welf Society.* 2019 Sep 13;**17**: 313–341. Available from: doi:10.1007/978-981-13-8844-6_15. PMCID: PMC7122948.
- 26 Check Hayden E. Genome sequencing: the third generation. *Nature.* 2009;**457**: 768–769.
- 27 Amarasinghe SL, Su S, Dong X, et al. Opportunities and challenges in long-read sequencing data analysis. *Genome Biol.* 2020;**21**: 30.
- 28 Chin CS, Alexander D, Marks P, et al. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Meth.* 2013;**10**: 563–569.
- 29 Feng Y, et al. Nanopore-based fourth-generation DNA sequencing technology. *Genom Proteom Bioinform.* 2015;**13**(1):4–16.

- 30 Lu H, Giordano F, Ning Z. Oxford Nanopore MinION sequencing and genome assembly. *Genom Proteom Bioinform.* 2016;**14**(5):265–279.
- 31 Midha MK, Wu M, Chiu KP. Long-read sequencing in deciphering human genetics to a greater depth. *Hum Genet.* 2019;**138**: 1201–1215.
- 32 Cao Y, et al. Nanopore sequencing: a rapid solution for infectious disease epidemics. *Sci China Life Sci.* 2019;**62**: 1101–1103.
- 33 Khoury JD, Solary E, Abla O, et al. The 5th edition of the World Health Organization classification of haematolymphoid tumours: myeloid and histiocytic/dendritic neoplasms. *Leukemia.* 2022 Jul;**36**(7):1703–1719. Available from: doi:10.1038/s41375-022-01613-1. Epub 2022 Jun 22. PMID: 35732831; PMCID: PMC9252913.
- 34 Alaggio R, Amador C, Anagnostopoulos I, et al. The 5th edition of the World Health Organization classification of haematolymphoid tumours: lymphoid neoplasms. *Leukemia.* 2022 Jul;**36**(7):1720–1748. Available from: doi:10.1038/s41375-022-01620-2. Epub 2022 Jun 22. Erratum in: *Leukemia.* 2023 Sep;**37**(9):1944–1951. PMID: 35732829; PMCID: PMC9214472.
- 35 Milbury CA, Creeden J, Yip WK, et al. Clinical and analytical validation of FoundationOne® CDx, a comprehensive genomic profiling assay for solid tumors. *PLoS One.* 2022 Mar 16;**17**(3):e0264138. Available from: doi:10.1371/journal.pone.0264138. PMID: 35294956; PMCID: PMC8926248.
- 36 Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, et al. DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol.* 2007;**57**: 81–91.

IntechOpen