7 Next-Generation Sequencing Methods

T. Rajesh[1,*], M. Jaya[2]

[1]THE UNIVERSITY OF NEW SOUTH WALES, SYDNEY, NSW, AUSTRALIA; [2]PSG INSTITUTE OF MEDICAL SCIENCES AND RESEARCH, COIMBATORE, TAMIL NADU, INDIA

7.1 Introduction

The dideoxy sequencing method developed by Sanger in 1977 has been the most commonly used DNA sequencing technique. This method is based on the DNA-dependent polymerase synthesis of a complementary strand in the presence of 2′-deoxynucleotides (dNTPs) and 2′,3′-dideoxynucleotides (ddNTPs). Whenever a ddNTP is incorporated during DNA synthesis, the reaction is terminated. When the resulting products are resolved on a polyacrylamide gel by electrophoresis, the terminal ddNTPs are identified and the DNA sequence of the template strand is revealed [1,2]. Following its inception, several modifications such as the use of fluorescently labeled dNTPs, capillary gel electrophoresis, and capillary arrays were incorporated. Limitations that remained with the Sanger sequencing method were the requirements of cloning and in vivo amplification of DNA fragments to be sequenced, the cost of fluorophores, and the involvement of intensive labor. However, in 2001, the complete human genome was sequenced entirely based on the Sanger sequencing reaction using fluorescently labeled dNTPs [3,4]. However, owing to the massive time consumption in terms of sequencing reaction and interpretation of results, the need for a highly robust, automated, and inexpensive preparation process surfaced. This led to the development of next-generation sequencing (NGS) technologies, with major advancements over the automated Sanger sequencing method. In 2005, the first NGS technology, termed pyrosequencing, was released by 454 Life Sciences. In 2006, the Solexa/Illumina sequencing platform was released and commercialized [5–7]. Successively, Applied Biosystems released a newer technology called Sequencing by Oligo Ligation Detection (SOLiD). All these new platforms represented major improvements such as high speed, cell-free library construction, ability to run thousands to millions of sequencing reactions in parallel, direct detection without the need for electrophoresis, and sequencing in real time. Despite their ability to produce only very short reads, NGS

*Corresponding Author.

Current Developments in Biotechnology and Bioengineering: Functional Genomics and Metabolic Engineering
http://dx.doi.org/10.1016/B978-0-444-63667-6.00007-9

technologies have revolutionized genome analysis, and here we briefly discuss the NGS methods that have dominated since 2005.

7.2 Methods of Next-Generation Sequencing

7.2.1 Massively Parallel Signature Sequencing

Massively parallel signature sequencing as a high-throughput sequencing technology was introduced following the completion of the human genome sequence with the aim of addressing the efficiency and accurate identification of sequence variation on a large scale [8]. The basis of this method is the combination of in vitro cloning of template DNA onto microbeads by creating a complex mixture of template–oligonucleotide tag conjugates, followed by cycles of ligation-based DNA sequencing (Fig. 7.1). Initially, a mixture of adapters that includes all possible overhangs with a type IIs restriction endonuclease is annealed to the target sequence, resulting in a single adapter with a perfect complementary overhang being ligated. As each adapter has a unique label, the

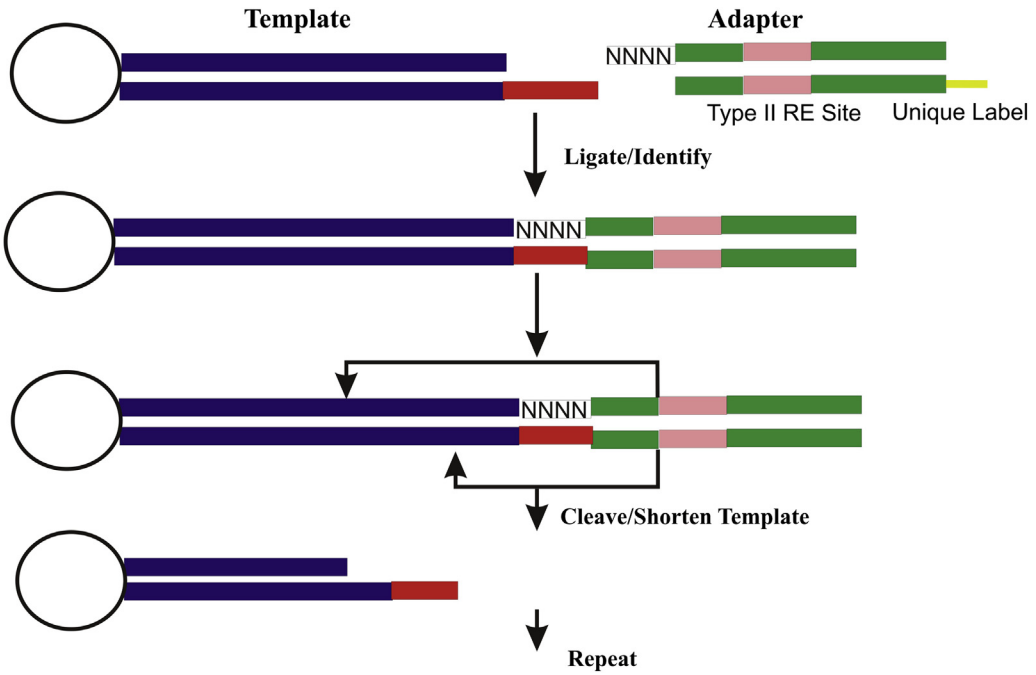

FIGURE 7.1 Schematic representation of steps involved in massively parallel signature sequencing. A template attached to the microbead with an overhang (shown in red (gray in print versions)) is exposed to adapters with a type II restriction site and a unique label. Once the adapter with a perfect complementary sequence is ligated, the sequence overhang it represents is identified. Cleavage at a type II restriction site now shortens the template sequence to be ligated to another adapter. The entire process is repeated until the complete template sequence is identified.

overhangs they represent are first detected after ligation. As the numbers of oligonucleotide tags are presented at a magnitude of 100 times more than that of the template, it is ensured that every template sample gets conjugated to a unique tag. By monitoring successful adapter ligations onto a surface of microbeads in a flow cell, signatures are obtained. Millions of microbeads are employed in this process and as a single copy of the template is attached to each bead, the subsequent sequencing reaction results in the generation of millions of signature sequences. Later, cleavage with a type IIs restriction endonuclease further exposes other bases for identification in subsequent cycles. As each microbead is subjected to successive cycles of ligation, signature identification, and cleavage, the use of the flow cell ensures that all the microbeads remain in a closely packed monolayer. Fluorescent signals from the array of microbeads are acquired onto a CCD (charge-coupled device) camera leading to a digital representation of each microbead, and image-processing software then tracks the positions of fluorescent signals from individual microbeads in the flow cell.

