A Next-Generation Sequencing Primer—How Does It Work and What Can It Do?

Yuriy O. Alekseyev, PhD1, Roghayeh Fazeli, MD1, Shi Yang, MD, PhD1, Raveen Basran, DPhil1, Thomas Maher, MS1, Nancy S. Miller, MD1, and Daniel Remick, MD1

Abstract

Next-generation sequencing refers to a high-throughput technology that determines the nucleic acid sequences and identifies variants in a sample. The technology has been introduced into clinical laboratory testing and produces test results for precision medicine. Since next-generation sequencing is relatively new, graduate students, medical students, pathology residents, and other physicians may benefit from a primer to provide a foundation about basic next-generation sequencing methods and applications, as well as specific examples where it has had diagnostic and prognostic utility. Next-generation sequencing technology grew out of advances in multiple fields to produce a sophisticated laboratory test with tremendous potential. Next-generation sequencing may be used in the clinical setting to look for specific genetic alterations in patients with cancer, diagnose inherited conditions such as cystic fibrosis, and detect and profile microbial organisms. This primer will review DNA sequencing technology, the commercialization of next-generation sequencing, and clinical uses of next-generation sequencing. Specific applications where next-generation sequencing has demonstrated utility in oncology are provided.

Keywords
diagnostic test, microbiology, mutation, oncology, sequencing

Received August 14, 2017. Received revised February 14, 2018. Accepted for publication February 16, 2018.

Introduction

Although nucleic acid sequencing technology has only existed for about 40 years, the technology represents an outstanding example of progress resulting from continuous improvement and increases in cost efficiency. The newest sequencing technologies are frequently referred to as next-generation sequencing (NGS). The results from NGS testing have been translated into clinical laboratories to produce clinically relevant information that directly impacts patient care. Some molecular tests employ a “one-gene one-test” approach by using specific sets of primers and polymerase chain reaction (PCR) to detect one specific mutation. In contrast, NGS is able to detect thousands or even hundreds of thousands of genetic variants in a single test run. This primer is written to provide an introduction to NGS for those health-care professionals who may have heard of the technology in the lay press or in grand rounds. While not an exhaustive review, it does lay the foundation for understanding the power of this innovative technology.

First-Generation DNA Sequencing Technology

After the discovery of the chemical composition of DNA in the late 19th century, nearly 50 years passed before the structure of DNA was elucidated1 and another quarter of a century elapsed prior to developing methods to sequence DNA.2,3 The principal method published in 1977 involves sequencing by synthesis (SBS) of a radioactively labeled DNA strand complimentary to the interrogated template strand using the dideoxy chain termination method.
technique. The resulting fragments were then analyzed by polyacrylamide gel electrophoresis. This method, known as Sanger sequencing, became the basis for the “first-generation” sequencing technology. The original Sanger sequencing method has subsequently been automated and commercialized. Major innovations include the introduction of fluorescent-labeled nucleotides instead of radioactivity, replacement of gel electrophoresis with capillary electrophoresis, and improvement of the DNA polymerases. Additional progress was achieved through adoption of molecular biology techniques, such as recombinant DNA technology and the PCR, which allowed production and amplification of DNA fragments. The Sanger method-based sequencing technology was used to sequence a number of increasingly large genomes starting with bacteria and phages, and eventually mammalian and human genomes.

One of the major limitations of Sanger sequencing is that only one sequence reaction can be analyzed per electrophoresis lane or capillary tube, hence the necessity to divide the DNA from a biological sample into individual template fragments. This was achieved by randomly cloning the fragmented DNA from a biological sample (by insertion into vectors, transformation of the bacteria, and extraction of pure individual fragments from the resulting colonies). This very labor-intensive process was one of the reasons why the first human genome project took more than 10 years and cost US$2.7 billion. Subsequent improvements allowed another human genome to be sequenced using the same technology for approximately US$10 million.

Despite these advances, the efficiency of this method has approached its limit and further use of this technology was considered time and cost prohibitive. It should be noted that Sanger sequencing remains the gold standard for confirming DNA sequences due to the stability of the technology and is still broadly used for targeted re-sequencing in research and clinical laboratories.

Next-Generation Sequencing

The terms NGS (sometimes subdivided into second- and third-generation sequencing) massively parallel sequencing or high-throughput sequencing usually refers to technologies that allow sequencing without the physical separation of individual reactions into separate tubes, capillaries, or lanes. Instead, the sequencing reactions occur in parallel on a solid surface (such as glass or beads, depending on the technology) and are only spatially separated. Thus, billions of sequencing reactions occur and are analyzed simultaneously, dramatically improving the throughput and decreasing the labor compared to Sanger sequencing. Regardless of the platform, NGS involves several common steps (see reviews for details), which are outlined in Figure 1.

Commercialization of Next-Generation Sequencing Technology

These novel approaches introduced early in the 21st century were rapidly adopted resulting in strong competition in the NGS market. There are several technical differences in the technologies. A glossary of terms used in molecular biology terms is provided in Table 1.

The first commercial NGS technology was introduced in 2004 by 454 Life Sciences (later purchased by Roche). This technology utilized luminescent detection of a pyrophosphate released upon incorporation of a correct nucleotide during SBS and produced relatively long sequences (called “reads”). This technology was used to sequence the genome of James Watson and the price dropped from US$10 million with Sanger sequencing to about US$2 million. Within 2 years, other platforms emerged (Illumina/Solexa and ABI SOLiD); however, they only produced very short reads. Illumina utilizes an SBS technology originally developed by a company called Solexa which uses reversibly terminated fluorescently labeled nucleotides. Illumina scientists managed to significantly increase the sequencing read length and dramatically improve accuracy and

