Biomolecular Detection and Quantification 14 (2017) 1–6

Contents lists available at ScienceDirect

Biomolecular Detection and Quantification

journal homepage: www.elsevier.com/locate/bdq

Review Article

Next-generation sequencing applications in clinical bacteriology

Yair Motro[a], Jacob Moran-Gilad[a,b,c,*]

[a] Department of Health System Management, Faculty of Health Sciences, Ben-Gurion University of the Negev, Beer-Sheva, Israel
[b] Public Health Services, Ministry of Health, Jerusalem, Israel
[c] ESCMID Study Group for Genomic and Molecular Diagnostics (ESGMD), Basel, Switzerland

ARTICLE INFO

Handled by Jim Huggett

Keywords:
Next generation sequencing
Clinical microbiology
Clinical diagnostics
Public health
Applications
Workflow

ABSTRACT

With the rapid advances in next generation sequencing (NGS) technologies, clinical and public health microbiology laboratories are increasingly adopting NGS technology in their workflows into their existing diagnostic cycles. In this bacteriology focused review, we review aspects and considerations for applying NGS in the clinical microbiology settings, and highlight the impact of such implementation on the analytical and post-analytical stages of diagnosis

1. Introduction

Clinical and public health microbiology laboratories help to lessen the burden of infectious disease by detecting and characterizing pathogens in infected patients or those pathogens circulating in the community. In this scenario, implementation of next generation sequencing (see Heather and Chain [1] for an encompassing review of NGS technologies) can potentially assist in clinical and public health decisions by determining the causative agent of infectious disease and/or the epidemiology and evolution of various infecting pathogens in the hospital or community settings [2]. With its multitude of benefits, NGS is becoming the gold standard in bacteriology, however since it is not yet fully accessible (particularly in low resource settings), currently NGS is mainly used at a level of reference microbiology rather than routine [3].

Traditionally, the clinical diagnostic cycle consists of three phases, namely, a pre-analytical phase (including, patient visit, examination and provisional diagnosis, collection of sample, identity, requisition, transport, records), an analytical phase (including, macroscopy, microscopy, culture, biochemical identification, serology, molecular analysis), and post-analytical phase (including, reporting of identification and antibiotic susceptibility testing, clinical interpretation of the results and patient treatment). With the rapid advances in NGS technologies and capabilities, clinical microbiologists are recognising that the influence of NGS on the diagnostic cycle will be in the scale of a “disruptive technology”, potentially reducing the time from diagnosis to clinical treatment, while also reducing the requirement for wet laboratory-based analyses performed in tandem. In addition, a NGS-based analytical phase will provide the opportunity to apply a broad repertoire of tools, including subtyping, resistome and virulome mapping, phenotypic inference, detection of new variants and toxins, among others [4]. This review will focus on the application of NGS in the clinical microbiology context, with emphasis on the potential role of NGS in the analytical and post-analytical phases of the clinical diagnostic cycle. It should be noted that this review focuses on bacteriology, though the importance of NGS in virology and other fields is of no lesser extent. So too, the NGS technology of focus in this review will be on the currently more widely used second generation sequencing [1].

2. The NGS workflow

Generally, the clinical laboratory workflow may be divided into several stages, namely, pathogen detection, identification, drug susceptibility, epidemiological typing [4], and detection of toxins and virulence factors that have clinical or prognostic implications. Note that bacterial and fungal isolates are detected through these steps, but virus detection and characterisation mainly relies on PCR-based assays. Also, each step involves a range of specialised tests that must be performed individually on each isolated organism [4].

There are several common steps that are shared among the majority of NGS methods, with the exception of single-molecule real-time NGS. A typical NGS workflow in a clinical laboratory includes: sample collection and preparation, nucleic acid extraction, NGS library preparation, sequencing, data analysis, and data storage [4,5].

* Corresponding author at: Dept. of Health Systems Management, Faculty of Health Sciences, Ben-Gurion University of the Negev POB 653, Beer-Sheva, 8410501, Israel.
E-mail address: giladko@post.bgu.ac.il (J. Moran-Gilad).

http://dx.doi.org/10.1016/j.bdq.2017.10.002
Received 2 February 2017; Received in revised form 1 October 2017; Accepted 12 October 2017
Available online 23 October 2017

2.1. Clinical sample

The clinical sample (for example, swab, sputum, stool, urine, or tissue, that contains the nucleic acid (DNA/RNA) of interest) is stored and transported to the clinical microbiology laboratory. The clinical specimen type depends on the patient's clinical syndrome, where ideally sample are collected during disease progression. Notably, the eventual NGS method that will be required (for example, whole genome sequencing (WGS) or metagenomics) will also influence the type of treatment of the sample (i.e. sample culturing or direct sample extraction).