7.2.2 Polony Sequencing

Polony sequencing technology was initially developed with the aim to address the cost efficiency of genome resequencing. This automated nonelectrophoretic DNA sequencing method comprises four basic steps that include construction of an in vitro paired-tag library from genomic DNA, generation of polonies by clonal amplification of library molecules on microbeads by emulsion polymerase chain reaction (PCR), enrichment and two-dimensional arraying of amplicon beads within an acrylamide gel matrix onto a microscope cover glass, and generation of short sequence reads by sequencing by ligation [9] (Fig. 7.2). For generation of a shotgun paired-tag library, sheared DNA fragments of 135 bp are attached to mate-paired 17- or 18-bp tags of target genomic sequence that are flanked and separated by universal sequences complementary to the amplification or sequencing primers employed in subsequent steps. This results in a library containing millions of unique mate-paired sequences. Amplification of the library is then performed by emulsion PCR in which a biotinylated or fluorescence-labeled PCR primer is immobilized on streptavidin-coated or paramagnetic beads. This results in the generation of billions of beads owing to the emulsification of a PCR mix in mineral oil, which acts as a microreactor. Within each bead with an immobilized PCR primer and single trapped template molecule, amplification occurs, resulting in thousands of single-stranded copies of the same PCR product. To eliminate the population of beads that do not bear an amplified template, an enrichment step is performed wherein either fluorescence-activated cell sorting (if fluorescently labeled primers are employed) or hybridization to capture beads that have amplified product is performed. The amplified and enriched beads are then arrayed two-dimensionally on microscope slides on which individual beads settle into a single monolayer, thereby ensuring that the DNA-bearing beads are accessible to reagents for enzymatic sequencing reactions. Finally, for DNA sequencing by ligation, an anchor primer is first

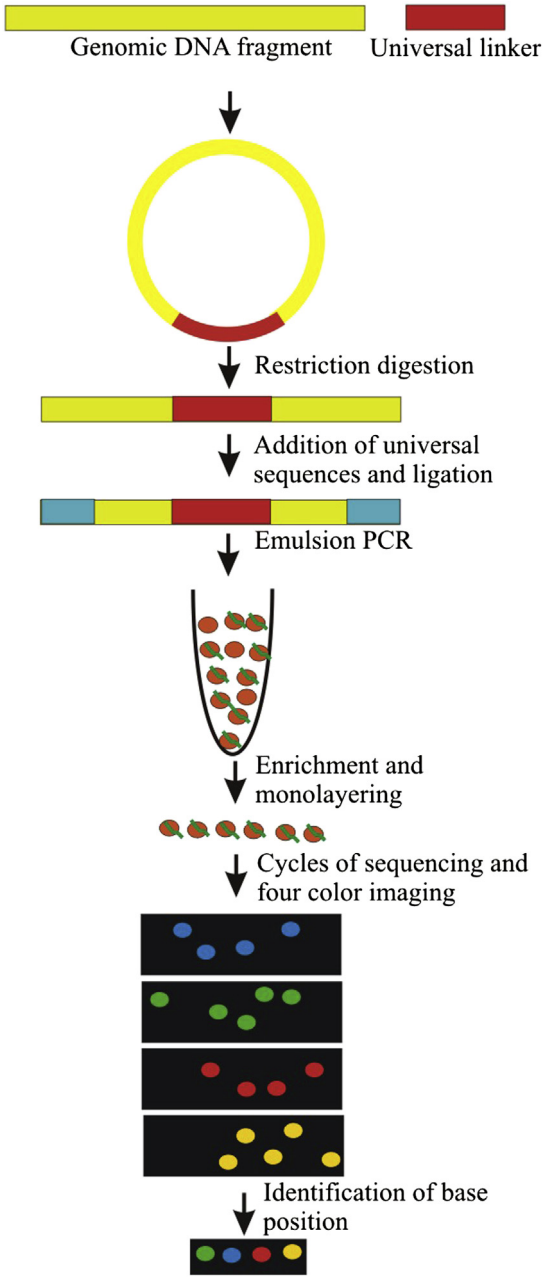

FIGURE 7.2 Steps involved in polony sequencing. Initially, size-selected genomic fragments (yellow (light gray in print versions)) are circularized with the aid of a universal linker or mate pair tags (red (dark gray in print versions)). Further, universal sequences (blue (gray in print versions)) with complementary regions for sequencing primers or amplification primers are then added to each fragment. Following enrichment by emulsion PCR and monolayering by immobilization onto flow cells automated sequencing is performed. At each sequencing cycle, four-color imaging is performed across several hundred raster positions to determine the sequence of each amplified bead at a specific position.

hybridized to a known sequence, or a primer site, within the single-stranded template. A ligation reaction is then performed with a pool of completely degenerate, fluorescently labeled nonanucleotides. During the ligation reaction, each bead becomes tagged with a fluorophore that indicates the identity of the base present at the position to be identified. With successive four-color fluorophore imaging, the anchor primer/nonamer complexes are stripped off, resulting in a new cycle by replacing the anchor primer. Now as a new mixture of fluorescently tagged nonamers is introduced, the query position one base farther into the genomic DNA tag is identified. The entire process results with a sequence read of 26 bp per bead, which is then normalized.