---

Figure 1. Schematic of sample processing for next-generation sequencing. See Table 1 for definitions of specific terms.
| Term                                           | Definition                                                                 |
|------------------------------------------------|---------------------------------------------------------------------------|
| Adaptors                                       | Single-stranded or double-stranded synthetic oligonucleotides that can be ligated to the ends of other DNA or RNA molecules |
| Bridge amplification                           | A polymerase chain reaction (PCR) technique that embeds DNA on an oligo-decorated solid surface for cloning |
| Cluster generation or template generation      | The product for the NGS sequencing step, it is platform dependent          |
| Copy number variation (CNV)                   | Variation in the number of copies of a particular gene compared to a reference standard |
| DNA sequencing                                 | Determining the sequential order of nucleotides in DNA                    |
| DNA fragmentation                              | Separating or breaking DNA strands into pieces                             |
| Emulsion PCR                                   | A PCR technique that is conducted on a bead surface within tiny water bubbles floating on an oil solution |
| Exome                                          | Includes the coding region within the genome and does not include the introns or noncoding regions |
| Fluorescent in situ hybridization (FISH)       | A molecular cyto genetic technique that uses fluorescent probes to bind to parts of DNA that have a high degree of sequence complementarity |
| Fusion gene                                    | A hybrid gene formed from partial or complete sequences of 2 previously separate genes |
| Homopolymer                                    | Repetitive stretch of single-nucleotide types (eg, TTT or GGGGGG)          |
| Hybridization                                  | The bonding of single-stranded DNA or RNA to form double-stranded DNA or RNA |
| Incidental findings                            | Variants identified that are not directly relevant to the diagnostic question |
| Insertion deletion                             | Insertion or deletion of nucleotide base(s) into the genome of an organism |
| Library generation                             | DNA/RNA prepared into a form compatible with the sequencing system used. This includes DNA fragmentation, shearing the DNA into smaller fragments, and adding common adapters to the DNA fragments |
| Microarray                                     | A technology used to detect expression or copy number of many genes simultaneously |
| Point mutation                                 | A mutation that affects a single-nucleotide base                           |
| Productivity                                   | Number of bases sequenced per run                                         |
| Quantitative PCR                               | An extremely sensitive PCR-based laboratory technique that allows the accurate measurement of the amount of specific nucleic acids in a sample |
| Multiplex PCR                                  | Amplification of several different DNA sequences in a single PCR experiment |
| Output                                         | Total length of sequenced genomic region                                  |
| PCR                                            | A technique in which segments of DNA can be amplified, generating thousands to millions of copies of a particular DNA sequence |
| Polyacrylamide gel electrophoresis            | A technique used to separate biological molecules, usually proteins or nucleic acids, based on their molecular weight |
| Pyrosequencing                                 | A method of DNA sequencing by measuring the synthesis of the complementary DNA strand (sequencing by synthesis) |
| Read                                           | Segment of DNA that has been sequenced                                    |
| Reference sequence                             | This is a consensus sequence of the DNA bases of an organism. The NGS sample is compared to the reference sequence to look for alterations |
| Sequencing by synthesis (SBS)                  | A technique in which sequencing is performed by detecting the nucleotide incorporated by a DNA polymerase |
| Sequencing depth                               | Number of times a given nucleotide in the genome has been read during the sequencing run (often referred to as depth of coverage) |
| Sequence coverage                              | Proportion of the targeted genomic region that is actually sequenced (found in the sequences from the generated data) |
| Single nucleotide variant (SNV) and single nucleotide polymorphism (SNP) | Variation of a single nucleotide of a particular gene between individuals (SNV). If this variant is present with some degree of frequency in a population, it referred to as a single-nucleotide polymorphism (SNP) |
| Sanger sequencing (chain termination method)   | A technique for DNA sequencing based on the selective incorporation of chain-terminating deoxy nucleotides (ddNTPs) by DNA polymerase during in vitro DNA replication |
| Targeted sequencing                            | Analyzing a panel of genes related to a disorder                          |
| Template DNA                                   | The noncoding strand of DNA                                               |
| Variant                                        | Alteration of a DNA sequence as compared to the reference sequence that may or may not be associated with a disease state. Classification of variants include: |
|                                               |   - Pathogenic variant: Genetic variation with sufficient evidence to classify it as pathogenic (capable of causing a disease) |
|                                               |   - Likely pathogenic (LP): Genetic variation with strong evidence in favor of its pathogenicity |
|                                               |   - Variant of unknown significance (VUS): Genetic variation that cannot be definitively determined to be benign or pathogenic |
|                                               |   - Likely benign (LB): Genetic variation with strong evidence against its pathogenicity |
|                                               |   - Benign (B): Genetic variation with very strong evidence against its pathogenicity |

(continued)
More recently, Illumina introduced 2 new instruments, the HiSeq X and NextSeq 500. The first is able to sequence a human genome at ×30 coverage for less than US$1000,29,30 and the latter does the same for a slightly higher price but in less than 20 hours. Moreover, a new series of instruments introduced in 2017 (NovaSeq) should reduce the costs by almost another order of magnitude.

A conceptually different sequencing platform called Ion Torrent21 was introduced in 2011. This SBS technology detects the minute changes in pH caused by H⁺ ions released during the incorporation of the correct nucleotide in the microenvironment around the beads with the attached clonally amplified DNA template molecules. Consequently, it does not require fluorescently labeled nucleotides and expensive optics to detect the fluorescence (see the study by Heather and Chain21, Reuter et al22, and Morey et al23 for details). Currently, the most popular applications from this company (now part of ThermoFisher) are targeted disease panels used in clinical settings (eg, cancer).

In the last decade, the amount of sequencing data has increased exponentially, accelerating translational research, clinical usage of genomics findings, and development of new genomics tests to support precision medicine approaches.32,33 Progress in sequencing technologies was also facilitated by the expansion and adoption of improved molecular biology methods. While early protocols required microgram quantities of high-quality nucleic acid, now samples with very low (ie, picogram) quantities of nucleic acid may be sequenced. Sequencing of samples from formalin-fixed paraffin-embedded material has become routine. These advances also led to the discovery of new circulating biomarkers, a revolution in prenatal diagnostics (genetic testing of fetal DNA from mother’s blood samples) and single-cell genomics approaches.38-40

### Comparison of Ion Torrent and Illumina

Illumina has developed an impressive line of instruments that differ in their productivity, speed, and price tags from small benchtop sequencers (producing 1.65-7.5 Gb per run) to production scale systems (producing thousands of Gb of data per run). All these instruments implement similar chemistry with the sequencing performed by synthesis using reversibly terminated fluorescently labeled nucleotides and capturing the fluorescent images after each nucleotide incorporation event. The sequencing data are deconvoluted from the image data based on the color of the labels.