2.2. Nucleic acids preparation

Nucleic acids can be prepared from clinical samples by using a variety of methodologies, some of which are dependent on the NGS system being used. Correct method selection is essential for a successful result, thereby lessening the introduction of biases and false negatives. Notably, the DNA quantity and quality required for NGS are higher than standard molecular assays and manual or automated platforms used for routine extraction for molecular diagnostics are not always fit for NGS. From our experience, extraction for NGS commonly requires tailoring or calibrating methods, at least for certain bacterial species, since a 'one size fits all' method is not readily applicable.

2.3. Nucleic acids sequencing

In general, two main NGS platform methods are currently used, namely short read platforms (including Illumina and Ion Torrent) and long reads platforms, including single molecule real time sequencing (Pacific Biosciences) and nanopore (Oxford Nanopore) sequencing (see Levy and Myers [6] and Kwong et al. [7] for detailed reviews and comparisons of these platforms). The input nucleic acid (for example, genomic DNA, reverse-transcribed RNA or cDNA, immunoprecipitated DNA) is firstly fragmented by methods such as sonication, nebulisation, or enzymatic digestions [8]. The fragments are then ligated to platform-specific oligonucleotide adapters to create a library of overlapping sequences, which is then hybridised to beads or a flow cell, followed by clonal amplification, such as emulsion PCR or bridge amplification (note that not all platforms require the clonal amplification phase or preparation of a DNA library). Enrichment procedures can also be completed at this stage to help select for a specific type of DNA if an organism is suspected. Of note, Becker et al. [9] compared six bacterial DNA extraction kits for a subsequent MiSeq sequencing run of a clinical *Klebsiella pneumoniae* sample, and noted that the choice of extraction kit had little effect on sequencing read quality and sequencing coverage, rather the extraction costs, extraction time, robustness and reproducibility as well as the potential for automation are the main factors for selecting a fitting extraction procedure.

2.4. Sequence data analysis

Depending upon the NGS platform, the clonally amplified templates are sequenced by various chemistries (such as pyrosequencing, reversible dye terminators, oligonucleotide probe ligation, and phospholinked fluorescent nucleotides), and following quality control and assurance of the sequence data, analysis is preformed to determine the composition of the DNA sequences for pathogen identification.

2.5. Data release and clinical report

The final stages of the NGS workflow are data release and dissemination of a clinically actionable report. Appropriate NGS analysis files should be stored or archived on- or off-site with patient privacy/confidentiality upheld [8], allowing for future re-examination upon request.

3. Impact of the NGS workflow on clinical microbiology

3.1. Less technical laboratory involvement

With the progression of the sample in the clinical laboratory workflow, the involvement of the hands-on technician at each successive step is required, particularly where additional challenges are posed by particular organisms, some of which may be of critical public health importance. For example, *Mycobacterium tuberculosis* complex bacteria are extremely slow growing, taking weeks to 1–2 months to achieve susceptibility results, thus delaying appropriate treatment and potentially negatively impacting the patient outcome [10]. Furthermore, many aetiological agents, such as *Borrelia burgdorferi* (causative agent of Lyme disease), *Bartonella* species, *Mycobacterium leprae*, and HIV elude conventional testing altogether [11] (for an example, see the latest developments in HIV clinical treatment and surveillance using NGS reported by Metzner [12] and Berg et al. [13]). Here NGS technologies may be applied for the identification of unculturable or difficult-to-culture microorganisms, including fastidious bacteria, anaerobes, and possible bioterrorism agents [10,14].

3.2. Reduction in patient diagnosis time

In addition, in the case of patient care, where time is critical, rapid infection identification and diagnosis is imperative. For example, in the case of encephalitis up to 60% of acute cases go undiagnosed, possibly due to a lack of assays that can test for the more than 100 aetiological agents that may cause the disease [11] as well as non-infectious aetiologies. In turn, prior to knowing the infecting pathogen, clinicians commonly are forced to educated guesses regarding the therapy, consequently leading to delays and a risk of ineffective treatment and further spread of infection. Furthermore, administration of broad-spectrum empirical therapy may be opted, potentially causing "collateral damage" by eliminating helpful gut microbiota while also accelerating antimicrobial resistance development. Here too NGS holds significant promise, offering potential faster and more reliable detection methods.

3.3. Wider diagnostics repertoire

NGS offers the capability of identifying a variety of organisms—bacterium, virus, fungus, yeast, or parasite, as opposed to a variety of individual tests traditionally required to identify a pathogen [4]. Unbiased or agnostic NGS amplifies all nucleic acids present in a clinical sample, including both host and microbes, without requiring primers for targeted amplification, and can potentially generate microbial sequence data for real-time patient management [4], providing great potential to impact patient care by assisting the customization of patient treatment, while in turn reducing the usage of ineffective drugs and selective pressure for resistance development [15].