7.2.3 454 Pyrosequencing

A novel method of DNA sequencing technology termed pyrosequencing was developed by 454 Life Sciences and later commercialized by Roche Diagnostics and is widely recognized as 454 pyrosequencing [6]. In this method, fragmented template DNA is attached to microbeads using adapters followed by denaturation and strong dilution of the resulting library (Fig. 7.3). Dilution is performed to a magnitude such that each bead contains a maximum of a single DNA fragment. Following dilution of the DNA-containing beads, amplification of individual DNA fragments is performed by emulsion PCR in which beads containing DNA fragments, adapters, and PCR reagents are mixed with emulsion oil to create water droplets in an oil emulsion, resulting in droplets of water containing a single bead with a single DNA fragment attached to it. The PCR then amplifies the DNA inside the water droplets in the oil solution, which represents a clonal colony resulting in 10^6 double-stranded copies of the library fragment in each bead. These beads containing amplified DNA are then loaded onto picotiter plates in such a way that each well contains a single bead and sequencing enzymes. To ensure the presence of a single bead in each well, individual wells are packed with packing beads. Pyrosequencing is then performed in a sequencing machine by the addition of one dNTP in each PCR cycle. The incorporation of a correct nucleotide results in an oxyluciferin signal from luciferase and generation of light, thereby leading to detection of the individual nucleotides added to the nascent DNA. Signals of each dNTP incorporated determine the sequence of the template fragment in each well and the combined data of resulting signals generate sequence readout of all fragments.

7.2.4 Reversible Terminator Sequencing by Synthesis

In this method, for polony generation, a process termed bridge amplification is used wherein template DNA molecules are first attached to adapters on a slide and then amplified. Sequencing is then performed by synthesis [10]. Initially for bridge amplification, randomly fragmented genomic DNA is ligated with adapters and denatured (Fig. 7.4). The resulting single-stranded fragments serve as templates for DNA amplification and are immobilized onto a flow cell surface that ensures high stability of surface-bound template and low nonspecific binding of fluorescently labeled nucleotides. Following DNA immobilization, forward and reverse primers complementary to

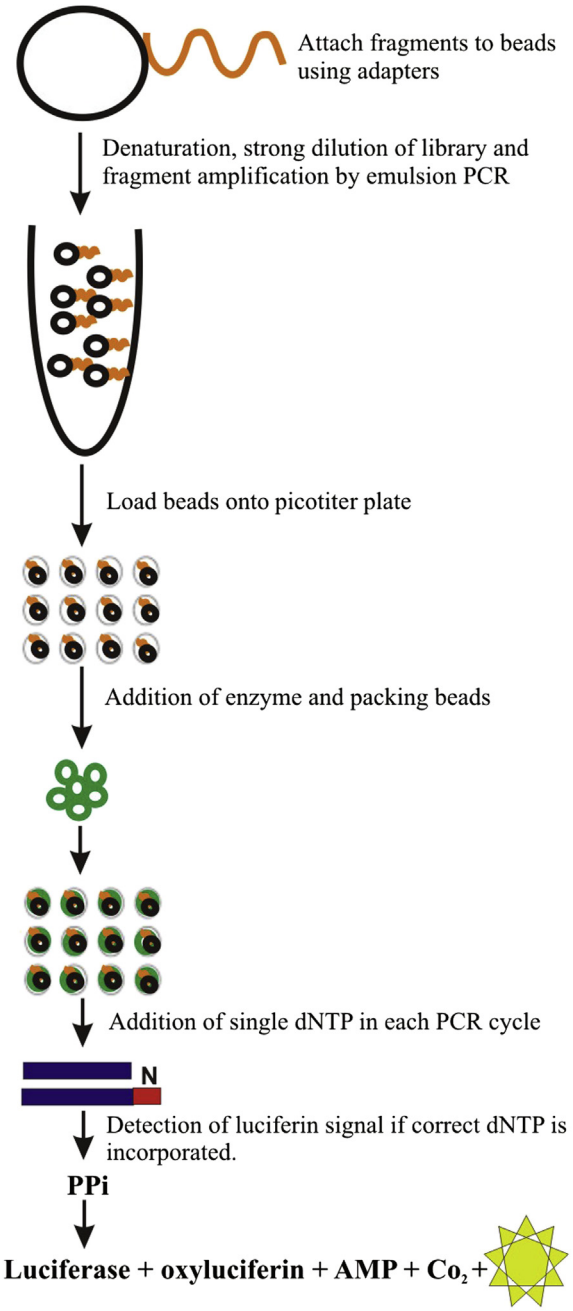

FIGURE 7.3 454 pyrosequencing. In this method, following size selection and attachment of a template fragment to beads, the double-stranded DNA template is strongly diluted, ensuring that the single-stranded fragment is amplified by emulsion PCR in the next step. The beads are then loaded onto picotiter plates and individual wells in the picotiter plate are loaded with enzyme and packing beads. Extension of the template is performed by flooding the wells with one nucleotide at a time. Whenever a correct base is incorporated, a pyrophosphate (PPi) is released, which, in turn, is locally converted via sulfurylase to ATP. Production of ATP enables the redox reaction that converts luciferin to oxyluciferin, releasing light. The luminescence produced is captured after each extension, identifying the base incorporated at a given location within each well. The extension is then iterated through the bases until completion. The yellow (light gray in print versions) symbol represents light.

