Trends in Biotechnology

CellPress OPEN ACCESS

Review

Sequential sequencing by synthesis and the next-generation sequencing revolution

Mathias Uhlen [1,2,@,*] and Stephen R. Quake [3,4,@]

The impact of next-generation sequencing (NGS) cannot be overestimated. The technology has transformed the field of life science, contributing to a dramatic expansion in our understanding of human health and disease and our understanding of biology and ecology. The vast majority of the major NGS systems today are based on the concept of 'sequencing by synthesis' (SBS) with sequential detection of nucleotide incorporation using an engineered DNA polymerase. Based on this strategy, various alternative platforms have been developed, including the use of either native nucleotides or reversible terminators and different strategies for the attachment of DNA to a solid support. In this review, some of the key concepts leading to this remarkable development are discussed.

Highlights

Next-generation sequencing (NGS) involving massively parallel DNA analysis has made an enormous impact on life science, medicine, and biotechnology, through a multitude of applications.

The vast majority of the major NGS systems are based on the concept of 'sequencing by synthesis' (SBS) with sequential detection of nucleotide incorporation using an engineered DNA polymerase.

The basic principles of SBS include attachment of DNA fragments to a solid support, conversion to a single-strand template and the annealing of a primer, the incorporation of complementary nucleotides by a polymerase, and detection of this incorporation.

The development of NGS spans several decades of innovations, from early systems using natural nucleotides to later systems for massively parallel sequencing systems using reversible fluorescent nucleotides.

The impact of NGS

There is no question that massively parallel sequencing (MPS) technology, often referred to as NGS, has had a unique and huge impact on life science research [1,2], with a large number of studies using this technology currently published every day. The number of DNA sequences in public databases has exploded since MPS by synthesis was commercially introduced in 2005 [3]. This has also led to a dramatic decrease in cost and efforts to sequence whole genomes. This is illustrated by the estimated cost for the first human genome published in 2001 using electrophoretic technology for base calling, which was estimated to be US$2–3 billion, while today it is possible to sequence a human genome with MPS for less than US$1000 [1], a cost reduction by a factor of over 1 million. This is remarkable and it is hard to find a similar example in scientific history.

This has led to an explosion of scientific data and the entrance of a new era in medicine and biology driven by 'big data' and data-driven research. In the field of genetics, the creation of maps covering the genetic diversity in various human populations [4] has greatly increased our understanding of the relationship of genes and diseases. Similar maps across the 'tree of life' [5] have increased our understanding of the parts list of human building blocks. The discovery of a new human ancestor, the Denisovans [6], was enabled by NGS, and population studies showed widespread remains of both Neanderthal and Denisovan DNA in our genomes. The technology has also led to clinical practice, both to understand and to treat cancers [7], but also to allow diagnosis in children with unknown disease-causing mutations, resulting in the design of drug treatments [8,9]. In the following, the various concepts enabling the rapid development of NGS are discussed.

The concept of SBS

The objective of SBS is to determine the sequencing of a DNA sample by detecting in a sequential manner the incorporation of nucleotides using an engineered DNA polymerase (Figure 1). An engineered polymerase is used to synthesize a copy of a single strand of DNA and the

[1]Science for Life Laboratory, Department of Protein Science, KTH Royal Institute of Technology, Stockholm, Sweden
[2]Department of Neuroscience, Karolinska Institutet, Stockholm, Sweden
[3]Departments of Bioengineering and Applied Physics, Stanford University, Stanford, CA, USA
[4]Chan Zuckerberg Initiative, Redwood City, California, USA, Stanford, CA, USA

*Correspondence:
mathias.uhlen@scilifelab.se (M. Uhlen).
@Twitter: @ProteinAtlas (Human Protein Atlas), @MathiasUhlen (M. Uhlen), and @StephenQuake (S.R. Quake).

© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

incorporation of each nucleotide is monitored. The key parts are highly similar for all embodiments of SBS and include the following:

(i) attachment of the DNA to be sequenced to a solid support, usually combined with amplification of the DNA to enhance the subsequent signal;
(ii) generation of single-stranded DNA on the solid support;
(iii) primer-dependent incorporation of complementary nucleotides using an engineered polymerase; and
(iv) detection of the incorporated nucleotide.