3.4. Further benefits from the NGS workflow

In an encompassing review [16], five main areas of benefit for clinical microbiologists from the applications of NGS were identified, including (a) clinical identification from primary samples or a pure culture [17], (b) infection control actions [18], (c) antimicrobial stewardship [19], (d) outbreak investigation in community and hospital settings to guide measures for containment [20], and (e) pathogen discovery [11]. Furthermore, in contrast to other microbial pathogen identification techniques, NGS metagenomics is not restricted to known organismal sequences, thus allowing for comprehensive pathogen detection without *a priori* knowledge of the target organism [21]. Additional NGS benefits include organism differentiation, novel organism discovery, virulence factors and resistance markers elucidation, host response characteristics to the offending microbe and administered

therapy, and operational aspects, including strain tracking and hospital infection control surveillance.

4. Applied NGS in clinical microbiology

4.1. NGS turnaround times: from sample to result

4.1.1. NGS on raw clinical samples

Rapid diagnostic testing on direct clinical samples has been widely utilized using various PCR and quantitative PCR methods, and rapid multiplex assays technologies [22,23]. Initially, NGS required the isolation of a bacterial species from a pure culture, however, advance have been made allowing direct detection from clinical samples [24], thus potentially reducing turnaround times from days or weeks to only a few hours, allowing such a procedure to be completed within an average clinical laboratory workday [4]. Though to date such studies involving direct clinical specimen sequencing or metagenomics have encountered difficulties, such as contaminating normal human microbiota and low-copy-number pathogens that require further evaluation, much effort is still being invested, particularly since NGS can establish a cause of infection and provide a potential actionable answer in cases where other technologies may not [11,20].

4.1.2. Reduced sequencing time

The first bacterial genome to be fully sequenced using Sanger sequencing was *Haemophilus influenzae*, required more than 1 year to complete at an estimated cost of $1 million [7]. With NGS, an entire bacterial genome can be sequenced in less than 1 day for less than $100 (depending on the platform used) [14]. On the other hand, when compared to existing molecular diagnostic assays that can take minutes to a few hours, NGS tests are relatively slow, though recent advances are proving that the par is reducible (for example, MinION [25]). Given such a reduction in time and cost, the main challenge lies in the NGS data, with the requirement for elucidation of clinically tangible information desired by the health professional.

4.1.3. Phenotypic resistance detection

Data obtained through NGS can be utilised to infer or predict phenotypic resistance (using known genetic determinants of antimicrobial resistance) in addition to pathogen identification (a process in which traditional antimicrobial susceptibility testing requires an extra day of laboratory workup, giving a turnaround time of 3 or more days) prompting rapid antibiotic treatment decision-making for the clinician and the patient (see Westblade et al. [26] for a concise review). However, currently phenotypic susceptibility prediction via NGS is still work in-progress [27–29]. Correlation between genotypic data and a clinical phenotype is not always necessarily true, for instance, mechanisms that involve inducible resistance, gene expression and regulation, post-translational modifications or combinations thereof. As a result, standardised growth-based susceptibility testing and rapid phenotypic testing methods are anticipated to be initially necessary to confirm an NGS result. Moreover, inferring the bacterial MIC on the basis of different combinations of resistance determinants is not readily achievable [30]. Similarly, a transcriptome-proteome combination may assist in extrapolating the genotypic and phenotypic connections [31]. In any case, NGS can certainly provide more information than other methods, though it is of note that the transition to NGS may create new challenges of reverse compatibility with older methods, thus requiring software solutions [32].

4.2. NGS in outbreak scenarios

4.2.1. Pathogen identification

Fast identification of the infecting pathogen and contaminated sources is a key to outbreak containment, in which NGS holds exceptional potential, for examples see Snitkin et al. [18], and McGann et al. [33] who achieved in 48 h (in a small-scale NGS laboratory) actionable reports of an unfolding vancomycin resistant *Enterococcus faecium* outbreak. Firstly, due to NGS' rapid pathogen detection when compared to traditional laboratory workup [4], thus providing the opportunity for quicker intervention strategies, such as patient isolation, contact precautions, or decolonisation. Secondly, due to NGS' vital pathogen information inference capabilities, providing direct clinical value and supporting individual and improved patient management via more accurate diagnosis and tailored therapy (potentially decreasing hospital stays, deterring AMR development, and ultimately decreasing mortality rates). Finally, by helping to promote infection control accountability (both within and between hospitals), NGS may potentially allow for curtailing the spread and infection rates of high-risk organisms in healthcare settings [27]. For example, determining whether the same bacterial strain was responsible for an outbreak in geographically distinct health care facilities [2].