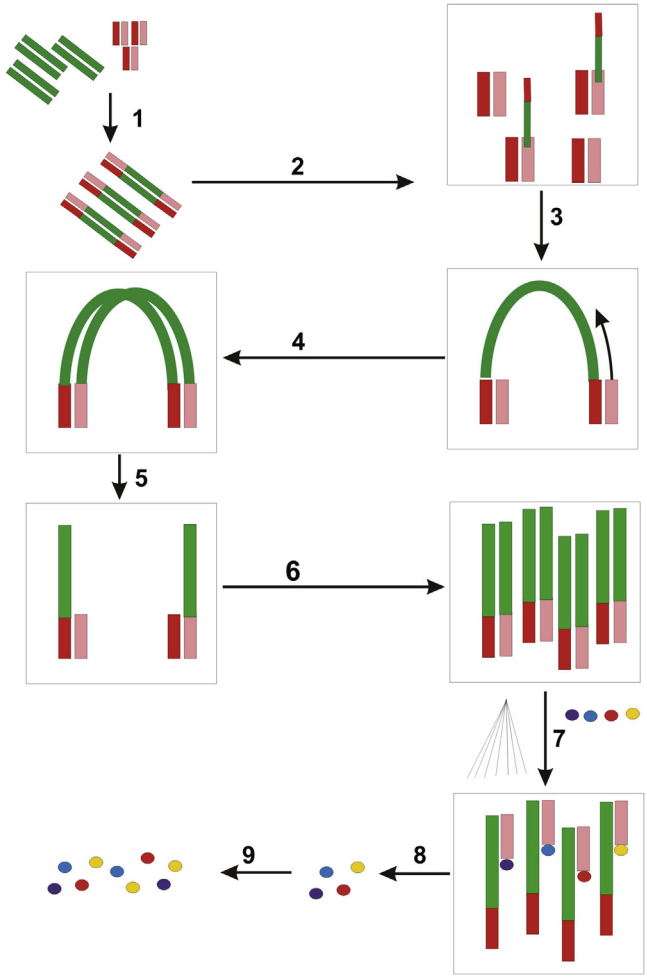

FIGURE 7.4 Steps involved in the reversible terminator sequencing method. (1) In the first step, randomly fragmented genomic DNA is ligated to adapters at both ends. (2) Fragments with attached adapters are then attached to the surface of the flow cell's previously attached primers. (3 and 4) PCR is then performed with unlabeled nucleotides to ensure bridge amplification leading to generation of multiple copies of double-stranded DNA fragments. (5 and 6) Repeated denaturation and amplification results in thousands of newly synthesized strands. (7 and 8) These DNA clusters are then subjected to sequencing, in which the first cycle consists of incorporation of a single fluorescent nucleotide, followed by high-resolution imaging of the entire flow cell. (9) The cycle is repeated, one base at a time, generating a series of images each representing a single base extension at a specific cluster. In silico superimposition of these images reconstructs the sequence of each fragment that generated a cluster.

the adapter sequence with a cleavage site in one among them are introduced into the library preparation step. Once unlabeled nucleotides are added, the enzyme incorporates these nucleotides to build double-stranded bridges on the flow cell surface. Cycles of denaturation and double-strand synthesis are repeated to achieve complete

amplification wherein several million dense clusters of double-stranded DNA are generated in each channel of the flow cells. Nucleotide sequence identification is then initiated by addition of four labeled reversible terminators, primers, and DNA polymerase. Laser excitation leads to fluorescence emission from each cluster and an image is captured that corresponds to the first base. Several cycles of sequencing over multiple chemistry cycles are repeated to determine the sequence of bases in each fragment, one base at a time.

7.2.5 Sequencing by Oligonucleotide Ligation Detection

SOLiD is an extension of previously developed nucleic acid sequencing chemistries in which a two-base encoding method utilizing a two-base color-encoding schema is employed for sequence identification of amplified fragments [11]. As with any other sequencing method, initial preparation of the in vitro sequencing library involves fragmentation of DNA samples to 400−850 bp, end repair, and ligation of two different DNA adapters to the ends of the library fragments. The fragments are then immobilized onto an excessive number of beads leading to single fragments being bound to individual beads. Emulsion PCR is then performed in microreactors for clonal amplification of fragments onto the beads. In the next step, the beads are deposited onto a glass slide and then the bases are read by probing the beads with mixtures of 5′-fluorescently labeled octamer probes of which the last two 3′ bases are known, whereas the rest are degenerate. With 16 possible combinations of two bases, there is a limited number of fluorescent dyes that can be used. Therefore, probes containing certain combinations of two 3′ (interrogatory) bases (di-base) are labeled with the same dye. The sequencing begins by adding an oligonucleotide primer complementary to the adapter, with its last 3′ base annealed to the last base of the adapter, i.e., an adapter−template junction. Then, a pool of probes labeled in the same color is added. If the two interrogatory bases anneal to a complementary sequence, DNA ligase ligates the probe to the primer and the fluorescent signal is recorded. Then, the last three 5′ bases of the probe are cleaved, followed by addition of the next pool of probes, labeled with the second dye. As the addition of four pools of probes is done n times, the ligated probes and oligonucleotide primer are removed and the new ones are annealed (Fig. 7.5). Because the new primer is shifted one base toward the 5′ end, the first of the two bases with the interrogated previous primer, the 5′ base, will be interrogated again, along with the new 5′ base. Resetting of the primer is done four times, thus enabling interrogations of all bases, including those covered by degenerate bases. Based on the order of the four colors, one for each base, a computer is able to reconstruct the sequence of bases. Using this approach, 50-bp reads can be obtained with the key advantage that each base is interrogated twice. The glass slides containing the beads bearing a DNA strand to be sequenced can be segmented in up to eight chambers to facilitate upscaling the number of samples to be analyzed. Such sequencing runs yield 2−4 Gb of DNA sequence data. Once the reads with quality values are base called, low-quality sequences can be removed. Further alignment of the reads to a reference genome enables a second stage of quality evaluation called two-base