In the clinical setting, the majority of the existing NGS tests provide a limited amount of sequence information. MiSeq is an instrument with relatively low productivity which fits the current need. Depending on the specific test, the instrument may produce from 500 Mb up to 15 Gb of data in 4 to 56 hours. Currently, Illumina produces a validated, Food and Drug Administration (FDA)-regulated custom amplicon kit that enables clinical laboratories to design custom NGS assays for the FDA-approved MiSeqDx and NextSeq550Dx instruments.

Ion Torrent has 3 instruments in their portfolio. The sequence is determined by measuring the change in pH in the microenvironment around the beads with the attached clonally amplified identical template molecules immediately after addition of a nucleotide (one at a time). There is no definitive stop at each position and the synthesis immediately continues in case of repeats on the template. The pH change in such cases is stronger than when just 1 nucleotide is incorporated allowing the calculation of the nucleotides in the repeat. However, this technology is more prone to homopolymer detection and framshift errors.41 The first instrument by this company, the Personal Genome Machine, is also approved by the FDA for clinical NGS tests (Ion PGM Dx). This instrument produces up to 2 Gb of sequencing data (200-400 bp) in 2 to 4 hours. The newer instrument S5 uses the same sequencing approaches and produces up to 15 Gb data (200-400 bp) in 2- to 4-hour runs. Both instruments require additional time and instrumentation for library preparation prior to sequencing.

Currently, Illumina has the largest market share and was the first to obtain FDA approval for their MiSeq instrument. A comparison of the 2 reveals advantages and disadvantages. The initial cost of both instruments is similar. The Ion Torrent will generate sequence data faster than Illumina, an important consideration for a clinical diagnostic test with urgent requests for results, that is, prenatal samples. The Ion Torrent system offers automated library preparation, template preparation, and automated chip loading with the purchase of a separate piece of equipment, the Ion Chef, and does not depend on a technician once the Ion Chef has been loaded. This should lead to more reproducible results by removing the variability between technologists. The manual library preparation workflow for the Illumina occupies the technican’s time and requires considerable molecular biology expertise. The actual applications for both instruments appear to be comparable and both will perform targeted resequencing and whole-exome sequencing (WES). The Illumina system has a lower cost per base of sequence. The FDA-approved Illumina MiSeqDx is able to generate a complete report for its cystic fibrosis assay, while...
the Ion Torrent needs manual analysis. An attractive feature of the Ion Torrent in the cancer field is the ability to generate a report that matches the sequence results with ongoing clinical trials. While this may be done with the Illumina data, the report requires third-party software. The S5 is the latest instrument from Ion Torrent; consequently, there are limited user data for comparison. It should be noted that the Ion Torrent’s underly-
ing technology has not changed with the release of the S5. Comparison of the Illumina and Ion Torrent systems has been published in several articles in areas such as clinical microbiology, germline variant detection, and prenatal testing as well as somatic variant in oncology. Both platforms performed well and the results are comparable.42,43

Requirements for Performing Clinical Next-Generation Sequencing

Next-generation sequencing results for clinical purposes have substantially increased the amount of information available by generating massive amounts of sequence data. As a result, the overall detection rate of disease-causing alterations has grown significantly. Next-generation sequencing has allowed the creation of targeted gene panels that will sequence hundreds of genes at once for less expense compared to the Sanger method or PCR assays. A clinical setting has many factors to consider in choosing a platform compared to a research setting. Factors to consider in a clinical laboratory system include specificity and sensitivity, reproducibility, and analytical accuracy to ensure clinicians receive accurate results.44,45 Next-generation sequencing technology yields massive amounts of data that require substantial analysis to produce a clinically relevant, concise result. This analysis requires appropriate infrastructure including analysis software, data storage, and accessibility.46–48 The report provided to the clinicians needs to be appropriately formatted. For example, a genetics or oncology report should include the classification of the variant (ie, pathogenic), literature describing the reportable variant(s), recommendations for further testing, and for oncology reports, indicate whether the variant is inherited or somatic. The data analysis usually involves massive genetic, genomic, and oncologic bioinformatics research and data analyses. Fortunately, there are jointly proposed guidelines from professional organizations for NGS testing, validation, proficiency testing,42 reporting, and quality assurance/quality control requirements and documentation.49 These articles will help standardize the proper application and interpretation of the NGS data for clinical utilization. Interested parties should refer to the extensive documentation in these articles for additional information.

Next-Generation Sequencing for Hereditary Disorders

Next-generation sequencing testing for hereditary disorders faces technical challenges, data management issues, reporting on incidental findings, and variant interpretation.50 Despite these challenges, NGS has been valuable in identifying the underlying molecular cause of disorders, especially for those diseases which are genetically heterogeneous. In addition, NGS has been valuable in the detection of rare variants after single-gene analysis has been negative or when multigene panels were too labor-intensive and costly for Sanger sequencing.51 For complex, rare phenotypes, NGS has also been a powerful tool in reducing the diagnostic odyssey often required to arrive at a diagnosis.52,53 For example, in Usher syndrome, an autosomal recessive disorder characterized by sensorineural hearing loss, retinitis pigmentosa, and vestibular dysfunction (in a subset of cases),54 NGS has allowed the development of targeted gene panels to survey all causative genes associated with Usher syndrome. Previously, a comprehensive analysis was limited by the labor-intensive, high cost of Sanger sequencing and turnaround times. Providing an earlier diagnosis for children with Usher Syndrome affords the opportunity for earlier medical management for patients and their families.54

