

# Making the Leap from Research Laboratory to Clinic: Challenges and Opportunities for Next-Generation Sequencing in Infectious Disease Diagnostics

Brittany Goldberg,<sup>a</sup> Heike Sichtig,<sup>b</sup> Chelsie Geyer,<sup>c</sup> Nathan Ledebor,<sup>d</sup> George M. Weinstock<sup>e</sup>

Division of Pediatric Infectious Diseases, Children's National Medical Center, Washington, DC, USA<sup>a</sup>; Division of Microbiology Devices, Food and Drug Administration, Office of In Vitro Diagnostics and Radiological Health, Silver Spring, Maryland, USA<sup>b</sup>; American Society for Microbiology, Washington, DC, USA<sup>c</sup>; Department of Pathology, Medical College of Wisconsin, Milwaukee, Wisconsin, USA<sup>d</sup>; The Jackson Laboratory for Genomic Medicine, Farmington, Connecticut, USA<sup>e</sup>

**ABSTRACT** Next-generation DNA sequencing (NGS) has progressed enormously over the past decade, transforming genomic analysis and opening up many new opportunities for applications in clinical microbiology laboratories. The impact of NGS on microbiology has been revolutionary, with new microbial genomic sequences being generated daily, leading to the development of large databases of genomes and gene sequences. The ability to analyze microbial communities without culturing organisms has created the ever-growing field of metagenomics and microbiome analysis and has generated significant new insights into the relation between host and microbe. The medical literature contains many examples of how this new technology can be used for infectious disease diagnostics and pathogen analysis. The implementation of NGS in medical practice has been a slow process due to various challenges such as clinical trials, lack of applicable regulatory guidelines, and the adaptation of the technology to the clinical environment. In April 2015, the American Academy of Microbiology (AAM) convened a colloquium to begin to define these issues, and in this document, we present some of the concepts that were generated from these discussions.

Use of next-generation DNA sequencing (NGS) (Table 1) in infectious disease diagnostics has progressed slowly over the past 10 years despite continued advances in sequencing technology. The first commercial NGS platform, the GS20 sequencer from 454 Life Sciences, which was originally released in 2005 (1, 2), resulted in a more than 100-fold increase in the amount of microbial genomic sequence data produced in a day compared to preceding instruments. Despite the growing body of literature and research broadly applying sequencing-based technology to disease pathophysiology, epidemiology, and clinical diagnostics, the clinical microbiology laboratory has yet to widely adopt NGS technology. As microbiology laboratories are faced with a wealth of innovative and often costly molecular technologies, the role of NGS in clinical infectious disease diagnostics needs to be carefully evaluated.

A number of highly publicized case reports and clinical studies have showcased the application of NGS as a single diagnostic tool with the potential to be broadly applicable to infectious disease diagnostics. Metagenomic (Table 1) sequencing has demonstrated its ability to identify microbial pathogens where traditional diagnostics have otherwise failed. For example, it is estimated that 63% of encephalitis cases go undiagnosed despite extensive testing (3). Several cases in the literature have successfully employed NGS to diagnose rare, novel, or atypical infectious etiologies for encephalitis, including cases of infection by *Leptospira* (4), astrovirus (5), and bornavirus (6). In one case, 38 different diagnostic tests had been conducted and failed to yield an actionable answer before a single NGS assay was performed, which identified the pathogen (4). Similarly, the utilization of metagenomic NGS identified divergent astrovirus clades in a pair of patients with encephalitis and demonstrated the unusual zoonotic potential of a group of these viruses (7).

Another promising application of NGS technology is hospital infection control surveillance programs and community outbreak investigations (8). By conducting whole-genome sequencing

(WGS) (Table 1), organisms can be identified at the subspecies/strain level based on the single nucleotide polymorphisms (SNPs) (Table 1) and other variants (Table 1) in their genotype. WGS through NGS technology offers greater precision than do more-traditional typing tools such as multilocus sequence typing and pulsed-field gel electrophoresis, which may assist in refining outbreak investigations and better guide infection control interventions (9). Because WGS analysis requires significant amounts of sequencing data, traditional sequencing methods preclude the use of WGS analysis for outbreak investigations. However, NGS platforms can generate the large volume of data needed for SNP or variant analysis and have led to a rapid expansion in the use of WGS for public health investigations. For example, WGS using NGS technology was applied to investigate an outbreak of hemolytic-uremic syndrome caused by an unusual strain of *Escherichia coli* in Germany (10), the origins of the 2010 Haitian *Vibrio cholerae* epidemic (11), a series of methicillin-resistant *Staphylococcus aureus* infections in a neonatal intensive care unit (12), and the origins of a series of nosocomial carbapenem-resistant *Klebsiella pneumoniae* infections (13), among many others.

While identification of causative microorganisms of disease is the chief responsibility of clinical microbiology laboratories, conducting antimicrobial resistance testing to guide therapy is among the most important tests conducted in the laboratory. NGS has the potential to suggest antimicrobial resistance through identifica-

Published 8 December 2015

**Citation** Goldberg B, Sichtig H, Geyer C, Ledebor N, Weinstock GM. 2015. Making the leap from research laboratory to clinic: challenges and opportunities for next-generation sequencing in infectious disease diagnostics. *mBio* 6(6):e01888-15. doi:10.1128/mBio.01888-15.

**Copyright** © 2015 Goldberg et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution-Noncommercial-ShareAlike 3.0 Unported license](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

Address correspondence to George M. Weinstock, [George.Weinstock@jax.org](mailto:George.Weinstock@jax.org).

TABLE 1 Glossary of terms used in DNA sequence analysis

Term	Abbreviation	Definition
16S rRNA gene		A slowly evolving gene in bacteria whose sequence is used for definition of taxa. It is a gene that is targeted for sequencing in microbiome analysis, where the goal is enumeration of the taxa present in a community.
Alignment		The process of comparing the sequence of a single sequencing read or a contig/whole genome following assembly to a reference genome. The goal is often to identify the organism from which a sequencing read came or to identify variants within the sequence.
Assembly		Reconstructing a genome, in whole or in part, from the fragment sequences produced by WGS (or mWGS).
Contig		A contiguous stretch of sequence produced when a series of overlapping sequence reads are merged to produce a single longer sequence.
Dideoxynucleotide sequencing		A “classical” method of DNA sequencing that preceded NGS and is frequently called Sanger sequencing.
Metagenomics		Analyzing a mixture of microbial genomes, a metagenome, without separating the genomes or culturing the organisms.
Metagenomic whole-genome shotgun sequencing	mWGS	The application of WGS to a metagenomics sample. DNA is extracted from the sample, producing a mixture of genomes, which are then subjected to WGS en masse.
Microbiome		A community of microbes comprising bacteria, viruses, and fungi and other eukaryotic microbes. Often the target of metagenomic analyses.
Next-generation sequencing	NGS	A collection of DNA sequencing methods that each use different biochemical approaches and instruments to produce data in vastly larger amounts, at greatly lower cost, in shorter time, and with less manual intervention than previous methods.
Reference genome		A genome sequence of a particular organism that can be used as a standard, e.g., for alignment or comparison of other genomes.
Read		The basic element produced by DNA sequencing. Sequencing of a DNA fragment produces a series of bases called a sequencing read.
Sanger sequencing		A “classical” method of DNA sequencing that preceded NGS but was almost exclusively used from the 1970s until the advent of NGS. Compared to NGS, it produced fewer data, was more expensive, and required more manual work.
Single nucleotide polymorphism	SNP	A difference of a single base compared to a reference genome. These can be substitutions of one base for another or insertion/deletion of a base (indel).
Variant		Any difference in a DNA sequence compared to a reference sequence. This can be a single-base difference (SNP) or insertions, deletions, inversions, or translocations of larger stretches of sequence (structural variants).
Whole-genome shotgun sequencing	WGS	Randomly fragmenting an entire genome and obtaining DNA sequence from the fragments to produce a collection of random DNA sequences. This can be applied to a single bacterium or to a mixture (metagenomc; see mWGS). These data can be used to identify variants following alignment of genes by comparison to sequence databases or to compare genome structures following assembly.