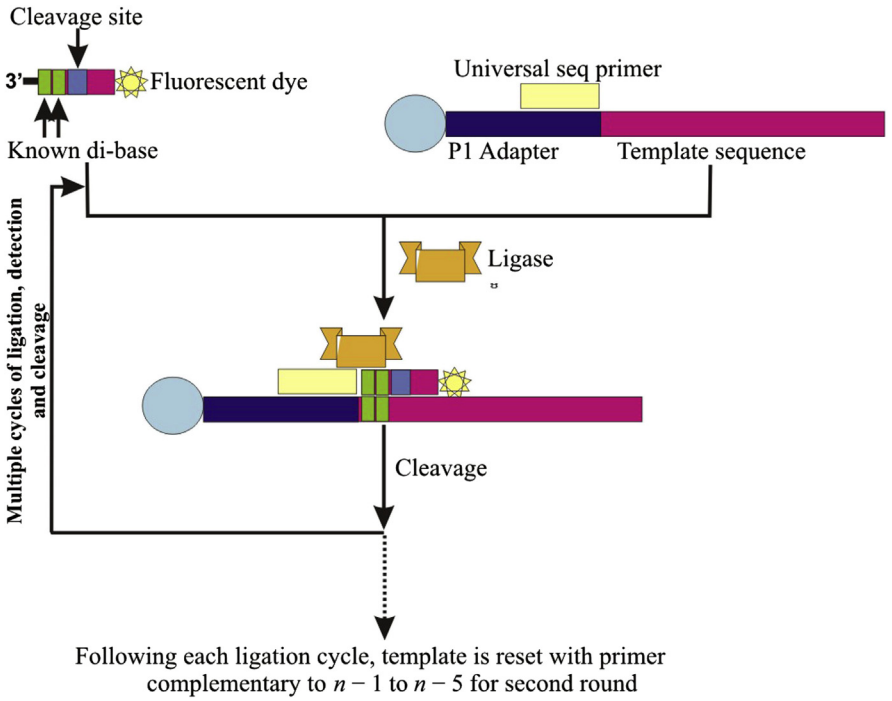

FIGURE 7.5 SOLiD technology. In this method, fragments in the DNA library attached to magnetic beads have a universal P1 adapter sequence ensuring that the starting sequence of every fragment is known and identical. Universal sequencing primers are then allowed to hybridize to the adapter sequence within the library template. A set of four fluorescently labeled di-base probes is added to the DNA fragments bound to beads, and each probe competes to ligate to the sequencing primer. Specificity of the di-base probe allows interrogation of every first and second base in each ligation reaction. Following a series of ligation cycles, the extension product is removed and the template is reset with a primer complementary to the $n - 1$ position for a second round of ligation cycles. Multiple cycles of ligation, detection, and cleavage are repeated until eventually the length of the fragment is read.

encoding to differentiate true single-base variants from base-calling errors. The quality and length obtained from SOLiD are comparable with those of other NGS methods, as the ligation-based method in the SOLiD system requires a complex panel of labeled oligonucleotides, and sequencing proceeds by off-set steps.

Minor disadvantages include more processing time owing to the dual base encoding method, slower sequencing chemistry, and lower read length. SOLiD has been widely applied in whole-genome resequencing, targeted resequencing, transcriptome research (gene expression profiling, small RNA analysis, and whole transcriptome analysis), and epigenome studies (like chromatin immunoprecipitation–sequencing and methylation). Like other NGS systems, SOLiD's computational infrastructure is expensive and not trivial to use; it requires an air-conditioned data center, computing cluster, skilled personnel in computing, distributed memory cluster, fast networks, and a batch queue system. The operating system used by most researchers is GNU/LINUX. Each solid sequencer run takes 7 days and generates around 4 TB of raw data.

7.2.6 Single-Molecule Real-Time Sequencing by Synthesis

Single-molecule real-time sequencing is yet another method of DNA sequencing by synthesis approach. In this method, DNA is synthesized in small microchips with well-like containers by zero-mode waveguides [12]. The zero-mode waveguide by itself is a confined circular structure with holes of ~70 nm in diameter and ~100 nm in depth. Nucleotide bases employed for DNA synthesis have a fluorescent dye molecule attached to their phosphate group. As a DNA polymerase initiates DNA synthesis within each zero-mode wave structure, the fluorescent dye molecule attached to each nucleotide is cleaved off, which is detected by the detector (Fig. 7.6). Once detected, instant diffusion of the cleaved fluorescent dye reduces its detection limit so that its fluorescent signal is no longer detected as further DNA synthesis occurs. The main advantage of this method is that the need for initial template amplification is eliminated. Pacific Biosciences first commercialized this technology.

7.2.7 Ion Torrent—Sequencing by Synthesis

Ion torrent is an amplicon sequencing technology, first introduced by Life Technologies. It is known to be simple, fast, scalable, and cost-effective [13]. In this method, sequence templates are generated on beads via emulsion PCR resulting in oil–water emulsion spheres. As each sphere contains one library molecule and reagents needed for amplification, an emulsion PCR results in amplification of individual fragments to millions of identical copies that are bound to the beads, allowing ultimate detection of the signal. A typical ion torrent chip consists of a flow compartment and microarrayed wells with built-in solid-state pH sensors. The basic principle of this technology relies on the release of a proton (H^+ ion) during the PCR in a DNA molecule, thereby leading to a change in local pH (Fig. 7.7). As individual dNTPs are applied in multiple cycles in consecutive order, a change in pH is detected if an applied base is at that position in the template. The presence of an identical stretch of bases results in a corresponding fold change in pH that is detected and interpreted accordingly.

In this method, high-quality reads of 225 bp are produced with 99% accuracy. The simplicity of this sequencing method allows it to be applied to sequence the template in both directions using a single chip. To achieve this, once the polymerase extension of the forward strand is done, the template for the second read is prepared directly on the ion chip by creating a nick in the original template resulting in a primer for the second read. Now the polymerase could extend the primer in the reverse direction, thus enabling sequencing in the reverse direction.

7.2.8 Clonal DNA Amplification in Solution—DNA Nanoballs

One of the newest ways of DNA sequencing in the realm of NGS is DNA nanoball sequencing. This method relies on rolling circle replication to amplify small fragments of genomic DNA into DNA nanoballs, followed by sequencing by ligation as in the case of