Next-generation sequencing has enabled diagnostic laboratories the ability to offer targeted disease panels for genetic disorders, such as connective tissue diseases, in addition to whole-genome sequencing (WGS) and WES. For patients with nonspecific clinical presentations, such as moderate to severe intellectual disability,55 WGS is recommended. The diagnostic rate of WES is approximately 25% to 31%53,56 similar to WGS57; however, WGS has the additional advantage of detecting larger numbers of copy number variations.57,58 Additionally, the reporting of incidental findings, defined as variants unrelated to the primary medical reason for testing,59 need to be addressed in the context of exome and genome sequencing. The American College of Medical Genetics and Genomics updated their guidelines in 2017 for reporting incidental or secondary findings in 59 medically actionable genes in which known or expected pathogenic variants were identified.59 Reporting known (or expected) disease causing mutations in conditions where preventive measures and/or treatments are available highlight the benefit of returning incidental findings to patients.60 However, there are also limitations and challenges in identifying and reporting such findings including the consent process, follow-up diagnostic evaluation, and additional laboratory resources.51,60,61

Next-Generation Sequencing for Detecting Microbial Organisms

The utility of NGS has been demonstrated for several applications involving pathogen biology and genomic epidemiology. These include targeted sequencing and unbiased interrogation of clinical samples for pathogen detection and identification (regardless of whether the organism can be cultivated or is viable), drug resistance profiling, strain typing and epidemiological outbreak investigation, microbiome studies, genomic determinant analysis of microbial functions including metabolism, and comparative ribosomal RNA phylogenetic studies. Next-generation sequencing brings added throughput, sensitivity, and informatics-based prowess to pathogen interrogation. It is emerging as a valuable diagnostic alternative when other
methods fail to identify an organism or cannot decipher complex specimens such as in patients with polymicrobial infections. The use of NGS to detect evidence of *Leptospira* in the cerebral spinal fluid of a critically ill pediatric patient was a landmark case that demonstrated the clinical utility of unbiased NGS to achieve an actionable diagnostic result when other approaches failed, including phenotypic, immunologic, and targeted PCR-based assays. Also, NGS can perform complete de novo genome sequencing for pathogens not yet fully characterized, providing reference genomes for further study. 63

HIV-1 genotyping for drug resistance prediction is a prototypical example of another value added by an NGS approach since it is more sensitive than Sanger sequencing and can detect small percentages of mutant quasi-species of potential importance to clinical management. A caveat is the technical and informatics challenges associated with authenticating minor variant calling in these applications. There are several challenges for pathogen testing such as separating microbial nucleic acid from human DNA, library preparation from non-sterile site samples, de novo sequence assembly of uncharacterized organisms, and assigning clinical significance to microbial sequences. Microbial profiling by NGS for clinical purposes is currently limited to laboratories with the expertise and resources to support independently developed assays since none have yet been commercialized to the extent necessary for widespread adoption. Several academic and commercial reference laboratories now offer NGS services to laboratories without the means to employ NGS technology on their own.

Next-generation sequencing can be applied to comparative microbiome characterizations in healthy and disease states or pre- and postinterventions. Next-generation sequencing characterization of the intestinal microbiome before and after fecal microbiota transplant (FMT) in cases of *Clostridium difficile* colitis has added to our knowledge of microbiome protection, microbial pathogenesis, and therapeutic efficacy of FMT. What we learn from NGS studies may create “personalized medicine for infectious diseases” by informing clinical management options and prognostication.

Next-generation sequencing utility has been demonstrated for microbial strain typing in epidemiological outbreak investigations, at least for those involving a limited number of strains or a single strain. A highly publicized example is the 2011 European shiga-toxin *Escherichia coli* outbreak during which NGS provided real-time de novo characterization of a novel outbreak strain.

Improvements are needed in order to make NGS an effective or adjunct tool for routine use in clinical microbiology, including commercialized, cost-competitive, user-friendly library preparation and instrumentation and software, standardized protocols and proficiency testing, well-curated reference genomes, and regulatory mandates revised to align with changing technology and practice. As previously mentioned, appropriate improvements in infrastructure may be required to accommodate the complex data to produce a succinct clinical report.

### Next-Generation Sequencing Applications in Oncology

Next-generation sequencing tests for diagnosing and managing oncology patients have been used since the technology was utilized to diagnose patients with solid tumors or hematologic abnormalities. The advantage of NGS lies in its ability to conduct large-scale inquiries for many sequence variants that are comprehensive, inclusive, and sensitive. Consequently, the technology can actually save costs compared to multiple, individual nucleic acid-based tests (such as fluorescent in situ hybridization, PCR/sequencing, etc). The small amount of tissue required may also obviate the need for an additional procedure, such as a repeat biopsy, to obtain sufficient material for analysis. Older methods required more nucleic acid which could not be extracted from biopsies, but the smaller amounts of tissue necessary for NGS may allow successful sequencing of the original biopsy. Next-generation sequencing offers clear advantages compared to the traditional one-gene one-test approach. Germline or somatic variants can be detected by NGS depending on the goal of testing. Libraries for somatic changes may be created from off-the-shelf panels or customized for individual types of malignancies. Previously, WGS and WES were not considered practical for routine clinical use; however, many academic laboratories and commercial vendors are developing test panels, using several genes or hundreds of genes, to detect a variety of genetic/somatic variants in cancers. Testing somatic variants in tumor specimens requires sequencing at a higher depth (ie, 1000× average coverage) that is offered by targeted panels, in contrast to germline testing in which a lower sequencing depth (ie, 30× average coverage) may be undertaken to reliably detect variants. To assess the clinical relevance of sequencing results, several determinants are considered, including single-nucleotide polymorphisms, point mutations or single-nucleotide variants, nucleotide insertions or deletions, gene fusion/rearrangements, and copy number variations (see Table 1 for definitions). Next-generation sequencing tests can be DNA or RNA based, or both, depending on the purpose and design of the test. As the technology matures, test panel costs are becoming affordable for routine clinical use and are being rapidly deployed in laboratories. This is especially true in the field of oncology for diagnostic and prognostic purposes, as well as the selection of appropriate therapies. The NGS tool has become an important part of a personalized medicine approach to target therapy.