tion of known resistance genes (14). Although WGS is in the early stages of development, studies have suggested that WGS can be used to predict antibiotic resistance with performance characteristics approaching those of traditional phenotypic testing. Stoesser et al. (15) demonstrated sensitivity and specificity of 96% and 97%, respectively, when using WGS to predict antibiotic resistance for clinical isolates of *E. coli* and *K. pneumoniae*. Comparative genomic sequencing has also been used to identify daptomycin resistance due to point mutations in metabolic genes that occurred during therapy for two cases of vancomycin-resistant enterococcal bacteremia which were ultimately fatal (16, 17). Earlier application of WGS might have detected these point mutations and guided therapy. However, it has not been established if WGS can be broadly applied to the full spectrum of pathogenic bacteria, particularly those with a diverse armamentarium of resistance mechanisms. WGS analysis of antimicrobial resistance genes could be particularly beneficial for slow-growing or difficult-to-culture organisms and organisms that elude phenotypic testing altogether. The use of WGS is not limited to the detection of bacterial resistance genes. WGS has also been applied for the following purposes: detection of low-level drug resistance among human immunodeficiency virus (HIV) “escape” variant populations (e.g., protease inhibitor [PI] and reverse transcriptase minor sequence variants) and coreceptor tropism (CCR5 and CXCR4) and analyses that are not possible using current genotypic and phenotypic HIV assays (18). These successes and potential applications of WGS analysis have been made possible by the advance of NGS technology, which provides the tools to produce useful WGS data.

Metagenomic shotgun sequencing (mWGS) (Table 1) is often used to study microbial communities in human disease in order to identify correlative or causative relations. Such communities comprise hundreds of different taxa of bacteria, viruses, and fungi and other eukaryotic microbes. Many of these organisms are difficult to culture, and culture-independent methods of performing comprehensive sampling of these complex communities have been the major obstacle to analysis. NGS is currently the best available analytical approach to profile microbiomes (Table 1) for this purpose. To date, NGS has helped elucidate the role of the lung and gut microbiomes in both general health and various diseases, including obesity, inflammatory bowel disease, cystic fibrosis (CF), metabolic syndrome, type II diabetes, and cardiovascular disease (19–23). As NGS continues to expand our knowledge base on metagenomics, microbiome analysis may produce diagnostic or prognostic biomarkers to guide therapeutic decisions.

There is a long-standing precedent in clinical microbiology laboratories to adopt new technology that complements or supplants existing “gold standard” testing. The viral culture bench is all but extinct in most clinical laboratories, with PCR-based molecular assays currently dominating viral diagnostics. Nonviral microbial diagnostics have been relatively slow to adopt molecular technology, but this is rapidly changing with a new generation of molecular diagnostics that utilize specific PCR primers for different bacterial, parasitic, and fungal targets. Several of these multiplex assays have already received U.S. Food and Drug Administration (FDA) clearance and offer laboratories an attractive and easy way to detect clinically relevant microbial pathogens. Similarly to the impact of PCR technology, microbial NGS diagnostics offer another step forward in the quality and quantity of information that could potentially be provided to clinicians and patients. In

April 2015, the American Academy of Microbiology (AAM) conducted a colloquium to critically evaluate the trends in the use of NGS for infectious disease diagnostics. Below, we describe some of the concepts that evolved from the AAM’s NGS colloquium (report to be published).

## NEXT-GENERATION SEQUENCING

DNA sequencing technology was previously dominated by procedures using the dideoxynucleotide chain termination method (also referred to as Sanger sequencing) (Table 1) for close to 30 years. In the mid-2000s, a number of new methods began to appear, commonly referred to as the “next-generation sequencing methods.” The first method that was widely commercialized was the “pyrosequencing” method—a method that formed the basis for the 454 Life Sciences instruments for about 10 years until the company closed in 2015. These instruments had a major impact on microbial genomics and became the leading platform for producing whole-genome sequences from individual bacteria as well as the platform of choice for sequencing 16S rRNA (Table 1) genes in microbiome analyses.

Since the advent of NGS, there has been an ongoing expansion of sequencing methods and instruments as well as continual improvements in the quality, quantity, and cost of sequences that are produced. The spectrum of common and available instruments current as of July 2015 is shown in Table 2, which indicates the considerable range in characteristics of NGS options.

An understanding of how these characteristics match the various NGS applications is useful to appreciate the strengths or limitations of different sequencing approaches. To sequence a single microbial genome, the genome is fragmented into manageable-size pieces, which are then used in WGS (Fig. 1). Depending on the sequencing platform, this can produce sequences (i.e., reads) (Table 1) that range from hundreds of bases to thousands or tens of thousands of bases in length. When the goal is to produce a microbial genome, the reads are assembled into a single genome sequence using a variety of computational strategies. Often, the assembly (Table 1) leaves gaps, and the individual chunks of genomic sequence that result are called contigs (Table 1). When identifying variants in the genome (e.g., for typing or tracking), the reads are directly aligned with a reference genome (Table 1) where the sequence variants are determined. The smaller the length of sequence read, the more contigs and gaps that are produced during the assembly process. Hence, there is a preference for longer sequence reads. To obtain a complete sequence without gaps, one often uses both short and long reads to combine the strengths of these various sequencing approaches (24). However, the platforms producing shorter reads are often less expensive and can yield large amounts of data. Moreover, the accuracy of base calls comprising the sequence read is also important. The lower the accuracy of the read, the more reads (i.e., coverage) that are required to obtain good contigs, ultimately driving up the sequencing costs. In addition, when reads have errors, it is harder to identify the true sequence variants. At present, platforms that produce longer reads tend to have lower accuracy and users have to consider the tradeoffs between the various platforms when designing a project.