Steps (iii) and (iv) are repeated and the sequence is assembled from the signals obtained in step (iv). This principle of SBS has been used for almost all MPS efforts and it has contributed to the vast majority of sequence information generated during the past decade [10].

The historical development of sequential SBS has been reviewed by others [10], which we briefly summarize here. The concept was first described in 1993 [11] in the form of a technique later known as pyrosequencing. In this case, nucleotide incorporation was detected by measuring pyrophosphate products of incorporation. In the first publication [11], all of the key concepts of SBS were introduced, including the amplification of DNA to enhance the subsequent signal and attachment of the DNA to be sequenced to a solid support, the generation of single-stranded DNA on the solid support, the incorporation of nucleotides using an engineered polymerase, and light detection of the incorporated nucleotide. This paper also outlines a vision of MPS: 'Automated on-line methods with multiple samples in parallel can be envisioned'. In a follow-up article [12], the concept was further developed, and a few years later Ronaghi, Uhlén and Nyrén [13] showed that non-incorporated nucleotides could be removed with a fourth enzyme (apyrase) allowing SBS to be performed without the need to wash away non-incorporated nucleotides. A commercial instrument based on SBS (called Pyrosequencing) was launched in 2000 with all key concepts for SBS with real-time detection and with a throughput of 96 samples in parallel

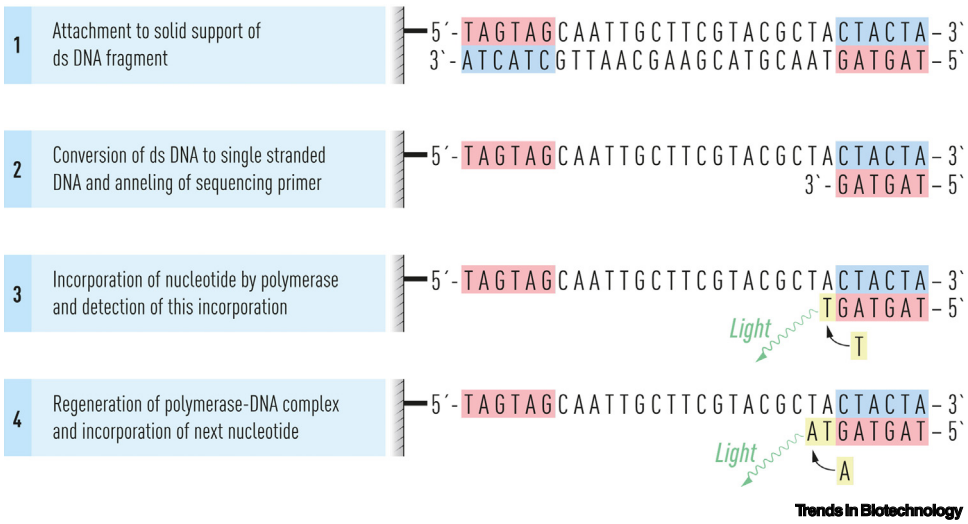

Figure 1. The concept of sequencing by synthesis. Amplified DNA fragments are attached to a solid support and single-stranded DNA is generated. A primer is annealed to the end of the fragment and the incorporation of a nucleotide corresponding to the template strand is achieved using a DNA polymerase without proof-reading capacity. The incorporation of the nucleotide is monitored (e.g., via light) and steps 3 and 4 are repeated. Abbreviation: ds, double-stranded.

Glossary

Gb: 1 billion bases (nucleotides) of DNA sequence data.
RNA-seq: sequencing of RNA using NGS.

[14]. A modified version of this instrument is still available and is used for many applications, including DNA methylation/epigenetics [15,16] and forensics studies [17].