Next-generation sequencing has been used as a molecular diagnostic test for many solid tissue cancers as well as hematologic malignancies. Test results can be helpful for the initial diagnosis, tumor classification, determining the origin of the cancer, and prognosis. Table provides a partial list of cancers where NGS information has provided value for managing patients. Thyroid nodules are a specific example where fine needle aspiration and cytologic examination may not yield a definitive diagnosis, while NGS has been shown to have high specificity and sensitivity for cancer detection. However, not
Table 2. Examples Where NGS Provides Additional Information.*

| Cancer | Clinical Application and/or Research Area |
|--------|------------------------------------------|
| Thyroid cancer | Clinical application: Adjunct test to fine needle aspiration for the diagnosis of indeterminate thyroid nodules |
| Breast, ovarian, and endometrial cancer | Clinical application: Diagnostic testing and identifying familial mutations |
| Melanoma | Clinical application: Therapeutic decisions, prognostic value, and identifying inherited cancer risk |
| Colon cancer | Clinical application: Differential diagnosis, therapeutic decision, and inherited risk assessment for cancer |
| Lung cancer | Clinical application: Therapeutic decisions for non-small cell lung cancer (NSCLC) |
| Leukemia | Clinical application: Classification and prognosis at the time of diagnosis, therapeutic decisions, and determination of a different gene mutation profile at the time of relapse |
| Lymphoma | Research area: Potential prognostic implication and association with response to treatment |
| Head and neck cancer | Clinical application: Diagnostic, prognostic, and therapeutic decisions |
| Sarcoma | Clinical application: Prognostic and therapeutic decisions |
| Brain tumors | Clinical application: Evaluation for hereditary paraganglioma/pheochromocytoma |
| Paraganglioma and pheochromocytoma | Clinical application: Inherited risks for cancer and therapeutic decisions |
| Prostate cancer | Clinical application: Inherited syndromes and prognostic and therapeutic decisions in KIT/ PDGFRA wild-type GIST |
| Gastrointestinal stromal tumors | |

Abbreviations: GIST, gastrointestinal stromal tumor; NGS, next-generation sequencing.

*Partial list of tumors/cancers where NGS has been shown to provide additional information.

all patients will derive enough clinical benefit to justify the cost of using NGS testing and careful test utilization is prudent.

A clear, clinically important use of NGS is to identify the most appropriate therapy for the individual patient. The National Cancer Institute’s Molecular Analysis for Therapy Choice trial is a good example of how NGS technology can be utilized in clinical practice. Despite the promising clinical utility of NGS, the influence of molecular profiling on individual patient’s targeted therapy has yet to reach its full potential. For example, the Integrated Molecular Profiling in Advanced Cancers Trial and Community Oncology Molecular Profiling in Advanced Cancers Trial (IMPACT/COMPACT) trial showed that only 5% of patients received targeted treatments based on their profiling results. While this is a relatively low number, the study did not comprehensively evaluate factors that may have influenced the targeted therapy. The trial was limited to specimens obtained many years prior to the molecular testing and did not profile the metastatic lesions, which may have yielded different molecular profile. Also some of the patients included in the study were heavily pretreated and were not well enough to receive further treatment based on the results of molecular testing. Molecular testing also did not include copy number variation or recurrent translocations, which may have influenced the therapeutic decision.

With the availability of clinical trials matching drugs targeting specific genetic alterations, many academic medical centers and even larger community hospitals have begun to adopt NGS into their routine practice. Companion tests for targeted therapy are also in development. Among the current obstacles, preventing even wider adoption are the initial cost to purchase the instrument and complex bioinformatics to interpret the sequence data. Recently, commercial laboratories have entered into this market and competition will ultimately lower the cost and improve the quality of products. Future development will allow the NGS technology to be more affordable with wider applications such as cell-free DNA for circulating tumor DNA detection or liquid biopsy. These can potentially be used for monitoring disease progression, finding secondary mutations (such as mutations in epidermal growth factor receptor), minimal residual disease management, and occult tumor detection.

Conclusion

Next-generation sequencing has become a widely used technology in the field of pathology. Several advances have reduced the time and cost of the test, while data analysis has extended the utility. Routine histopathology and diagnostic work will not be replaced in the near future, but NGS offers significant advantages in selected cases.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported in part by NIH grant R21 AI112887-01, R01GM117519, and T32 GM86308 (DGR).

ORCID iD

Thomas Maher http://orcid.org/0000-0002-6083-9487
Daniel Remick http://orcid.org/0000-0002-2615-3713

References

1. Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature. 1953;171:737-738.
2. Maxam AM, Gilbert W. A new method for sequencing DNA. *Proc Natl Acad Sci U S A.* 1977;74:560-564.

3. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. 1977. *Biotechnology.* 1992;24:104-108.

4. Hunkapiller T, Kaiser RJ, Koop BF, Hood L. Large-scale and automated DNA sequence determination. *Science.* 1991;254:59-67.

5. Smith LM, Sanders IZ, Kaiser RJ, et al. Fluorescence detection in automated DNA sequence analysis. *Nature.* 1986;321:674-679.

6. Prober JM, Trainor GL, Dam RJ, et al. A system for rapid DNA sequencing with fluorescent chain-terminating deoxyribonucleotides. *Science.* 1987;238:336-341.

7. Luckey JA, Drossman H, Kostichka AJ, et al. High speed DNA sequencing by capillary electrophoresis. *Nucleic Acids Res.* 1990;18:4417-4421.

8. Swerdlow H, Gesteland R. Capillary gel electrophoresis for rapid, high resolution DNA sequencing. *Nucleic Acids Res.* 1990;18:1415-1419.

9. Chen CY. DNA polymerases drive DNA sequencing-by-synthesis technologies: both past and present. *Front Microbiol.* 2014;5:305.

10. Jackson DA, Symons RH, Berg P. Biochemical method for inserting new genetic information into DNA of simian virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of *Escherichia coli.* *Proc Natl Acad Sci U S A.* 1972;69:2904-2909.