4.2.2. Pathogenicity biomarkers

Another powerful outbreak containment-associated function that applied NGS can provide is an insight into how bacterial populations respond to drug treatment, by determining bacterial virulence- or antibiotic resistance-associated genetic variation biomarkers, that could help infection control experts to prepare for current, emerging, or predicted infectious threats [18].

4.2.3. Outbreak detection and intervention

Beyond the hospital setting, NGS could assist in public health interventions in local and global community outbreaks. For example, Jackson et al. [34] reported the application of NGS to assist in identifying in real-time more than four times the number of listeriosis outbreaks identified by the traditional gold standard method. By providing the source and mode of pathogen transmission associated with a foodborne disease outbreak, NGS can then drive a proper public health response in the community, such as a food recall [35]. Current bacterial outbreak typing methods include PFGE, multilocus sequence typing (MLST), and multilocus variable-number tandem-repeat analyses (MLVA), all of which have proven effective, though their limited genetic resolution has rendered a substantial proportion of outbreaks unresolved [34]. These analyses target only selected genomic regions, as opposed to NGS which can provide a whole genome analysis, such as the core genome MLST approach, with increased resolution and detailed linkages for outbreak responsible isolates [36].

NGS provides confidence in the matches among relatively few isolates across time and space, allowing for early intervention, and providing links from past isolates from known foods and environmental samples. Unambiguously evaluation of the relatedness of isolates, coupled with an estimated mutation rate, allow for calculation of a common ancestor, which in turn, with epidemiological data (such as the patient's hospital admittance date) can help predict transmission event occurrence, thus allowing for better targeting of infection control resources [28], even to the extent of an international spread of pathogens of public health importance, for example *B. pertussis* [37].

Finally, NGS may be used to identify microevolutionary differences among clinical outbreak isolates, confirming or rejecting associations between isolates, a discrimination not offered by existing approaches [38]. Also, directed NGS data mining may unravel novel infectious agents and/or new targets, supporting outbreak investigations involving highly clonal pathogens. To summarize, the rapid and open release of NGS data will transform expert response to outbreaks, particularly by means of global and real-time data sharing, for example see GenomeTrakr project housed at the National Center for Biotechnology Information (NCBI, NLM, NIH) (http://www.ncbi.nlm.nih.gov/bioproject/183844). Notably, a drawback of NGS is the lack of standardised guidelines or models endorsed by regulatory agencies for responsible data sharing, combined with hesitation from hospitals (due to privacy issues). Subsequently, open data sharing is difficult, while

additionally, the diversity of sequencing platforms, bioinformatics tools and applications used are also a risk that needs to be managed [39], thus limiting the public health benefit of applied NGS in the clinic [40].

4.3. NGS interpretation pipeline

NGS data analysis is a non-trivial, yet critical, task that may result in a clinically actionable result, impacting patient care [41]. In their review, Orsini et al. [41] defined seven logical sections of the NGS data analysis workflow, though in general one may consider several broad categories, namely pre-processing, base calling, alignment of the sequence to a reference sequence/database or *de novo* assembly, and annotation and variant detection (see Orsini et al. [41] for a detailed description of the Galaxy workflow stages). Essentially, the data analysis process is end-user driven, since many different analysis paths exist, for example, in a clinical microbiological laboratory, full-genome assembly and annotation are not a necessarily required step. Though, it should be noted that each step of NGS analysis may employ a different tool or algorithm, leading to a potentially time expensive exercise. For example, algorithmic differences may exist between one company's different platforms, such as Illumina's NextSeq, MiSeq and HiSeq. Also for non-technical laboratory staff and healthcare professional's bioinformatics training would be required before NGS technologies could be routinely used in the clinical laboratory.

Following the completion of the sequencing phase, the subsequent raw sequencing reads undertake pre-processing to remove adapter and low-quality sequences, base calling, assembly of the genome, alignment to known and curated genomic sequences, and annotation/comparative phylogenetics [41,42], the following is a concise summary of these steps:

- Pre-processing: to allow for tracking and sorting of the large number of raw sequence reads obtained from the NGS platform, a unique 4–12 bp DNA sequence is ligated to each read. This multiplexing stage of the sequencing generates "barcodes" that are generally part of the PCR primer or adapter, which prior to sequence assembly and analysis require removal (termed "demultiplexing"). The demultiplexed sequence reads are then "trimmed" from the remaining adapter sequences and low quality sequence regions (according to user selected quality stringency) [43,44].
- Base calling: each base call is assigned a predicted level of accuracy, with indications of mixed signals or other read errors filtered or removed. This is initially performed by platform proprietary software, and consequently, the generated quality scores cannot be directly compared, since each platform uses a different algorithm [45]. Recent reports have shown progress in cross-platform data analysis (see Kaas et al. [46], for example, where they developed software that can perform phylogeny analysis from multiple platform data);
- Genome assembly: involves the creation of contigs by overlapping the numerous short reads (30–500 bp depending on the platform, see above) or long reads (> 5000 bp) generated from the millions of sequencing read fragments (Note that longer contigs provide a higher depth of sequence coverage, improving accuracy and sensitivity of detecting pathogens). The contigs are then aligned and compared against an available database of high-quality sequenced organisms (for example, Kraken, MG-RAST, MetaPhlan, MEGAN). Genome assembling software selection is dependent on the sequencing method type, for example, either long or short read sequencing (see Koren and Phillippy [47] for a list of genome assembly programs). Of note, there are bioinformatics solutions tailored for microbiology which allow inference directly from reads.
- De novo assembly: genomes without a close relative in the reference database may undergo *de novo* assembly for newly discovered genomes, a method that is also useful for the discovery of mobile accessory elements, such as plasmids and phages, which may not have been present in the reference genomes. Such assembled de novo genomes can be assembled by programs such as SPAdes, IDBA-UD, and Velvet [42]. Notably, the assembly software for this procedure is technology dependent, since it is critical to produce long overlapping reads. Also, a reference database assists in accurately identify the organism and the discovery of novel variants and genes.
- *Annotation*: users are often interested in finding genes of biological importance, such as genes conferring antibiotic resistance and virulence, therefore whole/complete or even partial/draft genome annotation may be undertaken. Common bacterial genome annotation web-based tools such as the annotation tools Prokka, RAST and Glimmer may be applied [42], where in RAST, the ordered contigs are uploaded and predicted open reading frames for likely genes are identified, creating a highly annotated genome. Annotation quality is largely dependent on the gene database applied.
- *Comparative genome analysis*: the comparison by alignment of multiple genomes is used to obtain regions of sequence homology or to identify unique regions of the genomes [35]. Genetic anomalies such as genomic recombination and horizontal gene transfer could also be identified and used to explain, for example, chronic polyclonal infections mechanisms [48]. Examples of such software tools that enable annotated genomes to be compared and visualized include BRIG, Mauve, and ACT. Notably, there is a wide range of tools that enable bacterial phylotyping for public health surveillance and epidemiological studies [7]. These may include tools for SNP calling (for example see the tool Snippy [49], and a thorough review by Perez-Loada et al. [50]) or tools focusing on allele calling and gene-by-gene comparisons, allowing for high resolution phylogenetic inference via comparisons between several house-keeping genes (MLST), core genes (cgMLST), or even whole genomes (wgMLST, for example see the chewBBACA pipeline [51]),

4.4. Clinical relevance of NGS data: integration into care

Ongoing dialogue between clinicians, clinical microbiologists, and bioinformaticians is essential in order to provide extrapolation of clinical relevancy in a time-efficient manner. The clinician should receive data in a gradual and logical transition, using a familiar language based on traditional tests and assays. With the progression in this dialogue, more complex analyses may be presented. NGS output must be condensed into a reportable result, without any unnecessary information [52]. Due to the complexity of NGS data, it is easy to "select" an answer and make a false conclusion. Therefore, uncertainty will still be held concerning NGS generated results even with the great advancements and this requires thorough validation, training and ongoing communication. In turn, communication of an uncertain result to the clinician needs to be carefully considered and approached.

Potentially, NGS methods can both detect, identify and characterise infectious agents in a single run, for example, a single NGS assay may provide information regarding antibiotic resistance mechanisms and virulence determinants [4]. As mentioned earlier, antibiotic resistance prediction from genetic data is still in the early stages of development, with a limited number of cases showing the direct genotypic-phenotypic link [4,24,27], a study performed by Gordon et al. [53] highlighted an accurate genomic prediction tool for determining antimicrobial susceptibilities for *Staphylococcus aureus*, comparable to antimicrobial susceptibility testing methods both in specificity and sensitivity. Even though antimicrobial resistance genes may provide a predictive value, genetic mechanisms responsible for a resistant phenotype are not easily predicted. Assessing virulence gene content from NGS data is technically feasible but again its clinical correlation is not at all clear. It is anticipated that NGS will be a preferred method for virulence and antibiotic resistance mechanisms studies, furthering diagnostics, therapeutics, and vaccine design capabilities [7,54].