The concept of MPS

The first next-generation sequencers (Figure 2) were based on pyrosequencing chemistry and were commercialized by Rothberg and coworkers [3] from the company 454 Life Sciences in the USA. They showed that sequencing could be performed in a highly parallel manner, and in the paper they described the successful sequencing of the genome of *Mycoplasma genitalium*, which is also the first description of whole-genome sequencing using MPS. Many important applications were enabled by this pioneering instrument [18], including the analysis of the Neanderthal genome [19] in collaboration with Paabo and coworkers at the Max Planck Institute in Leipzig, Germany. Others used this approach to study microbial diversity in the deep sea [20], to study the microbial genomics of archaea in soil [21], to study the possible causes of honeybee hive collapse [22], and to perform the first single-cell genome sequencing of an unculturable organism [23]. This approach was also used to determine one of the first individual human genome sequences in 2008 [24].

The development of fluorescence chemistries

A further improvement of MPS by synthesis was the development of reversible and fluorescently labeled terminators to address the issue of homopolymers, which were less adequately determined by the Pyrosequencing detection system. The use of fluorescently labeled nucleotides for sequencing dates to the 1980s using conventional electrophoretic sequencing. Reversible terminators were first described by Metzker and colleagues [25], showing that 3′-modified nucleotides could be used for base-specific termination and photolytic removal of the 3′-protecting group. A few years later, the two UK chemists Balasubramanian and Klenerman filed a patent [26] describing 'Arrayed biomolecules and their use in sequencing' in which they proposed the use of fluorescently labeled nucleotides combined with reversible terminators to allow SBS. Around the same time, the USA-based Quake group was pursuing a similar strategy of fluorescence photobleaching sequencing, which they described in a grant proposal to the National Institutes of Health (NIH) (https://grantome.com/grant/NIH/R29-HG001642-04). Balasubramanian and Klenerman initially aimed to achieve single-molecule detection but later abandoned this strategy and the concept was combined with *in vitro*-amplified DNA on a solid support [27], whereas the Quake group began with the idea of using multiple copies of a DNA template on a surface but then became the first to demonstrate single-molecule SBS [28].

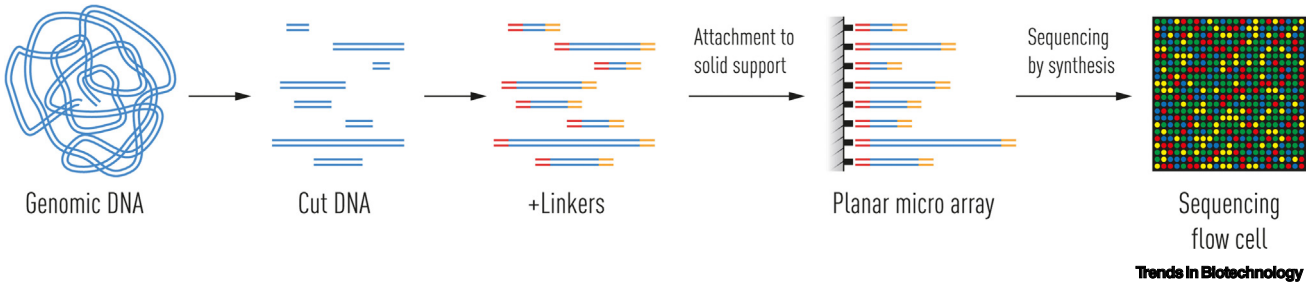

Figure 2. The concept of massively parallel sequencing. Genome DNA is fragmented and linkers are attached to the ends of the fragments. The double-stranded DNAs are attached to a planar microarray with each fragment separated physically in two dimensions on a microfabricated flow cell. The sequencing by synthesis protocol outlined in Figure 1 is followed and the light generated in each spot on the array is detected in a massively parallel manner. The sequence for each spot can be read based on the light emitted during each step of nucleotide incorporation.