Another major application for NGS is metagenomic sequencing. In this case, one starts with a particular specimen type such as stool, saliva, a nasal swab, etc., with the goal of determining the bacteria, viruses, fungi, and other microbes that are present. Fol-

TABLE 2 Current spectrum of popular, available NGS instruments (as of July 2015)<sup>a</sup>

Instrument	Read length (accuracy)	No. of reads/run	Run time	Run cost (\$)	No. of genomes/metagenomes per run	Comment
ILMN MiSeq	≤300 bases (high up to 250 bases)	≤20 million	≤2.5 days	≤1,500	40/4	Read pairs
ILMN NextSeq	≤150 bases (high)	≤400 million	≤1.2 days	≤5,000	400/60	Read pairs
ILMN HiSeq 2500	≤250 bases (high)	≤300 million	2.5 days	<8,000	300/80	Read pairs, rapid run mode
PB RSII	~10 kb (low for single pass)	70,000	≤4 h	≤1,200	3/TBD	Shorter (e.g., 2-kb) reads have higher quality; no paired ends
LT Ion PGM	≤400 bases	≤5 million	≤7 h	≤750	7/1	PGM318 chip; no paired ends
LT Ion Proton	≤200 bases	≤80 million	≤4 h	1,000	30/16	PI chip; no paired ends
ONT MinION	10 kb–100 kb (low)	Variable	Variable	1,000	TBD	No paired ends; beta testing

<sup>a</sup> Abbreviations: ILMN, Illumina; LT, Life Technologies; PB, Pacific Biosciences; ONT, Oxford Nanopore Technologies. TBD, to be determined. ILMN is from Illumina. LT is from <http://allseq.com/knowledgebank/sequencing-platforms/life-technologies-ion-torrent>. Genomes/run, 100× coverage of a 3-Mb genome. Metagenomes/run, 5 million reads (pairs)/metagenome.

lowing DNA extraction, a series of processing and bioinformatics steps are performed to produce NGS data for analysis (Fig. 2), analogous to the whole-genome sequence of an individual microbe described above; however, in this case the method is applied to a mixture/community of microbes. While the genomic se-

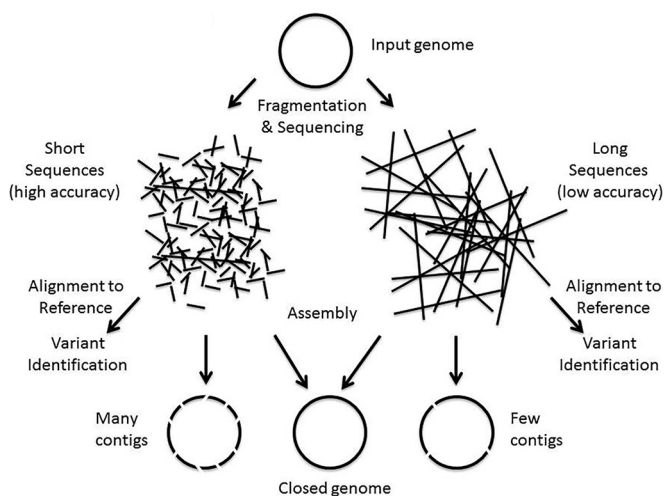


FIG 1 NGS genome analysis. The general process of using NGS for analysis of a single genome is depicted in this figure. Note that there are many variations on this approach. Purified DNA (i.e., input genome) is fragmented and run through the DNA sequencing process. The sequencing instrument produces either short (e.g., for Illumina) or long (e.g., for Pacific Biosciences) sequence reads depending on the platform used. When the goal is to produce a complete genome (e.g., for identifying virulence or antibiotic resistance genes or for comparative genomic studies), the reads are assembled into genomes using specialized software. Some gaps in the assembly may occur, leading to a draft genome sequence composed of many contigs. When the genome is assembled without gaps, it is said to be closed. When the goal is to identify variants (e.g., SNPs) with respect to a reference genome, the reads can be aligned directly with the reference genome and sequence variants can be identified using specialized programs. These variants serve to define the organism at the subspecies and strain level and are useful in epidemiological tracking as well as to identify mutations that occur.

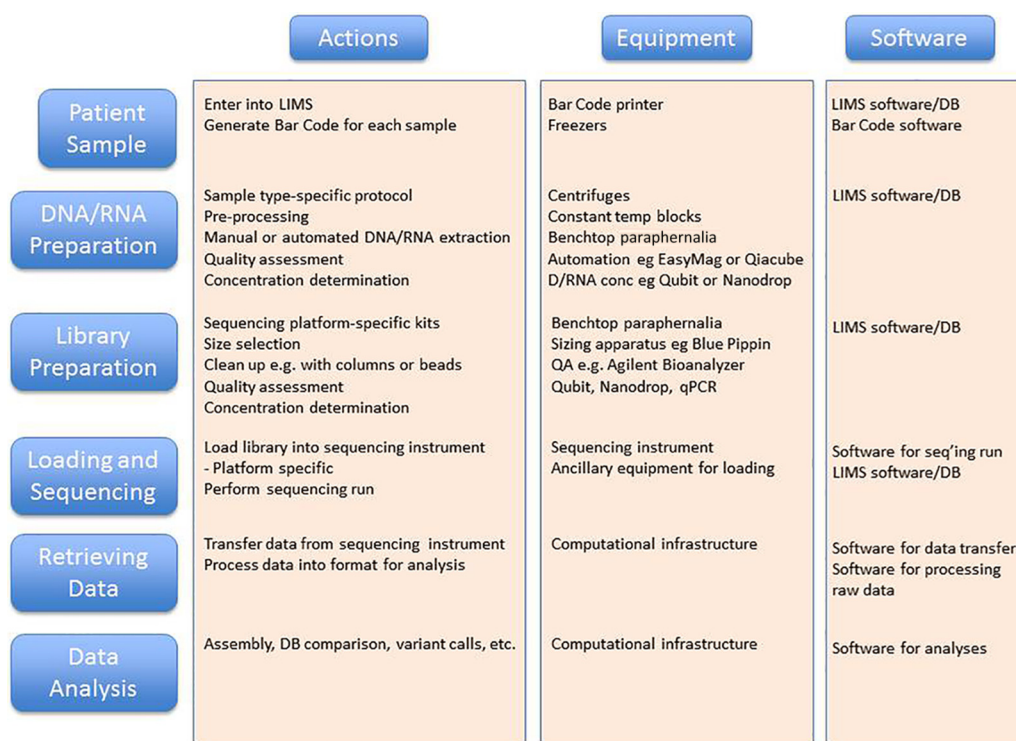
quences produced can be assembled, because of the various abundances of organisms, there is often not enough sequence from rarer organisms to produce a complete genome, and assembly results in very short contigs for these genomes in the mixture. More commonly, the sequence reads are aligned against databases of many (typically thousands) genome sequences to identify the likely source organism for the sequence. The abundance of such source organisms is estimated from the number of reads that match their genomes. In this way, one can identify and estimate the abundance of organisms that are present in the original patient sample.

The high throughput of NGS allows data to be produced from many individual bacteria/samples in a single sequencing run of hours to days. This is a vast improvement over the early days of bacterial sequencing when a single bacterial genome project could take years. Likewise, because a metagenomic sample is complex and contains hundreds of taxa at abundances ranging over many orders of magnitude, previously it was not possible to accurately assess the community structure by culture-based methods or traditional Sanger sequencing, which could handle only cultivable organisms or hundreds of sequences at best. With the availability of NGS, it is now possible to obtain millions or more of sequences per run, thus enabling deep sampling and more descriptive analysis of the microbial community on a truly useful level.