11. Saiki RK, Gelfand DH, Stoffel S, et al. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science.* 1988;239:487-491.

12. Adams MD, Celniker SE, Holt RA, et al. The genome sequence of *Drosophila melanogaster.* *Science.* 2000;287:2185-2195.

13. C. elegans Sequencing Consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science.* 1998;282:2012-2018.

14. Fleischmann RD, Adams MD, White O, et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science.* 1995;269:496-512.

15. Baer R, Bankier AT, Biggin MD, et al. DNA sequence and expression of the B95-8 Epstein-Barr virus genome. *Nature.* 1984;310:207-211.

16. Gibbs RA, Weinstock GM, Metzker ML, et al. Genome Sequencing Project Consortium. Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature.* 2004;428:493-521.

17. Mouse Genome Sequencing Consortium, Waterston RH, Lindblad-Toh K, Birney E, et al. Initial sequencing and comparative analysis of the mouse genome. *Nature.* 2002;420:520-562.

18. Lander ES, Linton LM, Birren B, et al; International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature.* 2001;409:860-921.

19. Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. *Science.* 2001;291:1304-1351.

20. Levy S, Sutton G, Ng PC, et al. The diploid genome sequence of an individual human. *PLOS Biol.* 2007;5:e254.

21. Heather JM, Chain B. The sequence of sequencers: the history of sequencing DNA. *Genomics.* 2016;107:1-8.

22. Reuter JA, Spacek DV, Snyder MP. High-throughput sequencing technologies. *Mol Cell.* 2015;58:586-597.

23. Morey M, Fernandez-Marmiesse A, Castineiras D, Fraga JM, Couce ML, Cocho JA. A glimpse into past, present, and future DNA sequencing. *Mol Genet Metab.* 2013;110:3-24.

24. Margulies M, Egholm M, Altman WE, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature.* 2005;437:376-380.

25. Wheeler DA, Srinivasan M, Egholm M, et al. The complete genome of an individual by massively parallel DNA sequencing. *Nature.* 2008;452:872-876.

26. Bentley DR, Balasubramanian S, Swerdlow HP, et al. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature.* 2008;456:53-59.

27. Kim HS, Lee H, Shin SJ, et al. Complementary utility of targeted next-generation sequencing and immunohistochemistry panels as a screening platform to select targeted therapy for advanced gastric cancer. *Oncotarget.* 2017;8:38389-38398.

28. D’Haene N, Le Mercier M, De Neve N, et al. Clinical validation of targeted next generation sequencing for colon and lung cancers. *PloS One.* 2015;10:e0138245.

29. van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C. Ten years of next-generation sequencing technology. *Trends Genet.* 2014;30:418-426.

30. Watson M. Illuminating the future of DNA sequencing. *Genome Biol.* 2014;15:108.

31. Rothberg JM, Hinz W, Rearick TM, et al. An integrated semiconductor device enabling non-optical genome sequencing. *Nature.* 2011;475:348-352.

32. McCarthy JJ, McLeod HL, Ginsburg GS. Genomic medicine: a decade of successes, challenges, and opportunities. *Sci Transl Med.* 2013;5:189sr4.

33. Green ED, Guyer MS; National Human Genome Research Institute. Charting a course for genomic medicine from base pairs to bedside. *Nature.* 2011;470:204-213.

34. Syn NL, Yong WP, Goh BC, Lee SC. Evolving landscape of tumor molecular profiling for personalized cancer therapy: a comprehensive review. *Expert Opin Drug Metab Toxicol.* 2016;12:911-922.

35. Parsons HA, Beaver JA, Park BH. Circulating plasma tumor DNA. *Adv Exp Med Biol.* 2016;829:259-276.

36. Lalani SR. Current genetic testing tools in neonatal medicine. *J Pediatr Neonatol.* 2017;6:1-12.

37. Xiong X, Milligan J, Shen Y, et al. High-throughput SNP genotyping assay on the Illumina platform as a screening platform to select targeted therapy for advanced gastric cancer. *Cytogenet Genome Res.* 2011;133:187-190.

38. Watanabe T, Morikawa S, Tada M, et al. Evaluation of the performance of sequencing technologies for clinical use. *Jpn J Genet.* 2014;89:13-23.

39. Wang J, Song Y. Single cell sequencing: a distinct new field. *Adv Exp Med Biol.* 2016;110:3-24.

40. Zhu W, Zhang XY, Marjani SL, et al. Next-generation molecular diagnostic: single-cell sequencing from bench to bedside. *Cell Mol Life Sci.* 2017;74:869-880.

41. Koshimizu E, Miyatake S, Okamoto N, et al. Performance comparison of bench-top next generation sequencers using
microdroplet PCR-based enrichment for targeted sequencing in patients with autism spectrum disorder. *PLoS One.* 2013;8:e74167.

42. Jennings LJ, Arcila ME, Corless C, et al. Guidelines for validation of next-generation sequencing-based oncology panels: a joint consensus recommendation of the Association for Molecular Pathology and College of American Pathologists. *J Mol Diagn.* 2017;19:341-365.

43. Misyura M, Zhang T, Sukhai MA, et al. Comparison of next-generation sequencing panels and platforms for detection and verification of somatic tumor variants for clinical diagnostics. *J Mol Diagn.* 2016;18:842-850.

44. Shim S, Kim Y, Oh SC, et al. Validation and optimization of the Ion Torrent S5 XL sequencer and Oncomine workflow for BRCA1 and BRCA2 genetic testing. *Onco Targets.* 2017;8:34858-34866.

45. Strom CM, Rivera S, Elzinga C, et al. Development and validation of a next-generation sequencing assay for BRCA1 and BRCA2 variants for the clinical laboratory. *PLOS One.* 2015;10:e0136419.

46. Yohe S, Thyagarajan B. Review of clinical next-generation sequencing. *Arch Pathol Lab Med.* 2017;141:1544-1557.

47. Mardis ER. The challenges of big data. *Dis Model Mech.* 2016;9:483-485.

48. Gogol-Doring A, Chen W. An overview of the analysis of next generation sequencing data. *Methods Mol Biol.* 2012;802:249-257.