5. Conclusion

The rapid advances in NGS technologies and capabilities has proven and further promises to be a game changer in diagnostic microbiology, significantly reducing the time from diagnosis to clinical treatment, while also broadening the repertoire of tools readily in-hand of clinical and public health microbiology laboratories, physicians and medical decision-makers.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] J.M. Heather, B. Chain, The sequence of sequencers: the history of sequencing DNA, Genomics 107 (1) (2016) 1–8.
[2] V. Sintchenko, E.C. Holmes, The role of pathogen genomics in assessing disease transmission, BMJ 350 (2015) h1314.
[3] D. MacCannell, Next generation sequencing in clinical and public health microbiology, Clin. Microbiol. Newsl. 38 (21) (2016) 169–176.
[4] X. Didelot, R. Bowden, D.J. Wilson, T.E. Peto, D.W. Crook, ransforming clinical microbiology with bacterial genome sequencing, Nat. Rev. Genet. Nat. Rev. Genet. 13 (9) (2012) 601–612, http://dx.doi.org/10.1038/nrg3226.
[5] C.A. Gilchrist, S.D. Turner, M.F. Riley, W.A. Petri, E.L. Hewlett, Whole-genome sequencing in outbreak analysis, Clin. Microbiol. Rev. 28 (3) (2015) 541–563.
[6] S.E. Levy, R.M. Myers, Advancements in next-generation sequencing, Annu. Rev. Genomics Hum. Genet. 17 (2016) 95–195.
[7] J.C. Kwong, N. McCallum, V. Sintchenko, B.P. Howden, Whole genome sequencing in clinical and public health microbiology, Pathol.-J. RCPA 47 (3) (2015) 199–210.
[8] M. Escalona, S. Rocha, D. Posada, A comparison of tools for the simulation of genomic next-generation sequencing data, Nat. Rev. Genet. 17 (8) (2016) 459–469.
[9] L. Becker, M. Steglich, S. Fuchs, G. Werner, U. Nübel, Comparison of six commercial kits to extract bacterial chromosome and plasmid DNA for MiSeq sequencing, Sci. Rep. 6 (2016) 28063, http://dx.doi.org/10.1038/srep28063.
[10] A. Mushtaq, Updates in tuberculosis diagnosis in the USA, Lancet Infect. Dis. 17 (2) (2017) 147–148.
[11] S.N. Naccache, K.S. Peggs, F.M. Mattes, R. Phadke, J.A. Garson, P. Grant, V. Gant, et al., Diagnosis of neuroinvasive astrovirus infection in an immunocompromised adult with encephalitis by unbiased next-generation sequencing, Clin. Infect. Dis. 60 (6) (2015) 919–923.
[12] K.J. Metzner, HIV whole-genome sequencing now: answering still-open questions, J. Clin. Microbiol. 54 (4) (2016) 834–835.
[13] M.G. Berg, J. Yamaguchi, E. Alessandri-Gradt, R.W. Tell, J.C. Plantier, C.A. Brennan, A pan-HIV strategy for complete genome sequencing, J. Clin. Microbiol. 54 (4) (2016) 868–882.
[14] R.J. Olsen, S.W. Long, J.M. Musser, Bacterial genomics in infectious disease and the clinical pathology laboratory, Arch. Pathol. Lab. Med. 136 (11) (2012) 1414–1422.
[15] Ø. Kommedal, M.T. Wilhelmsen, S. Skrede, R. Meisal, A. Jakovljev, P. Gaustad, Ø. Sæbø, et al., Massive parallel sequencing provides new perspectives on bacterial brain abscesses, J. Clin. Microbiol. 52 (6) (2014) 1990–1997 (JCM-00346).
[16] J.M. Rizzo, M.J. Buck, Key principles and clinical applications of next-generation DNA sequencing, Cancer Prev. Res. 5 (7) (2012) 887–900.
[17] H. Hasman, D. Saputra, T. Sicheritz-Ponten, O. Lund, C.A. Svendsen, N. Frimodt-Møller, F.M. Aarestrup, Rapid whole genome sequencing for the detection and characterization of microorganisms directly from clinical samples, J. Clin. Microbiol. 52 (1) (2014) 139–146 (JCM-02452).
[18] E.S. Snitkin, A.M. Zelazny, P.J. Thomas, F. Stock, D.K. Henderson, T.N. Palmore, NISC Comparative Sequencing Program, et al., Tracking a hospital outbreak of carbapenem-resistant Klebsiella pneumoniae with whole-genome sequencing, Sci. Transl. Med. 4 (148) (2012) (148ra116-148ra116).
[19] F.M. Aarestrup, E.W. Brown, C. Detter, P. Gerner-Smidt, M.W. Gilmour, D. Harmsen, K. Ijaz, et al., Integrating genome-based informatics to modernize global disease monitoring, information sharing, and response, Emerg. Infect. Dis. 18 (11) (2012) e1.