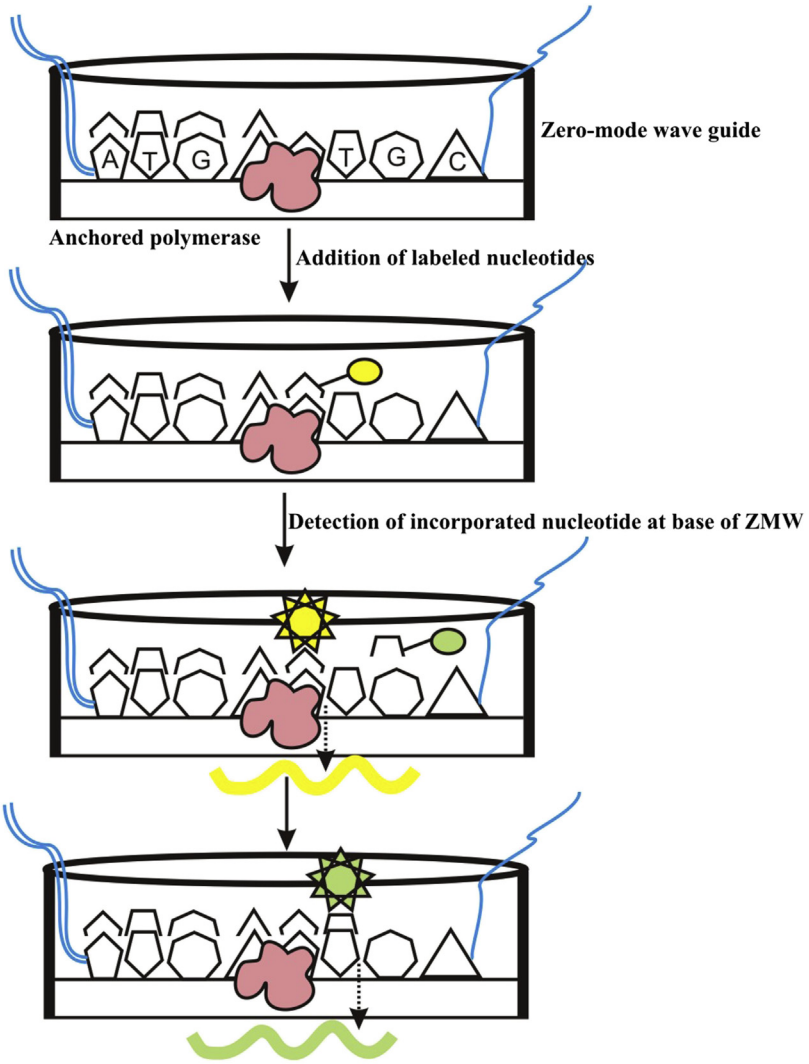

FIGURE 7.6 Single-molecule real-time sequencing. Single-molecule real-time sequencing utilizes a zero-mode waveguide (ZMW), which forms a smart cell. Within each cell, a DNA template–polymerase complex is immobilized at the bottom. Phospholinked nucleotides are then introduced into the ZMW chamber and each of the four nucleotides is labeled with a different colored fluorophore. When a nucleotide is incorporated by the DNA polymerase, the fluorescent tag is cleaved off and detected and diffuses out of the observation area of the ZMW to where its fluorescence is no longer observable. Base calling is done according to the corresponding fluorescence of the dye. Repeated cycles of nucleotide incorporation and detection ensure sequencing of the template DNA fragment.

SOLiD technology [14,15]. As shown in Fig. 7.8, initially, template DNA to be sequenced is sheared into random fragments and then a short DNA fragment called a primer is added to form a closed DNA loop. In the next step, the so-formed DNA loop is digested with a restriction enzyme that cuts these loop at specific locations recognizing 13-bp

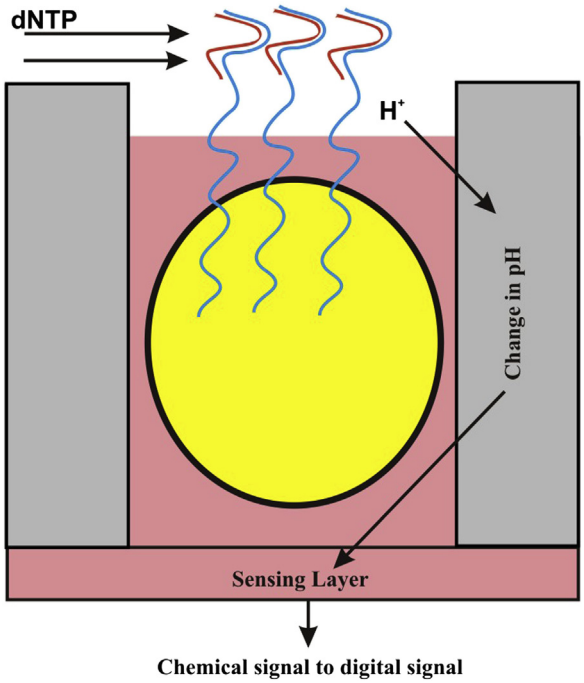

FIGURE 7.7 Principle behind ion torrent technology. As represented in a simplified version of an individual well, dNTPs are added to the bead (yellow (light gray in print versions)) containing a DNA template. When the added base is complementary, it is incorporated. This results in the release of a proton and a local pH change occurs. A sensor present at the base of the well then converts this chemical signal to a digital signal leading to the identification of the base at that position.

sequences. Once again a second set of primers is added to be ligated in the cleaved site of the first formed DNA loop. The process is repeated until four such primers are ligated to form a concatamer of template DNA–primer loop. A special DNA polymerase similar to phi29 then replicates these DNA loops on the order of n times by the rolling circle model of replication resulting in the formation of DNA nanoballs. Sequencing of these nanoballs is then performed by the ligation method. At each copy of DNA in the nanoball, a five-base fluorescently labeled DNA fragment complementary to the template DNA is ligated. Once the fragment is ligated, a signal is produced, which is then recorded, and the entire five-base DNA fragment is removed. After successive cycles of probing with DNA fragments and ligation, the base order of the template DNA is determined through the fluorescence from the ligated and bound probes. The amplification step in DNA nanoball sequencing avoids the cost and challenges of sequencing methods that rely on single fluorophore measurements used by single-molecule sequencing systems.