49. Li MM, Datto M, Duncavage EJ, et al. Standards and guidelines for the interpretation and reporting of variant sequences in cancer: a joint consensus recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. *J Mol Diagn.* 2017;19:4-23.

50. Matthijs G, Souche E, Alders M, et al. Guidelines for diagnostic next-generation sequencing. *Eur J Hum Genet.* 2016;24:1515.

51. Jamuar SS, Tan EC. Clinical application of next-generation sequencing for Mendelian diseases. *Hum Genomics.* 2015;9:10.

52. Daoud H, Luco SM, Li R, et al. Next-generation sequencing for diagnosis of rare diseases in the neonatal intensive care unit. *CMAJ.* 2016;188:E254-E260.

53. Sawyer SL, Hartley T, Dyment DA, et al. Utility of whole-exome sequencing for those near the end of the diagnostic odyssey: time to address gaps in care. *Clin Genet.* 2016;89:275-284.

54. Aparisi MJ, Aller E, Fuster-Garcia C, et al. Targeted next-generation sequencing for molecular diagnosis of Usher syndrome. *Orphanet J Rare Dis.* 2014;9:168.

55. Boycott K, Hartley T, Adam S, et al; Canadian College of Medical Geneticists. The clinical application of genome-wide sequencing for monogenic diseases in Canada: Position Statement of the Canadian College of Medical Geneticists. *J Med Genet.* 2015;52:431-437.

56. Lee H, Deignan JL, Dorrani N, et al. Clinical exome sequencing for genetic identification of rare Mendelian disorders. *JAMA.* 2014;312:1880-1887.

57. Stavropoulos DJ, Merico D, Jobling R, et al. Whole-genome sequencing expands diagnostic utility and improves clinical management in paediatric medicine. *NPJ Genomic Medicine.* 2016;1:15012.

58. Wright CF, Fitzgerald TW, Jones WD, et al. Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. *Lancet.* 2015;385:1305-1314.

59. Kalia SS, Adelman K, Bale SJ, et al. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. *Genet Med.* 2017;19:249-255.

60. Ackerman SL, Koenig BA. Understanding variations in secondary findings reporting practices across U.S. genome sequencing laboratories. *AJOB Empir Bioeth.* 2018;9:48-57.

61. Allyse M, Michie M. Not-so-incidental findings: the ACMG recommendations on the reporting of incidental findings in clinical whole genome and whole exome sequencing. *Trends Biotechnol.* 2013;31:439-441.

62. Wilson MR, Naccache SN, Samaya E, et al. Actionable diagnosis of neuroleptospirosis by next-generation sequencing. *N Engl J Med.* 2014;370:2408-2417.

63. Loman NJ, Constantinidou C, Chan JZ, et al. High-throughput bacterial genome sequencing: an embarrassment of choice, a world of opportunity. *Nat Rev Microbiol.* 2012;10:599-606.

64. Lefterova MI, Suarez CJ, Banaei N, Pinsky BA. Next-generation sequencing for infectious disease diagnosis and management: a report of the association for molecular pathology. *J Mol Diagn.* 2015;17:623-634.

65. Deurenberg RH, Bathoorn E, Chlebowicz MA, et al. Application of next generation sequencing in clinical microbiology and infection prevention. *J Biotechnol.* 2017;243:16-24.

66. Fisher KE, Zhang L, Wang J, et al. Clinical validation and implementation of a targeted next-generation sequencing assay to detect somatic variants in non-small cell lung, melanoma, and gastrointestinal malignancies. *J Mol Diagn.* 2016;18:299-315.

67. Hovelson DH, McDaniel AS, Cani AK, et al. Development and validation of a scalable next-generation sequencing system for assessing relevant somatic variants in solid tumors. *Neoplasia.* 2015;17:385-399.

68. Wagle N, Berger MF, Davis MJ, et al. High-throughput detection of actionable genomic alterations in clinical tumor samples by targeted, massively parallel sequencing. *Cancer Discov.* 2012;2:82-93.

69. Au CH, Wa A, Ho DN, Chan TL, Ma ES. Clinical evaluation of panel testing by next-generation sequencing (NGS) for gene mutations in myeloid neoplasms. *Diagn Pathol.* 2016;11:11.

70. Luthra R, Chen H, Roy-Chowdhuri S, Singh RR. Next-generation sequencing in clinical molecular diagnostics of cancer: advantages and challenges. *Cancers (Basel).* 2015;7:2023-2036.

71. Pant S, Weiner R, Marton MJ. Navigating the rapids: the development of regulated next-generation sequencing-based clinical trial assays and companion diagnostics. *Front Oncol.* 2014;4:78.

72. Quinn AM, Hickson N, Adaway M, et al. Diagnostic mutation profiling and validation of non-small-cell lung cancer small biopsy samples using a high throughput platform. *J Thorac Oncol.* 2015;10:784-792.
73. Luthra R, Patel KP, Routhbort MJ, et al. A targeted high-throughput next-generation sequencing panel for clinical screening of mutations, gene amplifications, and fusions in solid tumors. *J Mol Diagn.* 2017;19:255-264.

74. Coyne GO, Takebe N, Chen AP. Defining precision: the precision medicine initiative trials NCI-MPACT and NCI-MATCH. *Curr Probl Cancer.* 2017;41:182-193.

75. McNeil C. NCI-MATCH launch highlights new trial design in precision-medicine era [published online ahead of print July 03, 2015]. *J Natl Cancer Inst.* 2015;107.

76. Verri C, Borzi C, Holscher T, et al. Mutational profile from targeted NGS predicts survival in LDCT screening-detected lung cancers. *J Thorac Oncol.* 2017;12:922-931.

77. Nikiforov YE, Carty SE, Chiosa SL, et al. Impact of the multi-gene ThyroSeq Next-generation sequencing assay on cancer diagnosis in thyroid nodules with atypia of undetermined significance/follicular lesion of undetermined significance cytology. *Thyroid.* 2015;25:1217-1223.