[20] G. Palacios, J. Druce, L. Du, T. Tran, C. Birch, T. Briese, J.F. Simons, et al., A new arenavirus in a cluster of fatal transplant-associated diseases, New Engl. J. Med. 358 (10) (2008) 991–998.
[21] J. Shendure, H. Ji, Next-generation DNA sequencing, Nat. Biotechnol. Nat. Biotechnol. 26 (10) (2008) 1135–1145, http://dx.doi.org/10.1038/nbt1486.
[22] E.J. Kaleta, A.E. Clark, A. Cherkaoui, V.H. Wysocki, E.L. Ingram, J. Schrenzel, D.M. Wolk, Comparative analysis of PCR–electrospray ionization/mass spectrometry (MS) and MALDI-TOF/MS for the identification of bacteria and yeast from positive blood culture bottles, Clin. Chem. 57 (7) (2011) 1057–1067.
[23] E.B. Popowitch, S.S. O'Neill, M.B. Miller, Comparison of the Biofire FilmArray RP, Genmark eSensor RVP, Luminex xTAG RVPv1, and Luminex xTAG RVP fast multiplex assays for detection of respiratory viruses, J. Clin. Microbiol. 51 (5) (2013) 1528–1533.
[24] C.U. Köser, L.J. Fraser, A. Ioannou, J. Becq, M.J. Ellington, N.A. Gormley, S.J. Peacock, et al., Rapid single-colony whole-genome sequencing of bacterial pathogens, J. Antimicrob. Chemother. 69 (2014) 1275–1281.
[25] E.C. Hayden, Pint-sized DNA sequencer impresses first users, Nature 521 (7550) (2015) 15–16.
[26] L.F. Westblade, A. van Belkum, A. Grundhoff, G.M. Weinstock, E.G. Pamer, M.J. Pallen, W.M. Dunne, Role of clinicogenomics in infectious disease diagnostics and public health microbiology, J. Clin. Microbiol. 54 (7) (2016) 1686–1693.
[27] N.D. Pecora, N. Li, M. Allard, C. Li, E. Albano, M. Delaney, L. Bry, et al., Genomically informed surveillance for carbapenem-resistant Enterobacteriaceae in a health care system, MBio 6 (4) (2015) e01030–15.
[28] G.H. Tyson, P.F. McDermott, C. Li, Y. Chen, D.A. Tadesse, S. Mukherjee, T.S. Edrington, et al., WGS accurately predicts antimicrobial resistance in Escherichia coli, J. Antimicrob. Chemother. 70 (10) (2015) 2763–2769 (dkv186).
[29] P.T. Clausen, E. Zankari, F.M. Aarestrup, O. Lund, Benchmarking of methods for identification of antimicrobial resistance genes in bacterial whole genome data, J. Antimicrob. Chemother. 71 (9) (2016) 2484–2488.
[30] M.J. Ellington, O. Ekelund, F.M. Aarestrup, R. Canton, M. Doumith, C. Giske, J. Iredell, et al., The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST Subcommittee, Clin. Microbiol. Infect. 23 (1) (2017) 2–22.
[31] M.D. Ritchie, E.R. Holzinger, R. Li, S.A. Pendergrass, D. Kim, Methods of integrating data to uncover genotype-phenotype interactions, Nat. Rev. Genet. 16 (2) (2015) 85–97.
[32] M. Gordon, E. Yakunin, L. Valinsky, V. Chalifa-Caspi, J. Moran-Gilad, A bioinformatics tool for ensuring the backwards compatibility of Legionella pneumophila typing in the genomic era, Clin. Microbiol. Infect. 23 (5) (2017) 306–310.
[33] P. McGann, J.L. Bunin, E. Snesrud, S. Singh, R. Maybank, A.C. Ong, S. Yamada, et al., Real time application of whole genome sequencing for outbreak investigation–what is an achievable turnaround time? Diagn. Microbiol. Infect. Dis. 85 (3) (2016) 277–282.
[34] B.R. Jackson, C. Tarr, E. Strain, K.A. Jackson, A. Conrad, H. Carleton, B.J. Silk, et al., Implementation of nationwide real-time whole-genome sequencing to enhance listeriosis outbreak detection and investigation, Clin. Infect. Dis. 63 (3) (2016) 380–386 (ciw242).
[35] S. Reuter, M.J. Ellington, E.J. Cartwright, C.U. Köser, M.E. Török, T. Gouliouris, S.J. Peacock, Rapid bacterial whole-genome sequencing to enhance diagnostic and public health microbiology, JAMA Intern. Med. JAMA Intern. Med. 173 (15) (2013) 1397, http://dx.doi.org/10.1001/jamainternmed.2013.7734.
[36] J. Moran-Gilad, K. Prior, E. Yakunin, T.G. Harrison, A. Underwood, T. Lazarovitch, I. Grotto, et al., Design and application of a core genome multilocus sequence typing scheme for investigation of Legionnaires' disease incidents, Euro Surveill. 20 (21186) (2015) 10–2807.