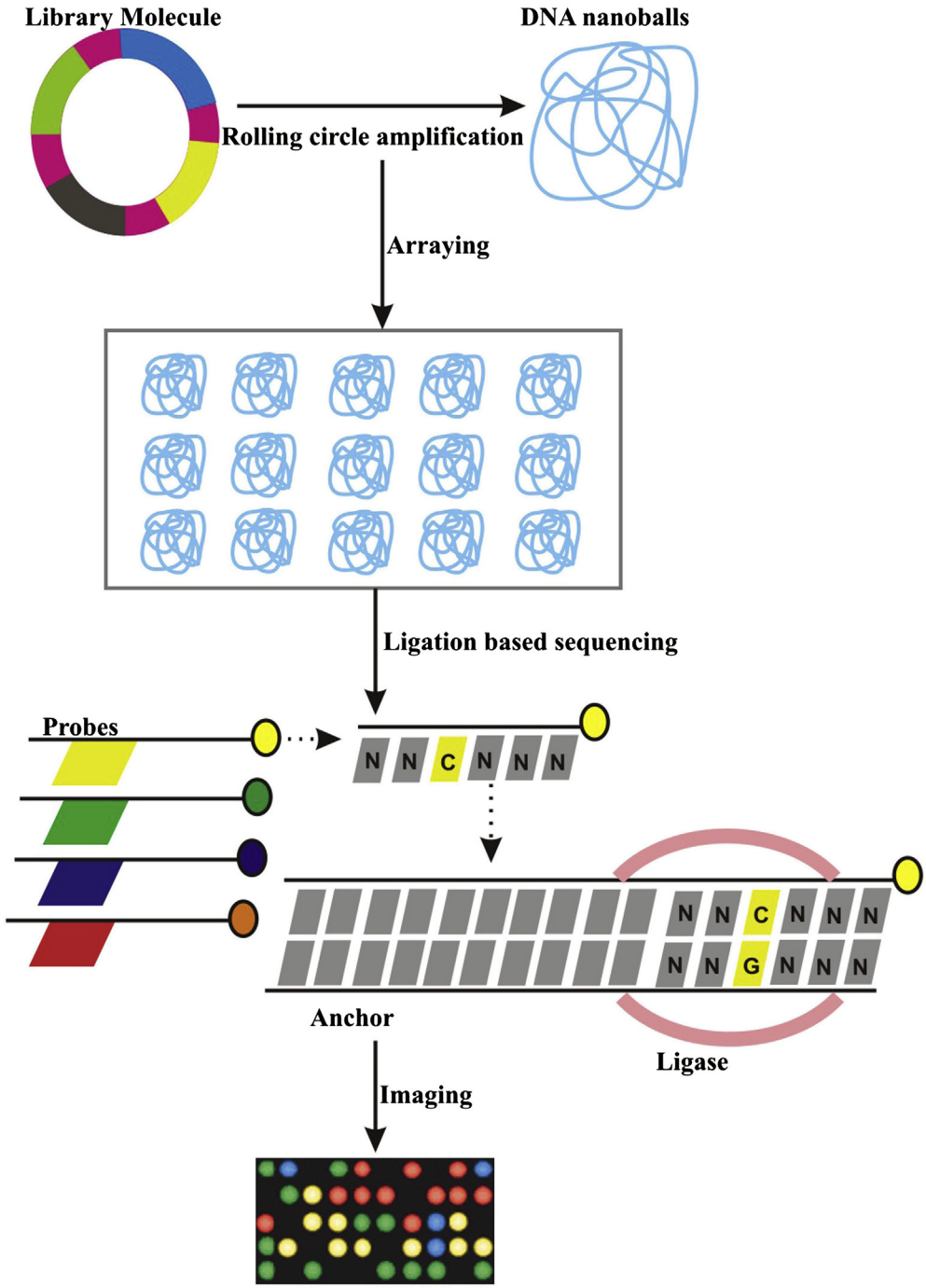

FIGURE 7.8 DNA nanoball sequencing. Sequencing by this method is performed in four steps. In the first step, genomic DNA is converted into library molecules by using DNA ligase and a set of four primers. Rolling circle amplification results in generation of concatamers of DNA nanoballs in the billions. These DNA nanoballs are added to a silicon slide that contains a grid-like pattern of binding sites. Ligation-based sequencing chemistry results in the interrogation of bases in the genomic DNA library. Digital imaging of signals from patterned arrays is performed after each sequencing reaction ensuring identification of the nucleotide sequence of that fragment.

7.3 Comparison of Sequencing Methods and Their Applications

NGS approaches have emerged as dominant genomics technologies in comparison to the Sanger sequencing method, mainly because of their cost-effectiveness and a number of wide applications. A typical massively parallel sequencing system shares good performance on throughput, accuracy, and cost compared to Sanger sequencing. This system exhibits better performance and its own advancements in terms of read length, accuracy, applications, consumables, labor requirement, and informatics infrastructure. The pyrosequencing method for genome sequencing has the capability of delivering explicit sequence information within minutes. This is an additional advantage that makes it an ideal choice for genetic analysis in clinical research. The output data from pyrosequencing is of optimal standards for genetic tests and is the best possible assurance of correct genetic tests. Currently, pyrosequencing is able to produce the longest reads of any NGS system, about 700 bp, and it approaches those generated by Sanger chemistry. Polony sequencing has been tested in a bacterial genome and the sequence read length was about 13 bases per colony [16]. Accounting for various other advantages, reversible termination sequencing technology has been widely accepted and used in multiple NGS platforms [10]. The application of the reversible termination strategy effectively solves the problem of accurately identifying homopolymeric runs (such as poly(A)), which is difficult to achieve using pyrosequencing technology. It has been reported that reversible termination sequencing could accurately read out more than 18 consecutive As [17,18]. SOLiD has similar throughput and cost per base compared to Illumina. It also has the best raw accuracy among the commercial NGS systems. It has the lowest error rates (−0.01) owing to two-base encoding. However, this technology is still limited by short read lengths (350 bp). Ion torrent promises higher output and longer reads. Nanopore is highly potential for long read lengths and short sequencing times. Zhao et al. reported that a single-nucleotide polymorphism could be detected by a change in the threshold voltage of a nanopore [19]. Current ion torrent chips can yield several hundred thousand reads with an average length of about 100 bp in less than 2 h. NGS technologies have already been used for various applications, ranging from whole-genome sequencing to exome sequencing, target resequencing, single-nucleotide polymorphisms, structural variation discovery, mRNA and noncoding RNA profiling, and protein–nucleic acid interaction assays, mutation screening, methylation profiling, and chromatin remodeling. NGS technologies are becoming a potential tool for gene expression analysis, especially for those species having reference genome sequences already available.