78. Nikiforova MN, Wald AI, Roy S, Durso MB, Nikiforov YE. Targeted next-generation sequencing panel (ThyroSeq) for detection of mutations in thyroid cancer. *J Clin Endocrinol Metab.* 2013;98:E1852-E1860.

79. Picarsic JL, Buryk MA, Ozolek J, et al. Molecular characterization of sporadic pediatric thyroid carcinoma with the DNA/RNA ThyroSeq v2 next-generation sequencing assay. *Pediatr Dev Pathol.* 2016;19:115-122.

80. Castera L, Krieger S, Rousselin A, et al. Next-generation sequencing for the diagnosis of hereditary breast and ovarian cancer using genomic capture targeting multiple candidate genes. *Eur J Hum Genet.* 2014;22:1305-1313.

81. Easton DF, Pharaoh PD, Antoniou AC, et al. Gene-panel sequencing and the prediction of breast-cancer risk. *N Engl J Med.* 2015;372:2243-2257.

82. Johnson DB, Dahlman KH, Knol J, et al. Enabling a genetically informed approach to cancer medicine: a retrospective evaluation of the impact of comprehensive tumor profiling using a targeted next-generation sequencing panel. *Oncologist.* 2014;19:616-622.

83. Judkins T, Leclair B, Bowles K, et al. Development and analytical validation of a 25-gene next generation sequencing panel that includes the BRCA1 and BRCA2 genes to assess hereditary cancer risk. *BMC Cancer.* 2015;15:215.

84. LaDuca H, Stuenkel AJ, Dolinsky JS, et al. Utilization of multi-gene panels in hereditary cancer predisposition testing: analysis of more than 2,000 patients. *Genet Med.* 2014;16:830-837.

85. Vasan N, Yelensky R, Wang K, et al. A targeted next-generation sequencing assay detects a high frequency of therapeutically targetable alterations in primary and metastatic breast cancers: implications for clinical practice. *Oncologist.* 2014;19:453-458.

86. De Luca F, Rotunno G, Salvianti F, et al. Mutational analysis of single circulating tumor cells by next generation sequencing in metastatic breast cancer. *Oncotarget.* 2016;7:26107-26119.

87. Badal B, Solovyov A, Di Cecilia S, et al. Transcriptional dissection of melanoma identifies a high-risk subtype underlying TP53 family genes and epigenome deregulation. *JCI Insight.* 2017;2:92102.

88. Carlson JA, Caldeira Xavier JC Jr, Tarasen A, et al. Next-generation sequencing reveals pathway activations and new routes to targeted therapies in cutaneous metastatic melanoma. *Am J Dermatopathol.* 2017;39:1-13.

89. de Abreu FB, Peterson JD, Amos CI, Wells WA, Tsongalis GJ. Effective quality management practices in routine clinical next-generation sequencing. *Clin Chem Lab Med.* 2016;54:761-771.

90. de Unamuno Bustos B, Murria Estal R, Perez Simo G, et al. Towards personalized medicine in melanoma: implementation of a clinical next-generation sequencing panel. *Sci Rep.* 2017;7:495.

91. Pinto R, De Summa S, Strippoli S, et al. The next generation of metastatic melanoma: uncovering the genetic variants for anti-BRAF therapy response. *Oncotarget.* 2016;7:25135-25149.

92. Belardini L, Capalbo C, Buffone A, et al. Validation of the Ion Torrent PGM sequencing for the prospective routine molecular diagnostic of colorectal cancer. *Clin Biochem.* 2015;48:908-910.

93. Darwanto A, Hein AM, Strauss S, et al. Use of the QIAGEN GeneReader NGS system for detection of KRAS mutations, validated by the QIAGEN Therascreen PCR kit and alternative NGS platform. *BMC Cancer.* 2017;17:358.

94. Stadler ZK, Battaglini F, Middha S, et al. Reliable detection of mismatch repair deficiency in colorectal cancers using mutational load in next-generation sequencing panels. *J Clin Oncol.* 2016;34:2141-2147.

95. Tops BB, Normanno N, Kurth H, et al. Development of a semiconductor sequencing-based panel for genotyping of colon and lung cancer by the Onconetwork consortium. *BMC Cancer.* 2015;15:26.

96. Xie N, Yao Y, Wan L, Zhu T, Liu L, Yuan J. Next-generation sequencing reveals lymph node metastasis associated genetic markers in colorectal cancer. *Exp Ther Med.* 2017;14:338-343.

97. Coco S, Truini A, Vanni I, et al. Next generation sequencing in non-small cell lung cancer: new avenues toward the personalized medicine. *Curr Drug Targets.* 2015;16:47-59.

98. DiBardino DM, Saqi A, Elvin JA, et al. Yield and clinical utility of next-generation sequencing in selected patients with lung adenocarcinoma. *Clin Lung Cancer.* 2016;17:517-522 e3.

99. Hagemann IS, Devarakonda S, Lockwood CM, et al. Clinical next-generation sequencing in patients with non-small cell lung cancer. *Cancer.* 2015;121:631-639.

100. Lim SM, Kim EY, Kim HR, et al. Genomic profiling of lung adenocarcinoma patients reveals therapeutic targets and confers clinical benefit when standard molecular testing is negative. *Oncotarget.* 2016;7:24172-24178.

101. Lindquist KE, Karlsson A, Leveen P, et al. Clinical framework for next generation sequencing based analysis of treatment predictive mutations and multiplexed gene fusion detection in non-small cell lung cancer. *Oncotarget.* 2017;8:34796-34810.

102. Richer AL, Friel JM, Carson VM, Inge LJ, Whitsett TG. Effective quality management practices in routine clinical next-generation sequencing. *Clin Chem Lab Med.* 2016;54:761-771.

103. Alikian M, Gale RP, Apperley JF, Foroni L. Molecular techniques for the personalised management of patients with chronic myeloid leukaemia. *Biomol Detect Quantif.* 2017;11:4-20.