An instrument called the 'Genome Analyzer' for MPS by synthesis was subsequently launched in 2006 by the company Solexa founded by Balasubramanian and Klenerman. The modified SBS strategy used reversible terminators and involved immobilizing the sequencing templates and primers on a solid support, primer extension by the incorporation of a nucleotide with a 3′-blocking group by a polymerase, detection of the color of the fluorophore carried by the extended base after washing away the unincorporated nucleotides to identify the incorporated nucleotide, and removal of the fluorescent tag and the 3′-blocking group and repeating the steps by incorporation of the next nucleotide. Solexa was acquired by Illumina in 2007, and although few peer-reviewed articles were published from the company, a paper was published in 2008 [29] reporting on the use of their instrument for human genome sequencing. Similarly, the single-molecule sequencing approach from the Quake group was commercialized by Helicos Biosciences, which published a paper demonstrating viral genome sequencing with their instrument [30], followed by a publication from the same group [31] determining the sequence of an entire human genome. Table 1 shows a short summary of some of the key milestones in the development of the concept of NGS.

Amplified DNA on solid support

For methods that do not use single-molecule sequencing, it became crucial to perform controlled amplification of individual templates. Several PCR-based alternatives for amplification of DNA have been successfully explored.

(i) Solid-phase sequencing. This strategy depended on PCR-amplified DNA fragments bound to solid support using the biotin-streptavidin system. This technology had already been

Table 1. Key publications and patents in the field of SBS

Year	Description	Refs
1993	The concept of SBS	[11]
1994	The use of reversible terminators	[25]
1998	Solution-based Pyrosequencing	[13]
1998	Array of molecules on surface (patent)	[26]
1998	Clonal amplification on surface (patent)	[27]
1999	*In situ* localized amplification	[36]
2003	Single-molecule sequencing	[28]
2003	Fluorescent *in situ* sequencing	[64]
2005	The concept of MPS	[3]
2006	Analysis of Neanderthal genome using SBS	[19]
2008	SBS using reversible terminators	[29]
2008	Transcriptomics analysis (RNA-seq)	[65]
2010	1000 Genomes Project published	[66]
2014	Single-cell genomics (transcriptomics)	[62]
2016	Population-based genome sequencing (Iceland population)	[4]
2016	Spatial transcriptomics	[63]
2017	The Human Cell Atlas project	[67]
2019	The 25 000 cancer genomes project	[43]
2021	Next-generation blood profiling using proximity extension assay	[50]
2021	The 100 000 Genomes Project for rare diseases	[44]
2023	Zoonomia: mammalian species genomes	[45]

described in the 1980s [32,33] and the approach was later used to develop Pyrosequencing, which was launched commercially in 2000. A variant of this concept was described by the Syvanen group [34] in which a oligonucleotide array was used for mini-sequencing to allow multiplex detection of mutations.

(ii) Emulsion PCR. This strategy depended on amplifying the template in microdroplets created by an emulsion technique. This concept was developed by the company 454 and was used to develop the 454/Pyrosequencing platform published [3] and launched in 2005.

(iii) Bridge-PCR. This strategy, also called *in vitro* molecular cloning, enables *in situ* localized amplification of DNA molecules using a solid support with synthesized primers. This was introduced in a patent application by Kawashima and coworkers [27] from the company Serano and the concept was further developed by the Church group to allow amplification of a single DNA molecule to form a 'polony' [35,36]. The bridge-PCR technology was later used for the launch of the Genome Analyzer by Solexa in 2006.

For most of the SBS platforms, the availability of an amplified DNA template on a solid support was one of the key concepts that enabled the introduction of MPS.

Alternative systems for MPS

After the introduction of the first MPS system in 2005 [3], many alternative systems were developed, including the Solexa/Illumina system based on reversible terminators [29], Helicos for single-molecule sequencing using reversible fluorescently labeled terminators [28], and PacBio for sequencing long reads using single DNA molecule read-out [37]. An approach with an alternative to SBS, called nanopore sequencing, was also developed [38]. Later, the reversible terminator strategy was combined with 'nanoballs' [39], also based on SBS. Table 2 shows some of the instruments for SBS, and it summarizes the amplification strategy used and the choice of reagents (nucleotides) in the assay. All of these use the concept of SBS.

MPS can also be performed without SBS, but with the same objective to allow sequence analysis to be performed in a scalable and parallel manner. Examples of this include 'sequencing-by-ligation' [40] and 'sequencing by hybridization' [41], as well as the nanopore system, which relies on the monitoring of nucleotides passing a protein nanopore. However, the concept of SBS has become the leading analytical platform for NGS, with the Illumina instruments completely dominating the genomics field during the past 10 years.