7.4 Future Perspectives

Advances in NGS methods have certainly revolutionized genome research. However, the need for further low-cost and faster methods is much expected. Considering the need to prepare a library and perform high-efficiency PCR, technologies with library- and PCR-

free methods are expected from potential users. Future biomedical research to an extent relies completely on advancements in NGS methods, for instance, in terms of personalized medicine. In this context, it should not be a matter of just producing better instruments for sequencing, but also technologies with the ability to produce longer and accurate reads in a very period of time. Further, technologies that would generate data that could be readily accessible without the need for a bioinformatics expert are highly desirable. Methods of such kind with robust accuracy and consistency would be ideal for customer-customized research.

References

[1] F. Sanger, A.R. Coulson, A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase, J. Mol. Biol. 94 (1975) 441–448.

[2] F. Sanger, S. Nicklen, A.R. Coulson, DNA sequencing with chain-terminating inhibitors, Proc. Natl. Acad. Sci. U.S.A. 74 (1977) 5463–5467.

[3] International Human Genome Sequencing Consortium, Initial sequencing and analysis of the human genome, Nature 409 (2001) 860–921.

[4] International Human Genome Sequencing Consortium, The sequence of the human genome, Science 291 (2001) 1304–1351.

[5] A. Ronaghi, M. Uhlen, P. Nyren, A sequencing method based on real-time pyrophosphate, Science 281 (1998) 363–365.

[6] M. Margulies, M. Egholm, W.E. Altman, S. Attiya, J.S. Bader, L.A. Bemben, et al., Genome sequencing in open microfabricated high density picoliter reactors, Nature 437 (2005) 376–380.

[7] K.L. Patrick, 454 Life Sciences: Illuminating the future of genome sequencing and personalized medicine, Yale J. Biol. Med. 80 (2007) 191–194.

[8] S. Brenner, M. Johnson, J. Bridgham, G. Golda, D.H. Lloyd, D. Johnson, S. Luo, S. McCurdy, M. Foy, M. Ewan, R. Roth, D. George, S. Eletr, G. Albrecht, E. Vermaas, S.R. Williams, K. Moon, T. Burcham, M. Pallas, R.B. DuBridge, J. Kirchner, K. Fearon, J. Mao, K. Corcoran, Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays, Nat. Biotechnol. 18 (2000) 630–634.

[9] J. Shendure, G.J. Porreca, N.B. Reppas, X. Lin, J.P. McCutcheon, A.M. Rosenbaum, M.D. Wang, K. Zhang, R.D. Mitra, G.M. Church, Accurate multiplex polony sequencing of an evolved bacterial genome, Science 309 (2005) 1728–1732.

[10] D.R. Bentley, S. Balasubramanian, H.P. Swerdlow, G.P. Smith, J. Milton, C.G. Brown, et al., Accurate whole human genome sequencing using reversible terminator chemistry, Nature 456 (2008) 53–59.

[11] E.R. Mardis, The impact of next-generation sequencing technology on genetics, Trends Genet. 24 (2008) 133–141.

[12] J. Eid, A. Fehr, J. Gray, K. Luong, J. Lyle, G. Otto, P. Peluso, D. Rank, P. Baybayan, B. Bettman, A. Bibillo, K. Bjornson, B. Chaudhuri, F. Christians, R. Cicero, S. Clark, R. Dalal, A. deWinter, J. Dixon, M. Foquet, A. Gaertner, P. Hardenbol, C. Heiner, K. Hester, D. Holden, G. Kearns, X. Kong, R. Kuse, Y. Lacroix, S. Lin, P. Lundquist, C. Ma, P. Marks, M. Maxham, D. Murphy, I. Park, T. Pham, M. Phillips, J. Roy, R. Sebra, G. Shen, J. Sorenson, A. Tomaney, K. Travers, M. Trulson, J. Vieceli, J. Wegener, D. Wu, A. Yang, D. Zaccarin, P. Zhao, F. Zhong, J. Korlach, S. Turner, Real-time DNA sequencing from single polymerase molecules, Science 323 (2009) 133–138.

[13] N. Rusk, Torrents of sequence, Nat. Methods 8 (2011) 44.

[14] R. Drmanac, A.B. Sparks, M.J. Callow, A.L. Halpern, N.L. Burns, B.G. Kermani, P. Carnevali, I. Nazarenko, G.B. Nilsen, G. Yeung, F. Dahl, A. Fernandez, B. Staker, K.P. Pant, J. Baccash, A.P. Borcherding, A. Brownley, R. Cedeno, L. Chen, D. Chernikoff, A. Cheung, R. Chirita, B. Curson, J.C.

Ebert, C.R. Hacker, R. Hartlage, B. Hauser, S. Huang, Y. Jiang, V. Karpinchyk, M. Koenig, C. Kong, T. Landers, C. Le, J. Liu, C.E. McBride, M. Morenzoni, R.E. Morey, K. Mutch, H. Perazich, K. Perry, B.A. Peters, J. Peterson, C.L. Pethiyagoda, K. Pothuraju, C. Richter, A.M. Rosenbaum, S. Roy, J. Shafto, U. Sharanhovich, K.W. Shannon, C.G. Sheppy, M. Sun, J.V. Thakuria, A. Tran, D. Vu, A.W. Zaranek, X. Wu, S. Drmanac, A.R. Oliphant, W.C. Banyai, B. Martin, D.G. Ballinger, G.M. Church, C.A. Reid, Human genome sequencing using unchained base reads in self-assembling DNA nanoarrays, Science 327 (2010) 78–81.

[15] J.G. Porreca, Genome sequencing on nanoballs, Nat. Biotechnol. 28 (2010) 43–44.

[16] S. Myllykangas, J. Buenrostro, H.P. Ji, Overview of Sequencing Technology Platforms, in: Bioinformatics For High Throughput Sequencing, 2012, pp. 11–25.

[17] D.R. Bentley, Whole-genome re-sequencing, Curr. Opin. Genet. Dev. 16 (2006) 545–552.

[18] S. Fields, Molecular biology. Site-seeing by sequencing, Science 316 (2007) 1441–1442.

[19] Q. Zhao, G. Sigalov, V. Dimitrov, B. Dorvel, U. Mirsaidov, S. Sligar, A. Aksimentiev, G. Timp, Detecting SNPs using a synthetic nanopore, Nano Lett. 7 (6) (2007) 1680–1685.