## CLINICAL APPLICATIONS

The potential clinical applications of NGS are vast and engage clinical microbiology at every level of the development process, such as outbreak tracking across the country, hospital infection control surveillance, pathogen discovery, mutation detection in a specific isolate, identification of multiple pathogens in a single sample, identification of viral quasispecies, and individual host response to infection (Fig. 3). These tests are poised to revolutionize infectious disease diagnostics and the practice of clinical microbiology. Initial reports of NGS applications for clinical diagnostics first appeared more than 5 years ago; however, the technology has not yet entered routine use within clinical micro-





**FIG 2** NGS workflow. A high-level overview of the steps taken in the NGS data production process and some of the equipment and software used in this process. The specific instruments and programs may vary as there are multiple solutions (e.g., Table 2). DB, database; QA, quality assurance; seq'ing, sequencing; conc, concentration; LIMS, laboratory information management system.

biology laboratories. Reports continue to appear in the medical literature of individual case studies or small-scale applications of NGS to clinical microbiology, but experience remains limited. The technology itself is still hobbled by the enormous amount of basic research and resources needed to transform this technology into a robust clinical application for infectious disease diagnostics. To date, the medical community is still awaiting large-scale clinical trials to establish the utility of NGS in different clinical settings; however, comprehensive funding and incentives for these studies are limited.

Although NGS has the potential to provide a huge volume of information to health care practitioners with respect to both microbial characteristics and the host response to disease, it will be necessary to determine strategies to pragmatically deal with this data barrage. Many of the potential applications to clinical medicine are not known. In the field of bacteriology, NGS has particular limitations in determining which organisms are merely “innocent bystanders” or are colonizers rather than active pathogens. For instance, a potentially pathogenic microorganism identified from one specimen type, such as cerebrospinal fluid (CSF) or synovial fluids, which are typically thought to be sterile, may be normal in the skin, gut, or mouth.

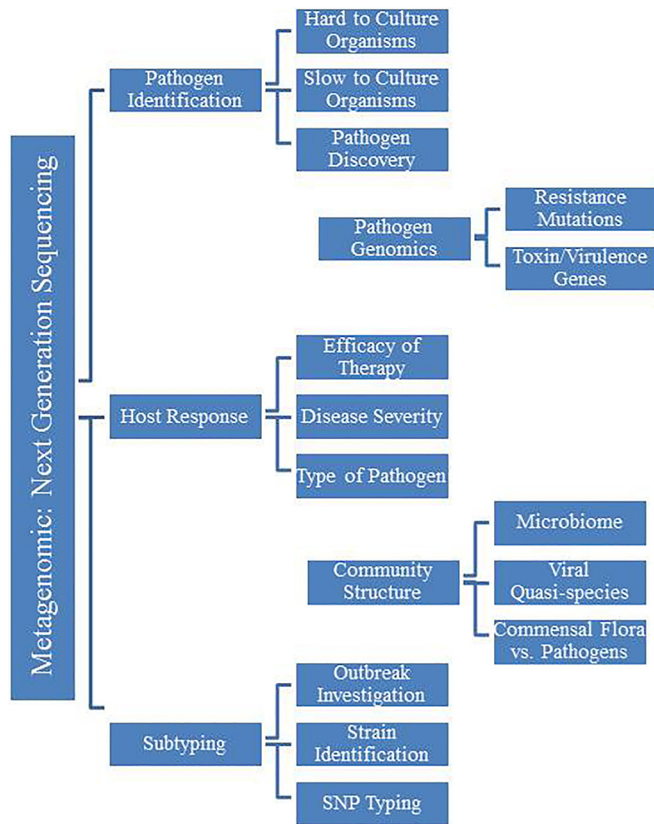
The potential implications for physicians/practitioners, and particularly antimicrobial stewardship, are staggering. One important question: if currently a certain percentage of practitioners treat patients with antibiotics even when lacking evidence of a bacterial pathogen, what will they do when presented with a pathogen that is likely (but not absolutely) a contaminant?

More importantly, a paradigm shift in the understanding of

molecular diagnostics will be necessary for clinicians and professionals in the clinical microbiology community. Both molecular diagnostics and traditional culture-based technology have their strengths and weaknesses. One of the strengths of NGS assays is the ability to detect many microorganisms directly from a patient sample without needing additional testing or *a priori* knowledge of the type of pathogen. NGS can potentially detect viruses, bacteria, yeast, fungi, and parasites without the need for additional individual testing. However, both traditional technologies and NGS assays have limitations in detection that can be clearly defined. But while culture absolutely indicates the presence of living organisms, the presence of organismal DNA is less definitive.

Ultimately, clinical studies and eventually clinical trials examining patient outcomes will be necessary to reinforce the cost savings and benefit of using NGS-based clinical tests for the diagnosis of infectious disease. A prominent issue with such studies is the scarcity of funding. Because of how expensive these analyses are to conduct, they can be unattractive to industry and require a certain amount of patience and determination from the clinical community to perform correctly and in an economically wise manner. Moreover, such studies require establishing fundamentals such as standard operating procedures for NGS diagnostics. At present, there is no standardization between studies, much less between different institutions. All aspects of the NGS diagnostic pipeline remain variable, including clinical specimen collection, sequencing parameters, data analysis/interpretation, and data reporting. Such details must be established before clinical outcomes can be systematically examined.

Determining the optimal approach to advance the use of NGS



**FIG 3** Potential clinical applications for metagenomics sequencing. There are numerous potential applications of NGS technology to the clinical microbiology laboratory. Each entry in the chart represents a potential area for the utilization of NGS diagnostics and/or future research.

poses serious challenges. For example, how would a clinical laboratory determine the analytical validation of a device that could potentially pick up any known or novel pathogens? While the majority of clinical infections are typically attributed to a finite number of bacterial, viral, and fungal pathogens, a main attraction of NGS for clinical diagnosis is the ability to capture rare or unsuspected, yet potentially actionable, pathogens. It is tremendously appealing to identify both pathogen and patient genomic sequence information for everything from virulence genes to strain-level sequence variants to an assessment of the patient response to infection or effectiveness of antibiotic therapy. The advancement of NGS technology into pathogen genomics could potentially provide supplemental descriptive or predictive information regarding the potential antimicrobial resistance gene profile of the microorganism and allow for more rapid selection of appropriate antimicrobial therapy. Although early studies have reported promising results for antimicrobial resistance detection, the most valuable piece of antimicrobial resistance information is the antimicrobials to which the organism is susceptible (15, 25), and it is unclear what the strength of NGS for predicting susceptibility as opposed to resistance will be. Additionally, as bacteria and viruses are constantly evolving new resistance mechanisms, reference databases would have to be continuously updated to include novel resistance genes and mutations in order for NGS to be an effective predictor. Thus, it seems unlikely that the clinical laboratory will be moving away from phenotypic confirmation in

the near future. Additional clinical studies are needed to explore the capabilities of NGS assays for these applications and the potential partnership between NGS and culture-based assays.

As NGS technology continues to mature, the research and medical communities will need to establish the analytical performance of each assay, the clinical validity for different pathogens, and, most importantly, the clinical niche for NGS. The shift to molecular diagnostics will require a change in the thinking of clinicians and clinical microbiology professionals. Upcoming generations of physicians, clinical providers, and clinical microbiologists will need to adjust to the strengths and limitations of these new tools. Like any new powerful technology, molecular diagnostics pose challenges in data interpretation and reporting but could more quickly and completely diagnose infection and provide critical information for clinical management. Despite all of the potential limitations and hurdles to the clinical applications of NGS assays, they present a remarkable opportunity to advance the field of clinical microbiology.