Applications of MPS

The impact of MPS technology on science has been overwhelming and the applications can roughly be divided into three separate research fields.

Whole-genome sequencing

MPS systems have allowed genomes to be assembled in an efficient manner. The approach usually relies on the concept of 'shotgun' sequencing, introduced in 1994 by the Myers and Venter groups [42] in the context of Sanger sequencing. In this approach, both random sequence fragments and paired ends from larger molecules are sequenced and assembled by bioinformatics to yield the complete whole genome. This approach, and use of the SBS concept, has led to an explosion of sequenced genomes ranging from bacteria, fungi, and plants to mammalian species, including the sequencing of whole populations of humans [4]. The strategy was used by the International Cancer Genome Consortium (ICGC) project to sequence 25 000 cancer genomes [43] and in the 100 000 Genomes Project to analyze the whole genomes of people affected by rare disease [44]. Recently, the analytical platform has also been used to explore the genetic variability across a large number of mammalian species and their relationship with complex phenotypes [45].

Table 2. Examples of instruments using sequential SBS

Year	Instrument/system	Amplification	Nucleotides	Comment
2000	PSQ96 (Pyrosequencing)	Magnetic beads	Natural	SBS concept
2005	454 Instrument	Microdroplets	Natural	MPS concept
2006	Genome Analyzer (Solexa)	Bridge-PCR	Reversible	Improve homopolymers
2007	tSMS (Helicos)	Single molecule	Reversible	Single molecule
2010	Hiseq 2000 (Illumina)	Bridge-PCR	Reversible	200 **Gb** per run
2010	PacBio RS	Circular DNA	Fluorescent	Extreme long reads
2010	Ion Torrent instrument	Microdroplets	Natural	Semiconductor detector
2016	GenCode (10X Genomics)	Single molecule	Reversible	Single-cell analysis
2019	DNBSeq-T7 (MGI)	Nanoballs	Reversible	6000 Gb per run
2022	NovaSeq X (Illumina)	Bridge-PCR	Reversible	6000 Gb per run

Outstanding questions

Will the trend of lower costs for DNA sequencing continue and will this result in whole-genome sequencing of whole populations?

Will genome sequencing facilitate treatment choices by physicians based on personalized medicine strategies?

Is it possible to overcome the ethical concerns with whole-genome sequencing regarding the safety and security of the individual genome information?

Reference-based sequencing

Cost-effective technology to sequence short stretches of DNA has enabled a large number of new applications with huge impacts in life science. In these applications, MPS is used to analyze sequences that are subsequently compared with reference sequences, and thus the origin of the sequences can be inferred. By counting the presence of a particular sequence, it is possible to quantitatively estimate biological phenomena. An example of this application is transcriptomics [**RNA-seq** (see Glossary)] [46], in which the RNA profile in organs, tissues, and cells is estimated. Another example is microbiome analysis [47], in which microbial populations in various environmental niches, including the human gut, are studied. In addition, there is a long tail of various applications, including epigenetic studies involving methylation analysis, studies of transcription factor binding sites (ChIP-seq) and assay for transposase-accessible chromatin using sequencing (ATTACK-seq) to determine chromatin accessibility across the genome [48].

Barcode counting

An interesting application of SBS is the use of 'tag' library molecules [49] based on synthetic oligonucleotide sequences ('barcodes'). These applications take advantage of MPS as a way to quantify the number of tags in a given sample. This technology is now used in a number of new applications, including 'next-generation blood profiling' [50,51] allowing thousands of proteins to be simultaneously detected from a small drop of blood. A common strategy is to combine barcode counting with reference-based sequencing. One example of this is single-cell genomics [52], in which reference-based sequencing is used for transcript counting and the barcode is used to identify the cell origin of a sample. Another example is spatial transcriptomics [53], where reference-based sequencing is again used for transcript counting, while the barcode is used to localize RNA molecules on the tissue sample.