[37] S. Octavia, S.Z. Wu, S. Kaur, L. Valinsky, E. Marva, J. Moran-Gilad, R. Lan, Whole-genome sequencing and comparative genomic analysis of Bordetella pertussis isolates from the 2007–2008 epidemic in Israel, J. Infect. 74 (2) (2017) 204–207.
[38] S.J. Salipante, D.J. Sengupta, L.A. Cummings, T.A. Land, D.R. Hoogestraat, B.T. Cookson, Application of whole-genome sequencing for bacterial strain typing in molecular epidemiology, J. Clin. Microbiol. J. Clin. Microbiol. 53 (4) (2015) 1072–1079, http://dx.doi.org/10.1128/jcm.03385-14.
[39] J. Moran-Gilad, V. Sintchenko, S.K. Pedersen, W.J. Wolfgang, J. Pettengill, E. Strain, R.S. Hendriksen, Proficiency testing for bacterial whole genome sequencing: an end-user survey of current capabilities, requirements and priorities, BMC Infect. Dis. 15 (1) (2015) 174.
[40] J. Moran-Gilad, Whole genome sequencing (WGS) for food-borne pathogen surveillance and control?taking the pulse, Eurosurveillance 22 (2017) 23.
[41] M. Orsini, G. Cuccuru, P. Uva, G. Fotia, Bacterial genomic data analysis in the next-generation sequencing era, Data Mining Tech. Life Sci. (2016) 407–422.
[42] N.J. Loman, C. Constantinidou, J.Z. Chan, M. Halachev, M. Sergeant, C.W. Penn, M.J. Pallen, et al., High-throughput bacterial genome sequencing: an embarrassment of choice, a world of opportunity, Nat. Rev. Microbiol. Nat. Rev. Micro 10 (9) (2012) 599–606, http://dx.doi.org/10.1038/nrmicro2850.
[43] S. Shokralla, J.L. Spall, J.F. Gibson, M. Hajibabaei, Next-generation sequencing technologies for environmental DNA research, Mol. Ecol. 21 (8) (2012) 1794–1805.
[44] A.T. Vincent, N. Derome, B. Boyle, A.I. Culley, S.J. Charette, Next-generation sequencing (NGS) in the microbiological world: how to make the most of your money, J. Microbiol. Methods (2016), http://dx.doi.org/10.1016/j.mimet.2016.02.016.
[45] K.V. Voelkerding, S.A. Dames, J.D. Durtschi, Next-generation sequencing: from basic research to diagnostics, Clin. Chem. 55 (4) (2009) 641–658.
[46] R.S. Kaas, P. Leekitcharoenphon, F.M. Aarestrup, O. Lund, Solving the problem of comparing whole bacterial genomes across different sequencing platforms, PLoS One 9 (8) (2014) e104984.
[47] S. Koren, A.M. Phillippy, One chromosome: one contig: complete microbial genomes from long-read sequencing and assembly, Curr. Opin. Microbiol. 23 (2015) 110–120.
[48] N.L. Hiller, A. Ahmed, E. Powell, D.P. Martin, R. Eutsey, J. Earl, R. Sampath, et al., Generation of genic diversity among Streptococcus pneumoniae strains via horizontal gene transfer during a chronic polyclonal pediatric infection, PLoS Pathog. 6 (9) (2010) e1001108.
[49] T. Seemann, Snippy: Fast Bacterial Variant Calling from NGS Reads, (2015) https://github.com/tseemann/snippy.
[50] M. Pérez-Losada, P. Cabezas, E. Castro-Nallar, K.A. Crandall, Pathogen typing in the genomics era: MLST and the future of molecular epidemiolog, Infect. Genet. Evol. 16 (2013) 38–53.
[51] M. Silva, M.P. Machado, M. Rossi, J. Moran-Gilad, S. Santos, M. Ramirez, J.A. Carrico, chewBBACA: A complete suite for gene-by-gene schema creation and strain identification, bioRxiv (2017) 173146.
[52] R.R. Singh, R. Luthra, M.J. Routbort, K.P. Patel, L.J. Medeiros, Implementation of

next generation sequencing in clinical molecular diagnostic laboratories: advantages, challenges and potential, Expert Rev. Precis. Med. Drug Dev. 1 (1) (2016) 109–120.
[53] N.C. Gordon, J.R. Price, K. Cole, R. Everitt, M. Morgan, J. Finney, M.J. Llewelyn, Prediction of Staphylococcus aureus antimicrobial resistance by whole-genome sequencing, J. Clin. Microbiol. 52 (4) (2014) 1182–1191.
[54] C.P. Brouwer, T.D. Vu, M. Zhou, G. Cardinali, M.M. Welling, N. van de Wiele, V. Robert, Current opportunities and challenges of next generation sequencing (NGS) of DNA; determining health and diseases, Br. Biotechnol. J. 13 (4) (2016).