### REGULATORY CONSIDERATIONS

The imminent adaptation of NGS-based infectious disease diagnostics in the clinical microbiology laboratory poses significant regulatory challenges for U.S. device developers and manufacturers. The technology itself is still evolving, and there is a need for new development of guidelines and streamlining of insurance reimbursement codes for the use of NGS testing in clinical microbiology laboratories. While the test is going through the approval process, an upgrade or new instrument will likely appear on the market that can perform that same test with a shorter turnaround time and at a lower cost. However, constant technology improvement is not unique to clinical microbial NGS. This is one of the current obstacles where resources are scarce to develop validated tests while demand for this novel technology is rising. As mentioned in the introduction, various publications showcase situations in which microbial NGS-based tests guided or improved diagnostic and therapeutic decisions. Such decisions cannot be made with current laboratory diagnostic devices.

Infectious disease NGS-based diagnostic devices are *in vitro* diagnostic (IVD) tests intended for use in the diagnosis of diseases or other conditions, including infections. Some tests are used in laboratory or other health professional settings, and other tests are intended for use by a patient to collect samples at home. IVDs are medical devices as defined in section 210(h) of the U.S. Federal Food, Drug, and Cosmetic Act and may also be biological products subject to section 351 of the Public Health Service Act. Similarly to other medical devices, IVDs are subject to premarket and postmarket controls. IVDs are also subject to the Clinical Laboratory Improvement Amendment of 1988 (CLIA '88) (26). In the United States, clinical diagnostic tests have to make an appropriate premarket submission and obtain approval or clearance for their test from the FDA prior to marketing. Sequence-based clinical diagnostic devices for the microbiology laboratory are raising new policy and regulatory issues; thoughts presented here are preliminary and do not represent finalized FDA policy. The FDA encourages submitters to contact the FDA, using the presubmission program to discuss the premarket submission strategy for their specific test (27).

During FDA's Microbial Sequencing workshop held on 1 April 2014, scientific and clinical community leaders emphasized the benefits of regulatory oversight of infectious disease NGS-based

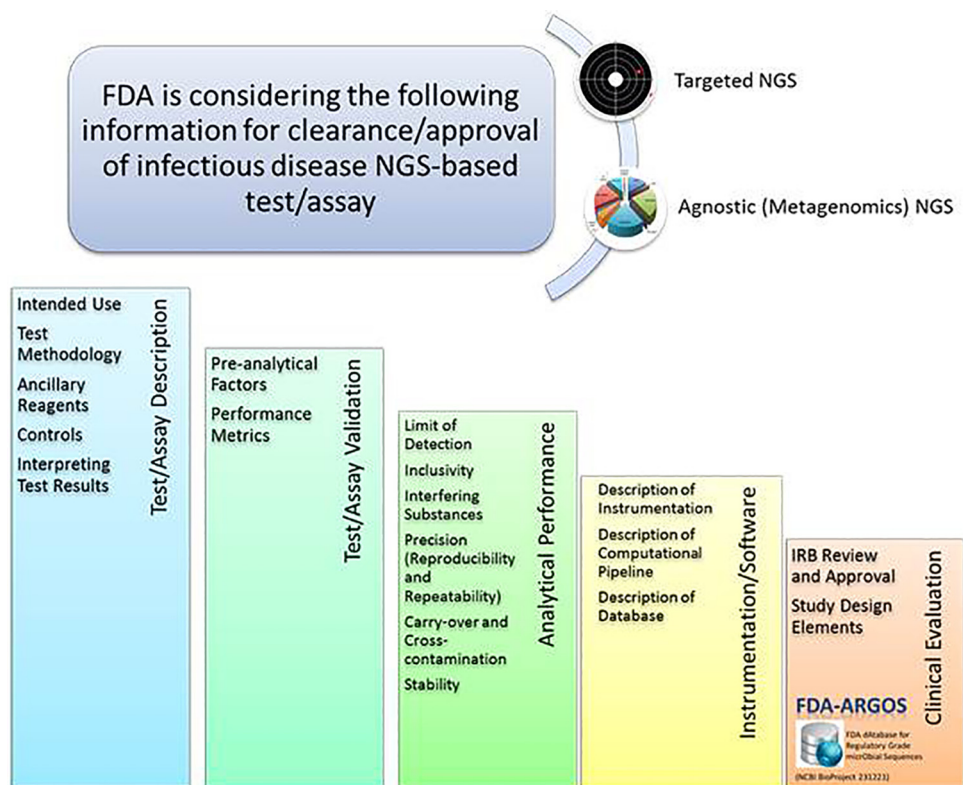


FIG 4 The FDA is considering the following information for the clearance/approval of an infectious disease NGS-based test/assay. The FDA presubmission process can be utilized for outstanding questions and to request additional information while policy is still being developed (26). IRB, institutional review board.

diagnostic devices due to challenges that these devices pose for patient management (28). The following challenges were identified at the workshop: (i) the absolute need for immediate and actionable results, (ii) the broad range of specimen types (e.g., urine, blood, CSF, stool, sputum, and others), (iii) the broad diversity of the infectious disease agents possibly present within a single specimen, and (iv) the dynamic nature of infectious disease agents. The request from scientific and clinical community leaders for guidance on infectious disease sequencing parallels the challenges faced by matrix-assisted laser desorption ionization (MALDI)–time of flight (TOF) mass spectrometry diagnostics prior to their widespread integration into the clinical microbiology laboratory.

Moreover, the publications referenced in the introduction and input from stakeholders at FDA's Microbial Sequencing workshop outlined that detection and identification of infectious disease organisms and antimicrobial resistance or virulence markers have progressed from culture-based methods to molecular methods using nucleic acid amplification and hybridization technologies. Today, single-approach high-throughput techniques or NGS may potentially replace previous methods which required several different tests. An infectious disease NGS diagnostic device differs from traditional devices that target specific organisms or virulence/resistance markers by being able to simultaneously detect every organism present in a sample during a single run. These challenges call for a novel regulatory approach tailored to the specific NGS technology and microbiology application used in the diagnostic assay. To date, the only *in vitro* diagnostic NGS system (i.e., assay and instrument) to have obtained FDA marketing au-

thorization is the Illumina MiSeqDx system. This authorization marks a significant milestone for NGS technology (29). The Illumina MiSeqDx system is tailored for the detection of cystic fibrosis (CF) (i.e., cystic fibrosis carrier screening for the general population and detection of the CF gene for CF patients); the Illumina MiSeqDx system has not been cleared by the FDA for microbial diagnostic use. Currently, the FDA is developing concepts for validation of NGS tests for infectious disease diagnostics and the detection of antimicrobial resistance and virulence markers. The FDA is also working on introducing models for streamlining clinical trials for the validation of infectious disease NGS diagnostic tests and other sequence-based microbial molecular diagnostics. Quality metrics and metadata parameters for microbial genomic sequence entries are currently being developed for use in regulatory decisions.