Concluding remarks and future perspective

In summary, the impact of sequential SBS in the life science field is enormous. The concept has allowed the rapid development of NGS platforms and thus transformed the field of life science, contributing to a dramatic expansion in our understanding of human health and disease as well as our understanding of biology and ecology. The vast amount of data has led to a new era of data-driven life science in which machine learning and other AI-based methods are and will be of increased importance to expand our life science knowledge base. The 25 000 cancer genomes project [43] and the 100 000 Genomes Project [44] are just two recent examples among many of how rapidly the field is moving forward. The trend of these 'big science' projects

generating open-access data makes it possible for researchers around the world to integrate omics data and analyze results based on both externally and internally generated data.

The exponential growth of sequencing data will most likely continue for many years, driven by even lower costs of whole-genome sequencing (see Outstanding questions). It is possible to envision whole-genome sequencing of all individuals at birth to facilitate later treatment choices by physicians based on personalized medicine strategies; however, such scenarios depend on ethical considerations and the safety of the information on the individual, and thus common rules and regulations to safeguard genome data must be in place [54]. It is also not unlikely that a large proportion of cancer patients will have their tumor genome sequenced to adapt treatment to the genetic make-up of the respective tumor [55].

Another interesting trend is the integration of SBS with various complementary technologies, sometimes referred to as multiomics [56]. Advances in proteomics [57], metabolomics [58], and bioimaging [59] combined with NGS data make it possible to study biology and medicine in ways impossible only a few years ago. A recent example is the development of next-generation blood protein profiling, in which proximity extension assays [50,51] and SomaScan assays [60,61] now allow thousands of proteins to be analyzed in a quantitative manner starting from only a small drop of blood.

A recent trend is the increased technical improvements in single-cell genomics [62] and spatial transcriptomics [53,63] to allow more depth and breadth in the investigation of the transcriptional landscape in cells, tissues, and organs. The technical improvements will most likely continue and thus advance our holistic understanding of the functional building blocks of life. In summary, SBS technology has led to a revolution in the field of DNA and RNA sequencing and in doing so transformed life science research and greatly contributed to an expansion of our knowledge of biology and medicine.

Acknowledgments

We are grateful to Per-Åke Nygren for critical reading and comments. The funding from the Knut and Alice Wallenberg Foundation is acknowledged.

Declaration of interests

M.U. is the cofounder of the company Pyrosequencing (Sweden) and S.R.Q. is the cofounder of Helicos Biosciences (USA).

References

1. Gibbs, R.A. (2020) The Human Genome Project changed everything. *Nat. Rev. Genet.* 21, 575–576
2. McGinn, S. and Gut, I.G. (2013) DNA sequencing – spanning the generations. *New Biotechnol.* 30, 366–372
3. Margulies, M. *et al.* (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437, 376–380
4. Gudbjartsson, D.F. *et al.* (2015) Sequence variants from whole genome sequencing a large group of Icelanders. *Sci. Data* 2, 150011
5. Hug, L.A. *et al.* (2016) A new view of the tree of life. *Nat. Microbiol.* 1, 16048
6. Meyer, M. *et al.* (2012) A high-coverage genome sequence from an archaic Denisovan individual. *Science* 338, 222–226
7. Mardis, E.R. and Wilson, R.K. (2009) Cancer genome sequencing: a review. *Hum. Mol. Genet.* 18, R163–R168
8. Stranneheim, H. *et al.* (2021) Integration of whole genome sequencing into a healthcare setting: high diagnostic rates across multiple clinical entities in 3219 rare disease patients. *Genome Med.* 13, 40
9. Marwaha, S. *et al.* (2022) A guide for the diagnosis of rare and undiagnosed disease: beyond the exome. *Genome Med.* 14, 23
10. Heather, J.M. and Chain, B. (2016) The sequence of sequencers: the history of sequencing DNA. *Genomics* 107, 1–8
11. Nyren, P. *et al.* (1993) Solid phase DNA minisequencing by an enzymatic luminometric inorganic pyrophosphate detection assay. *Anal. Biochem.* 208, 171–175
12. Ronaghi, M. *et al.* (1996) Real-time DNA sequencing using detection of pyrophosphate release. *Anal. Biochem.* 242, 84–89
13. Ronaghi, M. *et al.* (1998) A sequencing method based on real-time pyrophosphate. *Science* 281, 365
14. Harrington, C.T. *et al.* (2013) Fundamentals of pyrosequencing. *Arch. Pathol. Lab. Med.* 137, 1296–1303
15. Claus, R. *et al.* (2012) A systematic comparison of quantitative high-resolution DNA methylation analysis and methylation-specific PCR. *Epigenetics* 7, 772–780
16. Sadikovic, B. *et al.* (2021) Clinical epigenomics: genome-wide DNA methylation analysis for the diagnosis of Mendelian disorders. *Genet. Med.* 23, 1065–1074
17. Ghemrawi, M. *et al.* (2023) Pyrosequencing: current forensic methodology and future applications – a review. *Electrophoresis* 44, 298–312
18. Rothberg, J.M. and Leamon, J.H. (2008) The development and impact of 454 sequencing. *Nat. Biotechnol.* 26, 1117–1124