Since policy and regulatory issues for these devices are still evolving, the FDA presubmission process is a helpful tool for developers to inquire about specific information on studies required for evaluation of NGS-based devices. These studies are aimed at elucidating how NGS technologies can aid in infectious disease diagnostics and at gaining a better understanding of potential NGS clinical implementation strategies. The purpose of the presubmission program is to give a submitter an opportunity to discuss specific questions with the FDA regarding product development or application preparation. The FDA review of presubmission protocols leads to better-prepared submissions. Through the presubmission process, input can be given on possible approaches to validation studies and data for the evaluation of infectious disease NGS-based diagnostics (Fig. 4). The FDA can



provide guidance on the available regulatory pathways and assistance on how to evaluate performance using sequence outputs from infectious disease NGS-based devices. Quality metrics for reference databases are among the most critical validation issues as discussed by stakeholders during the FDA Microbial Sequencing workshop. Efforts toward generating an initial set of high-quality, regulatory-grade microbial genomic reference sequences through the FDA-ARGOS project in collaboration with various federal agencies are under way. The FDA's vision is a public, high-quality, regulatory-grade microbial reference database that contains qualified sequence data for use by developers and clinical end users.

As with other molecular biology-based diagnostic devices, FDA is considering using a "one-system" approach for the evaluation of infectious disease NGS diagnostic devices—from sample collection through the output of clinically actionable data. The components of the system generally include a specimen collection device, instruments, reagents, software (if applicable) used to generate the sequencing library or otherwise prepare the specimen for sequencing, the sequencing instruments along with the associated reagents and data collection elements that generate the raw sequence reads, and the data analysis pipeline (i.e., assembly, annotation, and variant calling, as applicable). Further, the FDA is considering using methods from the discipline of systems science to evaluate these devices. This approach will evaluate, in parallel, the "system" as a whole, from specimen collection to the individual steps in the sequencing data pipeline to the generation of clinically actionable data.

**FDA-ARGOS.** The FDA, in collaboration with various federal agencies, has developed the database entitled FDA-ARGOS (FDA database for regulatory-grade microbial sequences; BioProject 231221). To promote a least-burdensome regulatory approach for devices that incorporate infectious disease NGS diagnostic technology, the FDA considers the use of an alternative comparator method for clinical evaluation that relies heavily on public databases populated with regulatory-grade target genomic reference sequences. This database supplies a set of validated regulatory-grade microbial genomic sequence entries which is available at the National Center for Biotechnology Information (NCBI) website. Regulatory-grade microbial sequences are near-complete high-quality draft genomes with metadata requirements (30).

In order to advance from proof of concept to implementation of NGS-based diagnostics in U.S. clinical laboratories and hospitals, assays must undergo standardization, optimization, validation, and, eventually, automation. Once each step of the workflow has been optimized for efficiency, overall functionality and performance characteristics must be determined for the entire test. Validation requires the calculation of analytical sensitivity and specificity, accuracy, precision, and limit of detection (31). The National Institute for Standards and Technology (NIST) is developing whole-genome microbial reference standards to support the validation process for genomic sequencing-based diagnostic assays (32). Four microbial species that are clinically significant and vary in genome characteristics (e.g., size, GC content, and presence of accessory elements) are being characterized by NIST through evaluating genome assembly, base-level analysis, genomic purity, and genomic stability (33). These reference materials can be used during appropriate steps to support validation of a microbial NGS-based diagnostic assay. NGS is a multistep technology for which validation has to occur among all levels of

"wet" and "dry" processes described above. Guidance for developing specific protocols and quality control procedures that assess NGS tests is needed. For example, metrics should be established for the number of individual sequence reads that must match a genomic reference sequence or marker for a particular organism (32, 34). To aid in validating the bioinformatics or "dry" components of NGS, attendees at the AAM's colloquium suggested an *in silico*, characterized data set that could be downloaded and run through the clinical laboratory's bioinformatic pipeline. It may be necessary for a group of regulatory agencies such as the FDA, the Centers for Medicare and Medicaid Services (CMS), the National Institutes of Health (NIH), the Centers for Disease Control and Prevention (CDC), and the College of American Pathologists (CAP) to collaborate on the formulation of guidelines for the validation of infectious disease NGS-based diagnostic assays.

Another regulatory challenge that encompasses the transition of NGS into the clinical microbiology laboratory is management of patient data. Because microbial NGS data are generated from a patient's clinical sample, storage of the genomic information must be secure and uphold Health Insurance Portability and Accountability Act (HIPAA) regulations. Offsite remote cloud processing and data storage have become a popular alternative to local bioinformatics infrastructure, although no regulatory legislation is in place for these online tasks. It is well known that an established bioinformatics workflow, including storage and retrieval facilities for microbial genome sequence analyses, requires a substantial financial investment which may not be available or feasible for many clinical microbiology laboratories. Raw sequence reads consume considerable storage space, and thus, it was suggested at the colloquium to deposit raw data in the Short Read Archive (SRA) at the NCBI and the assembled/annotated genome in a central repository (35). It was also highly recommended by colloquium attendees to document only the clinically actionable result and analysis parameters in the patient's medical record, thereby decreasing the data footprint cost.

Not only do microbial genomic data hold immense value for the clinical microbiology field, but their power could be leveraged to help the public health and scientific research communities through implementation of open data sharing practices. Many discussions at the AAM's colloquium focused on the need for a publicly available, unified, genomic reference database that is filtered with metadata and maintains patient confidentiality and privacy laws. A centralized database with a continuous stream of new microbial genomes could serve as a vital resource (35, 36). It was suggested that guidelines for responsible data sharing and transfer be created by a group of relevant stakeholders to assist in maximizing public availability of microbial genomic data and to limit the creation of private genome sequence databases. Incorporation of new data repositories could occur with established databases such as the multiorganizational effort known as the International Nucleotide Sequence Database Collaboration (INSDC), an initiative that endorses data sharing among three existing databases: NCBI, European Nucleotide Archive (ENA), and DNA Data Bank of Japan (DDBJ). To truly understand the benefits that NGS diagnostic assays could have for the broader clinical microbiology community, data sharing through a constantly evolving, centralized reference repository is essential.



## CONCLUSION

Despite the enormously attractive potential of NGS for infectious disease diagnosis, there are many challenging, time-consuming, and costly bridges to cross before this technology can become mainstream and part of the clinical standard of care. However, the clinical literature has demonstrated that this technology can be successfully applied to solve medical diagnostic dilemmas and be likely useful for clinical cases that fail or challenge the limits of traditional laboratory testing. Although many potential impediments to the utilization of NGS-based diagnostics exist, it would be a loss to the medical community if this technology could not be applied to patient care in some capacity. The burden is on the infectious disease and genomics communities to bring these tools to the world of clinical diagnostics and to keep pace with the ongoing evolution of the analytical and computational aspects of NGS. Ten years ago, it would have been inconceivable that a single human genome could be sequenced in a day and for less than \$1,000. Yet, today the clinical microbiology laboratory is facing the same scenario, and one cannot fully envision all other infectious disease applications that will be possible in the near future. The rapid evolution of NGS challenges both the regulatory framework and the development of laboratory standards and will require additional funding and incentives to drive tangible improvements and progress. Cooperation among the medical and research communities, industry, regulatory bodies, and professional societies will help to develop innovative solutions to these challenges and to realize the potential benefit that NGS has for patients and their families.

## ACKNOWLEDGMENTS

We extend our grateful appreciation to AAM staff, including director Marina Moses and program assistant Daniel Peniston. We also thank all participants of the AAM colloquium for stimulating discussions and particularly ASM Secretary Joseph Campos for helping to bring the colloquium to fruition. The Academy report *Applications of Clinical Microbial Next-Generation Sequencing* is forthcoming, and the complete list of participants will be included there.