19. Green, R.E. *et al.* (2006) Analysis of one million base pairs of Neanderthal DNA. *Nature* 444, 330–336
20. Sogin, M.L. *et al.* (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc. Natl. Acad. Sci. U. S. A.* 103, 12115–12120
21. Leininger, S. *et al.* (2006) Archaea predominate among ammonia-oxidizing prokaryotes in soils. *Nature* 442, 806–809
22. Cox-Foster, D.L. *et al.* (2007) A metagenomic survey of microbes in honey bee colony collapse disorder. *Science* 318, 283–287
23. Marcy, Y. *et al.* (2007) Dissecting biological "dark matter" with single-cell genetic analysis of rare and uncultivated TM7 microbes from the human mouth. *Proc. Natl. Acad. Sci. U. S. A.* 104, 11889–11894
24. Wheeler, D.A. *et al.* (2008) The complete genome of an individual by massively parallel DNA sequencing. *Nature* 452, 872–876
25. Metzker, M.L. *et al.* (1994) Termination of DNA synthesis by novel 3′-modified-deoxyribonucleoside 5′-triphosphates. *Nucleic Acids Res.* 22, 4259–4267
26. Balasubramanian, S. and Klenerman, D. Solexa Ltd. Arrayed biomolecules and their use in sequencing, WO 00/06770
27. Kawashima, E. *et al.* Method of nucleic acid amplification, WO 98/44151
28. Braslavsky, I. *et al.* (2003) Sequence information can be obtained from single DNA molecules. *Proc. Natl. Acad. Sci. U. S. A.* 100, 3960–3964
29. Bentley, D.R. *et al.* (2008) Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 456, 53–59
30. Harris, T.D. *et al.* (2008) Single-molecule DNA sequencing of a viral genome. *Science* 320, 106–109
31. Pushkarev, D. *et al.* (2009) Single-molecule sequencing of an individual human genome. *Nat. Biotechnol.* 27, 847–850
32. Stahl, S. *et al.* (1988) Solid phase DNA sequencing using the biotin-avidin system. *Nucleic Acids Res.* 16, 3025–3038
33. Hultman, T. *et al.* (1989) Direct solid phase sequencing of genomic and plasmid DNA using magnetic beads as solid support. *Nucleic Acids Res.* 17, 4937–4946
34. Pastinen, T. *et al.* (1997) Minisequencing: a specific tool for DNA analysis and diagnostics on oligonucleotide arrays. *Genome Res.* 7, 606–614
35. Shendure, J. *et al.* (2005) Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* 309, 1728–1732
36. Mitra, R.D. and Church, G.M. (1999) *In situ* localized amplification and contact replication of many individual DNA molecules. *Nucleic Acids Res.* 27, e34
37. Rhoads, A. and Au, K.F. (2015) PacBio sequencing and its applications. *Genom. Proteomics Bioinform.* 13, 278–289
38. Branton, D. *et al.* (2008) The potential and challenges of nanopore sequencing. *Nat. Biotechnol.* 26, 1146–1153
39. Porreca, G.J. (2010) Genome sequencing on nanoballs. *Nat. Biotechnol.* 28, 43–44
40. Pfeifer, G.P. and Riggs, A.D. (1996) Genomic sequencing by ligation-mediated PCR. *Mol. Biotechnol.* 5, 281–288
41. Drmanac, S. *et al.* (1998) Accurate sequencing by hybridization for DNA diagnostics and individual genomics. *Nat. Biotechnol.* 16, 54–58
42. Adams, M.D. *et al.* (1994) A model for high-throughput automated DNA sequencing and analysis core facilities. *Nature* 368, 474–475
43. Zhang, J. *et al.* (2019) The International Cancer Genome Consortium Data Portal. *Nat. Biotechnol.* 37, 367–369
44. 100,000 Genomes Project Pilot Investigators *et al.* (2021) 100,000 Genomes Pilot on rare-disease diagnosis in health care – preliminary report. *N. Engl. J. Med.* 385, 1868–1880
45. Kaplow, I.M. *et al.* (2023) Relating enhancer genetic variation across mammals to complex phenotypes using machine learning. *Science* 380, eabm7993
46. Watson, A. *et al.* (1998) Technology for microarray analysis of gene expression. *Curr. Opin. Biotechnol.* 9, 609–614
47. Falony, G. *et al.* (2016) Population-level analysis of gut microbiome variation. *Science* 352, 560–564
48. Yan, F. *et al.* (2020) From reads to insight: a hitchhiker's guide to ATAC-seq data analysis. *Genome Biol.* 21, 22
49. Schuster, S.C. (2008) Next-generation sequencing transforms today's biology. *Nat. Methods* 5, 16–18
50. Zhong, W. *et al.* (2021) Next generation plasma proteome profiling to monitor health and disease. *Nat. Commun.* 12, 2493
51. Darmanis, S. *et al.* (2011) ProteinSeq: high-performance proteomic analyses by proximity ligation and next generation sequencing. *PLoS One* 6, e25583
52. Sandberg, R. (2014) Entering the era of single-cell transcriptomics in biology and medicine. *Nat. Methods* 11, 22–24
53. Chen, A. *et al.* (2022) Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* 185, 1777–1792.e21
54. McGuire, A.L. *et al.* (2008) Research ethics and the challenge of whole-genome sequencing. *Nat. Rev. Genet.* 9, 152–156
55. Rosenquist, R. *et al.* (2022) Clinical utility of whole-genome sequencing in precision oncology. *Semin. Cancer Biol.* 84, 32–39
56. Hasin, Y. *et al.* (2017) Multi-omics approaches to disease. *Genome Biol.* 18, 83
57. Cox, J. and Mann, M. (2007) Is proteomics the new genomics? *Cell* 130, 395–398
58. Liu, X. and Locasale, J.W. (2017) Metabolomics: a primer. *Trends Biochem. Sci.* 42, 274–284
59. Ellenberg, J. *et al.* (2018) A call for public archives for biological image data. *Nat. Methods* 15, 849–854
60. Gold, L. *et al.* (2012) Advances in human proteomics at high scale with the SOMAscan proteomics platform. *New Biotechnol.* 29, 543–549
61. Candia, J. *et al.* (2022) Assessment of variability in the plasma 7k SomaScan proteomics assay. *Sci. Rep.* 12, 17147
62. Shalek, A.K. *et al.* (2014) Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature* 510, 363–369
63. Stahl, P.L. *et al.* (2016) Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* 353, 78–82
64. Mitra, R.D. *et al.* (2003) Fluorescent *in situ* sequencing on polymerase colonies. *Anal. Biochem.* 320, 55–65
65. Mortazavi, A. *et al.* (2008) Mapping and quantifying mammalian transcriptomes by RNA-seq. *Nat. Methods* 5, 621–628
66. 1000 Genomes Project Consortium *et al.* (2010) A map of human genome variation from population-scale sequencing. *Nature* 467, 1061–1073
67. Rozenblatt-Rosen, O. *et al.* (2017) The Human Cell Atlas: from vision to reality. *Nature* 550, 451–453