All authors contributed to the conception and design of the project. C.G., B.G., and N.L. wrote the introduction. G.M.W. wrote Next-Generation Sequencing and designed Fig. 1 and 2 and Tables 1 and 2. B.G. wrote Clinical Applications and provided Fig. 3. H.S. wrote Regulatory Considerations and created Fig. 4. All authors contributed to the editing and polishing of all sections in the final version.

## REFERENCES

- Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Goodwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer MLL, Jarvie TP, Jirage KB, Kim J-B, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna M, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM. 2005. Genome sequencing in open microfabricated high density picoliter reactors. *Nature* 437:376–380.
- Liu L, Li Y, Li S, Hu N, He Y, Pong R, Lin D, Lu L, Law M. 2012. Comparison of next-generation sequencing systems. *J Biomed Biotechnol* 2012:1–11. <http://dx.doi.org/10.1155/2012/251364>.
- Brown JR, Morfopoulou S, Hubb J, Emmett WA, Ip W, Shah D, Brooks T, Paine SML, Anderson G, Virasami A, Tong CYW, Clark DA, Plagnol V, Jacques TS, Qasim W, Hubank M, Breuer J. 2015. Astrovirus VA1/HMO-C: an increasingly recognized neurotropic pathogen in immunocompromised patients. *Clin Infect Dis* 60:881–888. <http://dx.doi.org/10.1093/cid/ciu940>.
- Wilson MR, Naccache SN, Samayoa E, Biagtan M, Bashir H, Yu G, Salamat SM, Somasekar S, Federman S, Miller S, Sokolic R, Garabedian E, Candotti F, Buckley RH, Reed KD, Meyer TL, Seroogy CM, Galloway R, Henderson SL, Gern JE, DeRisi JL, Chiu CY. 2014. Actionable diagnosis of neuroleptospirosis by next-generation sequencing. *N Engl J Med* 370:2408–2417. <http://dx.doi.org/10.1056/NEJMoa1401268>.
- Naccache SN, Peggs KS, Mattes FM, Phadke R, Garson JA, Grant P, Samayoa E, Federman S, Miller S, Lunn MP, Gant V, Chiu CY. 2015. Diagnosis of neuroinvasive astrovirus infection in an immunocompromised adult with encephalitis by unbiased next-generation sequencing. *Clin Infect Dis* 60:919–923. <http://dx.doi.org/10.1093/cid/ciu912>.
- Hoffmann B, Tappe D, Höper D, Herden C, Boldt A, Mawrin C, Niederstrasser O, Müller T, Jenckel M, van der Grinten E, Lutter C, Abendroth B, Teifke JP, Cadar D, Schmidt-Chanasit J, Ulrich RG, Beer M. 2015. A variegated squirrel bornavirus associated with fatal human encephalitis. *N Engl J Med* 373:154–162. <http://dx.doi.org/10.1056/NEJMoa1415627>.
- Quan P, Wagner TA, Briese T, Torgerson TR, Hornig M, Tashmukhamedova A, Firth C, Palacios G, Baisre-De-Leon A, Paddock CD, Hutchison SK, Egholm M, Zaki SR, Goldman JE, Ochs HD, Lipkin WI. 2010. Astrovirus encephalitis in boy with X-linked agammaglobulinemia. *Emerg Infect Dis* 16:918–925. <http://dx.doi.org/10.3201/eid1606.091536>.
- GenomeWeb. 7 August 2015. CDC earmarks \$2.3 million for NGS, bioinformatic approaches to combat infectious disease. GenomeWeb, New York, NY. <https://www.genomeweb.com/research-funding/cdc-earmarks-23m-ngs-bioinformatic-approaches-combat-infectious-disease>.
- Turabelidze G, Lawrence SJ, Gao H, Sodergren E, Weinstock GM, Abubucker S, Wylie T, Mitreva M, Shaikh N, Gautom R, Tarr PI. 2013. Precise dissection of an *Escherichia coli* O157:H7 outbreak by single nucleotide polymorphism analysis. *J Clin Microbiol* 51:3950–3954. <http://dx.doi.org/10.1128/JCM.01930-13>.
- Rasko DA, Webster DR, Sahl JW, Bashir A, Boisen N, Scheutz F, Paxinos EE, Sebra R, Chin C, Iliopoulos D, Klammer A, Peluso P, Lee L, Kislyuk AO, Bullard J, Kasarskis A, Wang S, Eid J, Rank D, Redman JC, Steyert SR, Frimodt-Møller J, Struve C, Petersen AM, Krogfelt KA, Nataro JP, Schadt EE, Waldor MK. 2011. Origins of the *E. coli* strain causing an outbreak of hemolytic–uremic syndrome in Germany. *N Engl J Med* 365:709–717. <http://dx.doi.org/10.1056/NEJMoa1106920>.
- Chin C, Sorenson J, Harris JB, Robins WP, Charles RC, Jean-Charles RR, Bullard J, Webster DR, Kasarskis A, Peluso P, Paxinos EE, Yamai-chi Y, Calderwood SB, Mekalanos JJ, Schadt EE, Waldor MK. 2011. The origin of the Haitian cholera outbreak strain. *N Engl J Med* 364:33–42. <http://dx.doi.org/10.1056/NEJMoa1012928>.
- Azarian T, Cook RL, Johnson JA, Guzman N, McCarter YS, Gomez N, Rathore MH, Morris JG, Salemi M. 2015. Whole-genome sequencing for outbreak investigations of methicillin-resistant *Staphylococcus aureus* in the neonatal intensive care unit: time for routine practice? *Infect Control Hosp Epidemiol* 36:777–785. <http://dx.doi.org/10.1017/ice.2015.73>.
- Snitkin ES, Zelazny AM, Thomas PJ, Stock F, Henderson DK, Palmore TN, Segre JA. 2012. Tracking a hospital outbreak of carbapenem-resistant *Klebsiella pneumoniae* with whole-genome sequencing. *Sci Transl Med* 4:148ra116. <http://dx.doi.org/10.1126/scitranslmed.3004129>.
- Dunne WM, Jr, Westblade LF, Ford B. 2012. Next-generation and whole-genome sequencing in the diagnostic clinical microbiology laboratory. *Eur J Clin Microbiol Infect Dis* 31:1719–1726. <http://dx.doi.org/10.1007/s10096-012-1641-7>.
- Stoesser N, Batty EM, Eyre DW, Morgan M, Wyllie DH, Del Ojo Elias C, Johnson JR, Walker AS, Peto TEA, Crook DW. 2013. Predicting antimicrobial susceptibilities for *Escherichia coli* and *Klebsiella pneumoniae* isolates using whole genomic sequence data. *J Antimicrob Chemother* 68:2234–2244. <http://dx.doi.org/10.1093/jac/dkt180>.
- Arias CA, Panesso D, McGrath DM, Qin X, Mojica MF, Miller C, Diaz L, Tran TT, Rincon S, Barbu EM, Reyes J, Roh JH, Lobos E, Sodergren E, Pasqualini R, Arap W, Quinn JP, Shamoo Y, Murray BE, Weinstock GM. 2011. Genetic basis for in vivo daptomycin resistance in enterococci. *N Engl J Med* 365:892–900. <http://dx.doi.org/10.1056/NEJMoa1011138>.
- Tran TT, Panesso D, Gao H, Roh JH, Munita JM, Reyes J, Diaz L, Lobos EA, Shamoo Y, Mishra NN, Bayer AS, Murray BE, Weinstock GM, Arias CA. 2013. Whole-genome analysis of a daptomycin-susceptible *Enterococcus faecium* strain and its daptomycin-resistant vari-

- ant arising during therapy. *Antimicrob Agents Chemother* 57:261–268. <http://dx.doi.org/10.1128/AAC.01454-12>.
18. Fisher R, van Zyl GU, Travers SAA, Kosakovsky Pond SL, Engelbrech S, Murrell B, Scheffler K, Smith D. 2012. Deep sequencing reveals minor protease resistance mutations in patients failing a protease inhibitor regimen. *J Virol* 86:6231–6237. <http://dx.doi.org/10.1128/JVI.06541-11>.
  19. Ley RE, Bäckhed F, Turnbaugh P, Lozupone CA, Knight RD, Gordon JI. 2005. Obesity alters gut microbial ecology. *Proc Natl Acad Sci U S A* 102:11070–11075. <http://dx.doi.org/10.1073/pnas.0504978102>.
  20. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, Mende DR, Li J, Xu J, Li S, Li D, Cao J, Wang B, Liang H, Zheng H, Xie Y, Tap J, Lepage P, Bertalan M, Batto JM, Hansen T, Le Paslier D, Linneberg A, Nielsen HB, Pelletier E, Renault P, Sicheritz-Ponten T, Turner K, Zhu H, Yu C, Li S, Jian M, Zhou Y, Li Y, Zhang X, Li S, Qin N, Yang H, Wang J, Brunak S, Doré J, Guarner F, Kristiansen K, Pedersen O, Parkhill J, Weissenbach J, MetaHIT Consortium, Bork P, Ehrlich SD, Wang J. 2010. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464:59–65. <http://dx.doi.org/10.1038/nature08821>.
  21. Graessler J, Qin Y, Zhong H, Zhang J, Licinio J, Wong M, Xu A, Chavakis T, Bornstein AB, Ehrhart-Bornstein M, Lamounier-Zepter V, Lohmann T, Wolf T, Bornstein SR. 2013. Metagenomic sequencing of the human gut microbiome before and after bariatric surgery in obese patients with type 2 diabetes: correlation with inflammatory and metabolic parameters. *Pharmacogenomics J* 13:514–522. <http://dx.doi.org/10.1038/tpj.2012.43>.
  22. Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, DuGar B, Feldstein AE, Britt EB, Fu X, Chung Y, Wu Y, Schauer P, Smith JD, Allayee H, Tang WHW, DiDonato JA, Lusis AJ, Hazen SL. 2011. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* 472:57–63. <http://dx.doi.org/10.1038/nature09922>.
  23. Price KE, Hampton TH, Gifford AH, Dolben EL, Hogan DA, Morrison HG, Sogin ML, O'Toole GA. 2013. Unique microbial communities persist in individual cystic fibrosis patients throughout a clinical exacerbation. *Microbiome* 1:27. <http://dx.doi.org/10.1186/2049-2618-1-27>.
  24. Koren S, Schatz MC, Walenz BP, Martin J, Howard JT, Ganapathy G, Wang Z, Rasko DA, McCombie WR, Jarvis ED, Phillippy AM. 2012. Hybrid error correction and de novo assembly of single-molecule sequencing reads. *Nat Biotechnol* 30:693–700. <http://dx.doi.org/10.1038/nbt.2280>.
  25. Gordon NC, Price JR, Cole K, Everitt R, Morgan M, Finney J, Kearns AM, Pichon B, Young B, Wilson DJ, Llewelyn MJ, Paul J, Peto TEA, Crook DW, Walker AS, Golubchik T. 2014. Prediction of *Staphylococcus aureus* antimicrobial resistance by whole-genome sequencing. *J Clin Microbiol* 52:1182–1191. <http://dx.doi.org/10.1128/JCM.03117-13>.
  26. US Food and Drug Administration. 19 March 2015. Overview of IVD regulation. US Food and Drug Administration, Silver Spring, MD. <http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/IVDRegulatoryAssistance/ucm123682.htm>.
  27. US Food and Drug Administration. 18 February 2014. Requests for feedback on medical device submissions: the pre-submission program and meetings with Food and Drug Administration staff. US Food and Drug Administration, Silver Spring, MD. <http://www.fda.gov/ucm/groups/fdagov-public/@fdagov-meddev-gen/documents/document/ucm311176.pdf>.
  28. US Food and Drug Administration. 11 July 2014. Public workshop—advancing regulatory science for high throughput sequencing devices for microbial identification and detection of antimicrobial resistance markers, April 1, 2014. US Food and Drug Administration, Silver Spring, MD. <http://www.fda.gov/MedicalDevices/NewsEvents/WorkshopsConferences/ucm386967.htm>.
  29. Collins FS, Hamburg MA. 2013. First FDA authorization for next-generation sequencer. *N Engl J Med* 369:2369–2371. <http://dx.doi.org/10.1056/NEJMp1314561>.
  30. US Food and Drug Administration. 11 December 2013. FDA database for regulatory grade microbial sequences (FDA-ARGOS): supporting development and validation of infectious disease Dx tests. National Center for Biotechnology Information, Bethesda, MD. <http://www.ncbi.nlm.nih.gov/bioproject/231221>.
  31. Aziz N, Zhao Q, Bry L, Driscoll DK, Funke B, Gibson JS, Grody WW, Hegde MR, Hoeltge GA, Leonard GDB, Merker JD, Nagarajan R, Palicki LA, Robertoye RS, Schrijver I, Weck KE, Voelkerding KV. 2015. College of American Pathologists' laboratory standards for next-generation sequencing clinical tests. *Arch Pathol Lab Med* 139:481–493. <http://dx.doi.org/10.5858/arpa.2014-0250-CP>.
  32. Sichtig H. 2014. High-throughput sequencing technologies for microbial identification and detection of antimicrobial resistance markers, p 1–21. In FDA Public Workshop on High Throughput Sequencing Devices in Microbial Diagnostics. US Food and Drug Administration, Silver Spring, MD.
  33. Olson N, Zook J, Vang L, Catoe D, McDaniel J, Salit M. 11 April 2014. Microbial genomic reference materials. US Food and Drug Administration, Silver Spring, MD. <http://www.fda.gov/downloads/MedicalDevices/NewsEvents/WorkshopsConferences/UCM393056.pdf>.
  34. Rehm HL, Bale SJ, Bayrak-Toydemir P, Berg JS, Brown KK, Deignan JL, Friez MJ, Funke BH, Hegde MR, Lyon E. 2013. ACMG clinical laboratory standards for next-generation sequencing. *Genet Med* 15:733–747. <http://dx.doi.org/10.1038/gim.2013.92>.
  35. Fricke WF, Rasko DA. 2014. Bacterial genome sequencing in the clinic: bioinformatic challenges and solutions. *Nat Rev Genet* 15:49–55. <http://dx.doi.org/10.1038/nrg3624>.
  36. Didelot X, Bowden R, Wilson DJ, Peto TEA, Crook DW. 2012. Transforming clinical microbiology with bacterial genome sequencing. *Nat Rev Genet* 13:601–612. <http://dx.doi.org/10.1038/nrg3226>.